Article

Tracing oncogene-driven remodelling of the intestinal stem cell niche

https://doi.org/10.1038/s41586-021-03605-0

Received: 22 October 2020

Accepted: 30 April 2021

Published online: 02 June 2021

Check for updates

Min Kyu Yum[1,2,10], Seungmin Han[1,2,10], Juergen Fink[2], Szu-Hsien Sam Wu[3,4], Catherine Dabrowska[2,5], Teodora Trendafilova[2], Roxana Mustata[2], Lemonia Chatzeli[1,2], Roberta Azzarelli[2,6], Irina Pshenichnaya[2], Eunmin Lee[7], Frances England[2,5], Jong Kyoung Kim[7], Daniel E. Stange[8], Anna Philpott[2,6], Joo-Hyeon Lee[2,5], Bon-Kyoung Koo[2,3,11]✉ & Benjamin D. Simons[1,2,9,11]✉

Interactions between tumour cells and the surrounding microenvironment contribute to tumour progression, metastasis and recurrence[1–3]. Although mosaic analyses in *Drosophila* have advanced our understanding of such interactions[4,5], it has been difficult to engineer parallel approaches in vertebrates. Here we present an oncogene-associated, multicolour reporter mouse model—the Red2Onco system—that allows differential tracing of mutant and wild-type cells in the same tissue. By applying this system to the small intestine, we show that oncogene-expressing mutant crypts alter the cellular organization of neighbouring wild-type crypts, thereby driving accelerated clonal drift. Crypts that express oncogenic KRAS or PI3K secrete BMP ligands that suppress local stem cell activity, while changes in PDGFR^{lo}CD81^{+} stromal cells induced by crypts with oncogenic PI3K alter the WNT signalling environment. Together, these results show how oncogene-driven paracrine remodelling creates a niche environment that is detrimental to the maintenance of wild-type tissue, promoting field transformation dominated by oncogenic clones.

The tumour microenvironment (TME) constitutes a complex ecosystem that comprises mutant and wild-type (WT) cells, as well as endothelial, immune and mesenchymal cells[1,2]. From the earliest phase of tumour initiation, the TME adapts to signals from tumour cells[3]. However, little is known about how these changes affect non-malignant cells. Much of our understanding of this topic has been informed by studies in *Drosophila*, in which WT cells eliminate less 'fit' cells through 'cell competition'[4,6]. Conversely, tumour cells bearing mutations in genes such as *Myc*, *Ras* or *Apc* can become 'super-competitors' that eliminate their WT neighbours[5]. In mammals, increasing emphasis has been placed on cell competition during tumour development[7], with oncogene-driven changes in both mutant cells and the TME[8] having the potential to influence neighbouring WT cells[9,10] (Fig. 1a).

Multicolour reporters with oncogenes

Cell labelling strategies have been devised to study fate changes during tumour development[11,12]. Although such methods provide insight into factors that drive field cancerization, it has been difficult to assess how neighbouring WT cells react to, and influence, tumour growth. To study cross-talk between mutant and WT cells, previous studies have used mosaic labelling approaches[13,14] to study gain and loss of function. However, challenges in the design of these approaches limit their potential. Here, to circumvent these difficulties, we adapted the multicolour *Confetti* reporter line[15], based on the *Brainbow-2.1* cassette, to include an additional cDNA cassette following the *tdimer2* red fluorescent protein (*RFP*) cDNA (Fig. 1b). This inducible Red2Onco system allows an oncogene to be co-expressed specifically in RFP^{+} clones. Upon Cre activation, the *R26R-Red2Onco* construct stochastically recombines to express one of four fluorescent proteins, and only RFP^{+} clones co-express the oncogenic cDNA (Fig. 1b, c, Supplementary Video 1).

To demonstrate the utility of the Red2Onco system, we developed lines for RFP-specific co-expression of the intracellular domain (ICD) of *Notch1*, *Kras^{G12D}* or *PIK3CA^{H1047R}*. We first confirmed that RFP expression overlapped with gene expression (NOTCH1-ICD) or downstream activation (p-ERK for KRAS(G12D); p-AKT for PIK3CA(H1047R)) in intestine (Extended Data Fig. 1a, b). Consistent with their role in driving hyperplasia[16,17], we found that RFP^{+} cells that co-expressed *Kras^{G12D}* or *PIK3CA^{H1047R}* showed increased proliferation (Extended Data Fig. 1c). Similarly, as expected[18], *Notch1^{ICD}* expression blocked differentiation of secretory cells (Extended Data Fig. 1d, e). We also confirmed the viability of the Red2Onco system in other organs (Extended Data Fig. 1f, g).

Mutant intestinal clone expansion

In mammals, the small intestine is composed of crypts and villi. All absorptive and secretory cell lineages are generated by

[1]Wellcome Trust–Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK. [2]Wellcome Trust–Medical Research Council Cambridge Stem Cell Institute, Jeffrey Cheah Biomedical Centre, University of Cambridge, Cambridge, UK. [3]Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna Biocenter (VBC), Vienna, Austria. [4]Vienna BioCenter PhD Program, Doctoral School at the University of Vienna and Medical University of Vienna, Vienna, Austria. [5]Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. [6]Department of Oncology, University of Cambridge, Hutchison–MRC Research Centre, Cambridge, UK. [7]Department of New Biology, DGIST, Daegu, Republic of Korea. [8]Department of Visceral, Thoracic and Vascular Surgery, University Hospital Carl Gustav Carus, Medical Faculty, Technische Universität Dresden, Dresden, Germany. [9]Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Cambridge, UK. [10]These authors contributed equally: Min Kyu Yum, Seungmin Han. [11]These authors jointly supervised this work: Bon-Kyoung Koo, Benjamin D. Simons. ✉e-mail: bonkyoung.koo@imba.oeaw.ac.at; bds10@cam.ac.uk

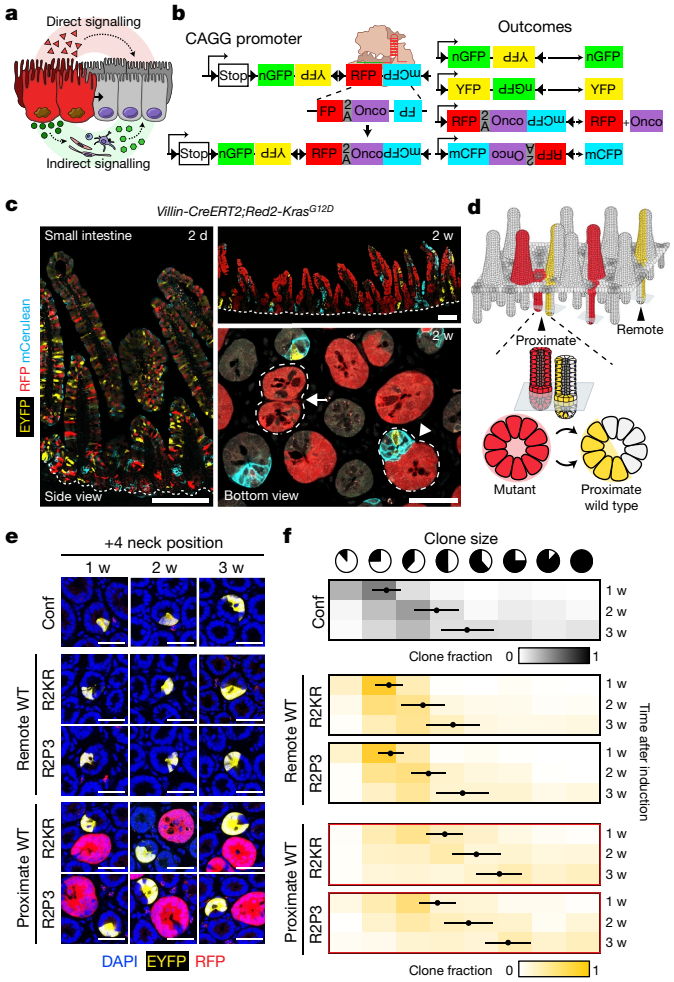


Fig. 1 | Red2Onco system: an oncogene-associated multicolour reporter. **a**, Schematic showing possible routes for cross-talk between mutant and neighbouring WT cells. **b**, *Red2Onco* knock-in strategy. The 2A peptide sequence and oncogene cDNA (*Kras*^{G12D}, *PIK3CA*^{H1047R} or *Notch1*^{ICD}) were cloned in-frame downstream of the *RFP* cDNA in the *R26R-Confetti* cassette[15], which encodes four fluorescent proteins. **c**, Representative images from sections (left, top right) or whole mounts (bottom right) of small intestine from *Villin-CreERT2;Red2-Kras*^{G12D} mice 2 days (d) or 2 weeks (w) after tamoxifen administration (representative of three independent experiments). White dashed line, mucosal lining. Bottom right shows crypt fission (arrow) and fusion events (arrowhead). **d**, Schematic illustration of a WT (YFP$^+$) clone in proximity to a fixed (monoclonal) mutant (RFP$^+$) crypt. Clone sizes quantified as defined in Extended Data Fig. 2g. **e**, Representative confocal images of WT (YFP$^+$) clones remote from, or proximate to, fixed mutant (RFP$^+$) crypts in whole-mount small intestine from *Villin-CreERT2;R26R-Confetti* (Conf) or *Red2Onco* mice (R2KR, *Red2-Kras*^{G12D}; R2P3, *Red2-PIK3CA*^{H1047R}) at 1, 2 and 3 w after tamoxifen administration. Images are representative of tissues quantified in **f**. **f**, Heat maps indicate relative clone fractions of indicated sizes (columns) at various times after induction (rows). Points and error bars denote mean ± s.e.m. $n = 6$ mice per group and time point. $n = 103, 80, 169$ clones (Conf); 93, 129, 178 (remote WT R2KR); 106, 94, 187 (remote WT R2P3); 90, 99, 241 (proximate WT R2KR); and 64, 137, 88 (proximate WT R2P3) at 1, 2, 3 w after induction, respectively. Remote clones are separated by more than two crypt diameters from mutant crypts. Proximate clones are neighbouring fixed mutant crypts. *R26R-Confetti* and remote WT data from Extended Data Fig. 2i are reproduced for comparison. Scale bars: 200 µm (**c**, left); 50 µm (**c**, right, **e**).

intestinal stem cells (ISCs) that reside at the crypt base. As ISCs divide, neighbouring cells become displaced from the niche and enter into a differentiation programme. As a result, ISC-derived clones stochastically expand and contract, leading to 'neutral drift' of clones until they are lost or the crypt becomes monoclonal[19,20].

Alongside WNT, mutations in the RAS–MAPK and PI3K–AKT pathways represent key drivers of colorectal cancer[11,21,22]. Unlike WNT, activation of RAS signalling does not lead to obvious changes in tissue architecture, leading to the assumption that mutations in WNT signalling may represent the first oncogenic hit. However, the discovery that KRAS mutations are abundant in the ageing human intestinal epithelium[10,23], as well as in patients with Crohn's disease[24,25], raised the question of whether mutations in RAS–MAPK or PI3K–AKT signalling components could also function as drivers of early mutant clone expansion. To investigate the effects of altered gene expression on clonal competition, we performed lineage tracing in *Villin-CreERT2;R26R-Red2Onco* mice using a clonal dosage of tamoxifen (0.2 mg per 20 g body weight, at which each crypt on average hosts less than one labelled, ISC-derived clone) (Extended Data Fig. 2a–f). Seven days after Cre induction, mutant clones already showed evidence of biased drift (Extended Data Fig. 2g–i). Although both mutant and WT clones showed a broad distribution of sizes, both the average size and the frequency of monoclonal (fixed) crypts increased more steeply in mutant than in WT clones (Extended Data Fig. 2j, k). These findings echoed the results of previous studies based on clonal activation of KRAS using a conventional reporter construct[26,27]. Quantitative comparison of isolated, WT (non-red) clones with the *Confetti* control (Extended Data Fig. 3a–c) indicated that the modified *Confetti* construct had no systemic effects on clone dynamics.

To resolve the dynamics of mutant clone expansion, we invoked a modelling-based approach, the validity of which has been tested previously[19,28]. Within this framework, crypts are modelled as a chain of N 'effective' ISCs that line the circumference of the crypt base. ISC loss (at rate λ) is correlated with the duplication of a neighbour, leading to neutral drift of clone size (Supplementary Theory). Applied to the *Confetti* control, with an effective stem cell number for the proximal small intestine of $N = 5$ (ref. [28]), a quantitative fit to the clone data obtained a loss-replacement rate of $\lambda = 0.9$ per week, similar to previous findings (Extended Data Fig. 3a). Following oncogene expression, mutant cells experience a survival advantage, leading to non-neutral competition. The resulting clone dynamics can be mapped to a refined model in which the loss-replacement rate acquires a bias $0 \le \delta \le 1$, with mutant ISCs replacing their WT neighbours at rate $\lambda(1+\delta)$, whereas WT ISCs replace mutant neighbours at rate $\lambda(1-\delta)$. Taking $N = 5$, the biased drift model predicted accurately the dynamics of clone sizes with $\lambda = 2.4$ per week and $\delta = 0.71$ for *Kras*^{G12D} (similar to previous estimates[26,27]), $\lambda = 1.9$ per week and $\delta = 0.64$ for *PIK3CA*^{H1047R}, and $\lambda = 1.1$ per week and $\delta = 0.36$ for *Notch1*^{ICD} (Extended Data Fig. 3d–f).

Active competition through apoptotic elimination of less-fit cells has been well documented in *Drosophila* and mouse epiblasts[4]. However, immunodetection of cleaved caspase-3 in crypts of *R26R-Red2Onco* mice showed no evidence of increased apoptosis (Extended Data Fig. 3g). By contrast, a two-hour pulse of 5-ethynyl-2′-deoxyuridine (EdU) showed an increased number of EdU$^+$ crypt base columnar cells in mutant clones in all three *Red2Onco* models (Extended Data Fig. 3h, i).

Oncogene-induced effects on neighbours

To study whether mutant clones influence the fates of WT clones in neighbouring crypts, we performed lineage tracing in *Villin-CreERT2;R26R-Red2Onco* mice induced at a near-clonal dosage of tamoxifen (2 mg per 20 g body weight), when more than 30% of crypts contained labelled cells. Unexpectedly, two weeks after induction, the frequency of fixed WT crypts in *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice had increased compared to controls (Extended Data Fig. 4a, b), which suggests that mutant cells mediate changes in the dynamics of WT cells. To dissect this phenotype, we used a clonal dosage of tamoxifen (0.2 mg per 20 g body weight) to compare the dynamics of WT clones in crypts that were either neighbouring (proximate) or remote (more than two crypt

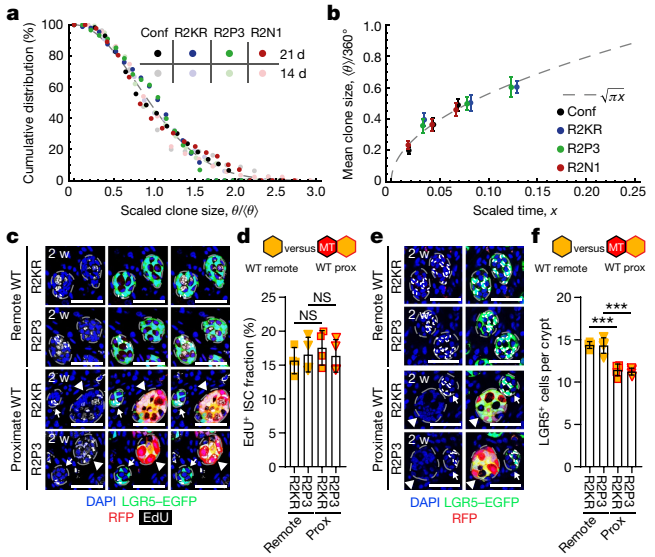


Fig. 2 | Reduced effective stem cell number leads to accelerated drift of WT clones in crypts that neighbour mutant crypts. a, Cumulative size distribution of WT (YFP$^+$) clones in crypts neighbouring crypts monoclonal for given mutant or *Confetti* control against the angular clone size θ rescaled by the average $\langle\theta\rangle$. Points show data from two time points and dashed line denotes scaling function, $\exp[-\pi(\theta/\langle\theta\rangle)^2/4]\times 100\%$, predicted by neutral drift model (see main text and Supplementary Theory). **b**, Corresponding average clone size $\langle\theta\rangle/360°$ as a function of time after induction, scaled by effective drift rates $x = t\lambda/N^2$ obtained from a fit to predicted square root dependence (dashed line) (Supplementary Theory). In **a**, **b**, experimental data as in Fig. 1f; for R2N1-proximate clones, $n = 6$ mice per group and time point, and 87, 128 clones were scored at 14, 21 d after induction. **c**, **d**, Representative confocal images (**c**) and quantification (**d**) of EdU$^+$ proliferating LGR5$^+$ stem cells in WT crypts neighbouring mutant (MT) crypts. Dashed white outlines, proliferating stem cells in WT crypts. $n = 4$ mice per group, 100 crypts analysed per mouse. **e**, **f**, Representative confocal images (**e**) and quantification (**f**) of LGR5–EGFP$^+$ stem cells. Dashed white outlines, EGFP$^+$ cells in WT crypts. $n = 5$ mice per group, 100 crypts analysed per mouse. **c**, **e**, Whole-mount small intestine imaged from *Lgr5-EGFP-IRES-CreERT2;Red2Onco* mice 2 w after tamoxifen administration. Arrows, proximate WT crypts; arrowheads, fixed mutant crypts. Dashed grey outlines, crypt borders. ***$P < 0.0001$; unpaired two-tailed *t*-test (**d**, **f**). NS, not significant. Data are mean ± s.e.m. (**b**) or mean ± s.d. (**d**, **f**) from biological replicates. For exact *P* values, see Source Data. Scale bars, 50 µm.

diameters) from crypts containing a mutant clone (Fig. 1d). Notably, WT clones that were proximate to mutant crypts showed accelerated drift in *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice, greatly exceeding that in remote WT crypts or *Confetti* controls (Fig. 1e, f). These observations were corroborated by quantification of clone size, which showed a corresponding increase in average clone size and fixation rate (Extended Data Fig. 4c, d). Furthermore, when segregated according to whether clones were oriented towards (inner) or away (outer) from a neighbouring mutant crypt, the clone size distributions were comparable, suggesting that accelerated dynamics take place in the context of neutral competition (Extended Data Fig. 4e–g).

To make a further quantitative assessment of WT clones in crypts that neighbour mutant crypts we noted that, before fixation, neutral clone dynamics are predicted to enter a scaling regime in which the size distribution depends only on the composite drift rate λ/N^2 (Supplementary Theory)[19]. Comparison of the clone size data showed convergence onto the predicted scaling form in both *Confetti* control and mutant conditions (Fig. 2a). From a fit to the average clone size, we found a drift rate that was 1.9 times larger in clones that were proximate to mutant crypts compared to *Confetti* controls (Fig. 2b, Extended Data Fig. 4h, Supplementary Theory).

To resolve the origin of accelerated drift, we considered whether the rate of ISC loss or replacement had increased (Extended Data Fig. 4i). However, measurements of the ISC proliferation rate based on EdU incorporation showed similar values (Fig. 2c, d), independent of the proximity of WT clones to mutant crypts, which suggests that the loss–replacement rate had not changed. We then estimated the number of ISCs on the basis of *Lgr5*-driven EGFP expression. When the majority of mutant crypts were already fixed, we found that the number of LGR5–EGFP$^+$ cells was 20% lower in WT crypts neighbouring mutant crypts than in remote WT crypts (Fig. 2e, f). Similar results were found when we used OLFM4 expression (Extended Data Fig. 4j). These findings suggest that accelerated clonal drift is associated with a reduction in the number of stem cells[29].

An observed reduction in the physical size of WT crypts neighbouring mutant crypts (Extended Data Fig. 5a–c) led us to investigate whether the change in ISC number is associated with deformation of surrounding tissue by the increase in size of mutant crypts. However, from measurements of the circularity of WT crypts, we found no evidence for an effect of local mechanical strain (Extended Data Fig. 5d, e). We then reasoned that, if the effect were mediated by paracrine factors from mutant cells, it should scale with the multiplicity of neighbouring mutant crypts (Extended Data Fig. 5f). Consistently, when proximate to multiple mutant crypts, WT clones from *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice showed increased rates of clonal drift and fixation alongside a further reduction in ISC number and crypt size (Extended Data Fig. 5g–l).

Finally, we questioned whether mutation-induced changes promote field expansion. Notably, in contrast to *Confetti* and *Red2-Notch1*^{ICD} mice, the fraction of mutant crypts increased steadily in *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice at the expense of WT crypts (Extended Data Fig. 5m, n). We then used the abundance of 'cojoined' crypts to estimate the relative frequency of crypt fission and fusion (Extended Data Fig. 5o). Following a previous study[30], we found that crypts mutant for *Kras*^{G12D} and *PIK3CA*^{H1047R} showed increased rates of fission and fusion compared to *Confetti* controls (Extended Data Fig. 5p), which suggests that the deleterious effects of mutant crypts on neighbouring WT crypts contribute to field transformation.

Comparative single-cell analysis

To define the mechanisms that mediate cross-talk, we exploited the design of the Red2Onco system for comparative single-cell transcriptomics (Fig. 3a). *Villin-CreERT2;R26R-Red2Onco* mice (*Kras*^{G12D} or *PIK3CA*^{H1047R}) were induced using a mosaic dosage of tamoxifen (4 mg per 20 g body weight). Following dissociation of intestine, we used fluorescence activated cell sorting (FACS) to sort individual mutant (RFP$^+$) and WT (YFP$^+$) epithelial cells, and profiled them alongside neighbouring stromal tissue (Extended Data Fig. 6a). By combining these results with single-cell RNA sequencing (scRNA-seq) data from control *Confetti* mice, we found evidence for mutation-driven changes in gene expression in epithelial, stromal and immune cells (Fig. 3a, Extended Data Fig. 6b; see Methods). Epithelial cells from *Red2-Kras*^{G12D}, *Red2-PIK3CA*^{H1047R} and *Confetti* control mice were clustered on the basis of marker-gene expression (Fig. 3b, Extended Data Fig. 6c–f, Supplementary Table 1). All major cell types were detected in ratios comparable to those reported previously[31–34]. Despite oncogene-induced transcriptional changes, cells from all three conditions overlapped within distinct clusters (Extended Data Fig. 6g; see Methods).

To explore the influence of oncogene activation on lineage specification, we examined the proportions of each cell type in *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice compared to *Confetti* controls. Notably, in both models, the proportion of mutant LGR5$^+$ stem cells was significantly decreased, whereas the proportion of mutant goblet cells was increased (Extended Data Fig. 6h). In *Red2-PIK3CA*^{H1047R} mice, there was also an increase in the proportion of mutant enterocytes. As with mutant cells, the WT population contained a lower proportion of stem

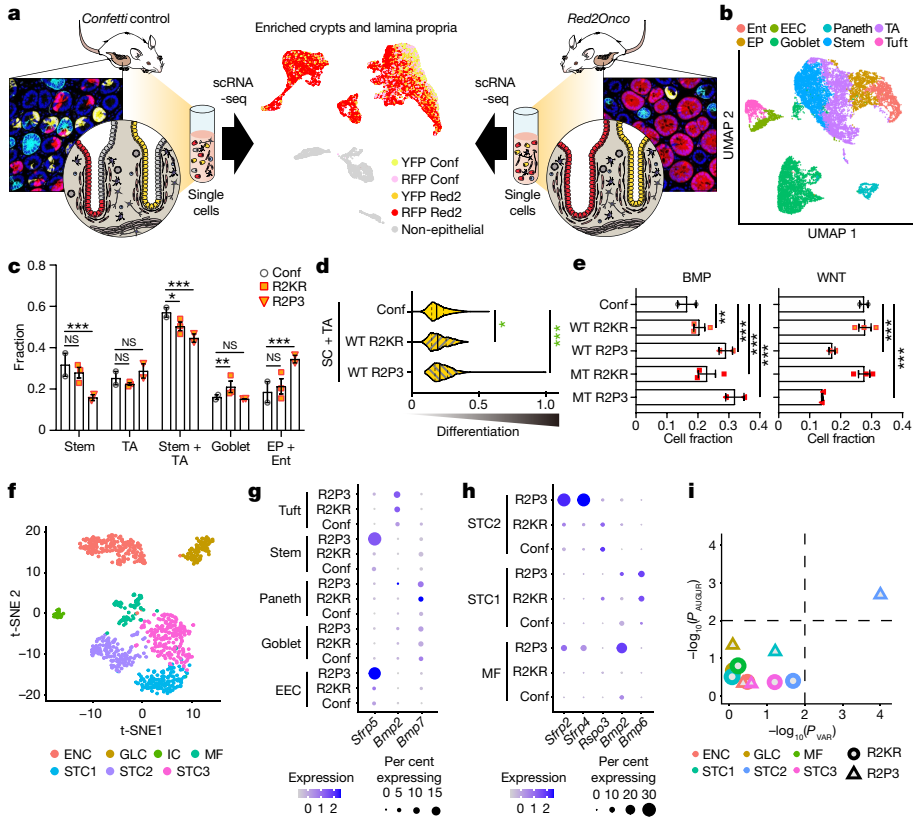


Fig. 3 | Comparative single-cell analysis identifies oncogene-driven niche changes. a, Schematic of comparative single-cell analysis. Mutant and WT epithelial, mesenchymal and immune cells are isolated from small intestine. By comparing profiles across models, the effects of oncogene expression on mutant and WT epithelial cells and the surrounding environment can be resolved. **b**, Uniform manifold approximation and projection (UMAP) plot showing clustering of epithelial cells based on marker expression: stem cells, transit-amplifying (TA) cells, enterocyte progenitors (EP), enterocytes (Ent), enteroendocrine cells (EEC), goblet cells, Paneth cells and tuft cells. **c**, Fractions of WT epithelial cell types in *Red2Onco* and *Confetti* mice. **d**, Priming scores of WT (YFP$^+$) stem and transit-amplifying cells towards enterocyte lineages in *Red2Onco* and *Confetti* mice. Centre line, 50th percentile; dashed lines, 25th and 75th percentiles. Green asterisks, higher in *Red2Onco* than *Confetti*. **e**, Fractions of 'active' cells with high enrichment scores for BMP (left) and WNT (right) pathways in mutant and WT epithelial cells from *Red2Onco* and *Confetti* mice. **f**, *t*-Distributed stochastic neighbour embedding (t-SNE) representing mesenchymal cell clusters. ENC, endothelial cell; GLC, glial cell; IC, interstitial cell of Cajal; MF, myofibroblast; STC1/2/3: stromal cell 1/2/3. **g**, **h**, Dot plots showing average expression of secreted factors known to modulate BMP and WNT signalling in mutant epithelial cells (**g**) and mesenchymal cells (**h**) from *Red2Onco* and WT cells from *Confetti* mice. Dot size denotes percentage of cells expressing given gene, colour denotes average expression across all cells of that type. **i**, Degree of transcriptomic change for each mesenchymal cell type in *Red2Onco* mice (see Methods). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$; two-sided likelihood ratio test (**c**, **e**), two-sided Kolmogorov–Smirnov test (**d**). NS, not significant. Data are mean ± s.e.m. (**c**, **e**) from biological replicates ($n = 2$ for Conf, R2P3; $n = 3$ for R2KR). For exact *P* values, see Source Data.

and transit-amplifying cells in both *Red2Onco* models (Fig. 3c, Extended Data Fig. 6i), consistent with a reduction in ISCs (Fig. 2). The reduction in ISC fraction was confirmed by immunohistochemistry and FACS (Extended Data Fig. 6j–n).

We then investigated whether changes in the fractions of stem and progenitor cells were accompanied by changes in their molecular characteristics. We first defined lineage-specific genes as molecular signatures of fate priming using scRNA-seq data from a previous study[32] (Supplementary Table 2). We then calculated the priming score towards differentiation for individual WT (YFP$^+$) and mutant (RFP$^+$) stem and transit-amplifying cells (see Methods). In *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice, both WT and mutant stem and transit-amplifying cells showed enhanced differentiation towards secretory and absorptive lineages, respectively (Fig. 3d, Extended Data Fig. 7a, b); this result was confirmed by marker-based quantitative PCR with reverse transcription (qRT–PCR) analysis and immunostaining or in situ hybridization (Extended Data Fig. 7c–j). These results suggested that stem cell priming towards differentiation is a potential driver of the reduced stem cell fraction in mutant and neighbouring WT crypts in both *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice.

To study the mechanisms that underlie biased differentiation, we then assessed the activities of major signalling pathways[2,35]. Gene set enrichment analysis (GSEA) (Supplementary Table 2; see Methods) revealed higher BMP signalling activity in both mutant and WT cells of *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice compared to *Confetti* controls, whereas WNT signalling activity was lower in both mutant and WT cells of *Red2-PIK3CA*^{H1047R} mice (Fig. 3e, Extended Data Fig. 7k). By contrast, NOTCH signalling activity was reduced in mutant cells, but not in WT cells, in both *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice (Extended Data Fig. 7l, m). On the basis of these findings, we focused on changes in BMP and WNT signalling to trace the basis of inter-crypt cellular cross-talk.

Mutant clones drive niche remodelling

We then examined direct (epithelial cell-to-cell) and indirect (via mesenchymal or immune cell) routes of communication that could affect the BMP and WNT pathways. Mesenchymal and immune cells were clustered on the basis of expression signatures (Fig. 3f, Extended Data Fig. 8a–c, Supplementary Table 3). Three clusters (labelled STC for stromal cell) showed marked expression of secreted factors for intestinal

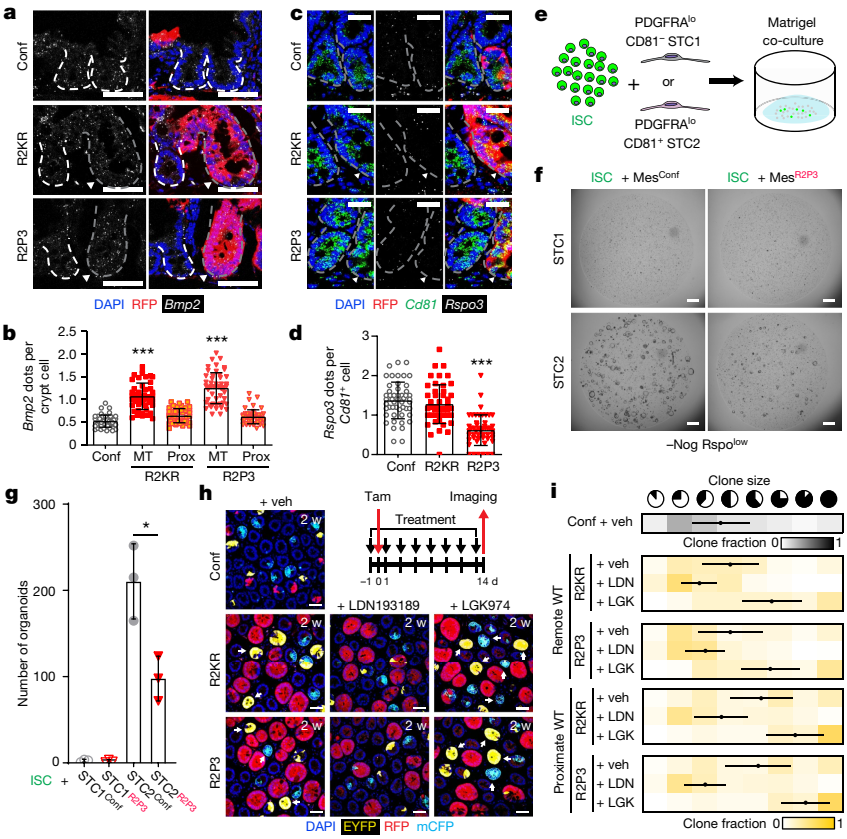


Fig. 4 | Functional validation of oncogene-driven niche remodelling. **a**, **b**, Representative in situ hybridization images (**a**) and quantification (**b**) of *Bmp2* expression. White dashed outline, WT crypts; grey dashed outline, mutant crypts. Prox, proximate WT crypts. For each group, 50 crypts analysed from $n = 3$ mice. **c**, **d**, Representative multiplexed in situ hybridization images (**c**) and quantification (**d**) of *Rspo3* expression in *Cd81*$^+$ cells. Grey dashed outlines, crypts. For each group, 50 crypt pairs analysed from $n = 3$ mice. **e**, Sorted LGR5$^+$ cells were cultured with either PDGFRA^{lo}CD81$^-$ STC1 cells or PDGFRA^{lo}CD81$^+$ STC2 cells from *Confetti* or *Red2-PIK3CA*^{H1047R} intestine. **f**, **g**, Representative bright-field images (**f**) and quantification (**g**) of intestinal organoids formed after 4 d of co-culturing. $n = 3$ independent experiments. Mes^{Conf}, Mes^{R2P3}: mesenchymal cells from *Confetti*, *Red2-PIK3CA*^{H1047R} mice. **h**, **i**, Representative confocal images (**h**) and heat map distribution of clone fractions (**i**) of whole-mount small intestine from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice at 2 w after induction. LDN193189 (LDN, BMP type I receptor blocking agent), LGK974 (LGK, porcupine inhibitor) or vehicle (veh) was administered following the dosing regimen shown (inset). White arrows, fixed WT crypts. Data are mean ± s.e.m. $n = 3$ mice per group. We scored 212 clones for Conf + veh; 85, 78, 96 clones for remote WT R2KR (in veh, LDN, LGK conditions, respectively); 91, 90, 113 for remote WT R2P3; 130, 108, 75 for proximate WT R2KR; 132, 81, 65 for proximate WT R2P3. **a**, **c**, Sections of small intestine from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice 2 w after induction. White arrowheads, fixed mutant crypts. $^*P < 0.05$, $^{***}P < 0.0001$; one-way ANOVA with Games-Howell's multiple comparisons test (**b**, **d**) and unpaired two-tailed *t*-test (**g**) from biological replicates. Data are mean ± s.d. (**b**, **d**, **g**). For exact *P* values, see Source Data. Scale bars: 50 µm (**a**, **h**), 25 µm (**c**), 500 µm (**f**).

stromal cells[36,37] (Extended Data Fig. 8d). Notably, *Bmp4* was highly expressed in cluster STC1, whereas STC2 expressed genes encoding WNT pathway-modulating factors such as *Rspo3*, as well as the BMP antagonist *Grem1*. We then searched for changes in the expression of secreted factors (Supplementary Table 4) that bind to receptors (Extended Data Fig. 8e, f) that modulate the BMP and WNT pathways. We noticed an increase in the expression of BMP ligand genes, such as *Bmp2* and *Bmp7*, that are found mainly in mutant secretory cells in *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice (Fig. 3g). We also detected an increase in the expression of WNT pathway antagonist genes, such as *Sfrp2* and *Sfrp4*, as well as a decrease in the expression of WNT pathway agonist genes, such as *Rspo3*, in cluster STC2 from *Red2-PIK3CA*^{H1047R} mice (Fig. 3h).

Changes in ligand expression in mesenchymal cells raised the possibility that the niche environment may become altered following oncogene activation. Following quantification of transcriptional changes in mesenchymal and immune cells (see Methods), we found that the STC2 population in *Red2-PIK3CA*^{H1047R} mice experienced significant alterations compared to other cell types (Fig. 3i, Extended Data Fig. 8g–l). In addition, this STC2 population showed increased levels of genes that regulate the extracellular matrix, such as *Fn1*, *Mmp2* and collagens (*Col6a2* and *Col5a2*) (Extended Data Fig. 8m). Together, these results showed that epithelial oncogene expression of either mutant KRAS or PI3K induces a BMP-rich signalling environment and, for PI3K, a WNT-poor environment (Extended Data Fig. 8n).

Blocking oncogenic niche remodelling

To validate these findings, we performed in situ hybridization for *Axin2* and *Id1*, which are downstream target genes of the WNT and BMP pathways, respectively. As expected from our single-cell analysis, we observed decreased expression of *Axin2* in both mutant and neighbouring WT crypts in *Red2-PIK3CA*^{H1047R} mice, but not in *Red2-Kras*^{G12D} mice (Extended Data Fig. 9a,b). *Id1* expression was increased in mutant and neighbouring WT crypts in both *Red2Onco* models (Extended Data Fig. 9c, d). These results were consistent with qRT–PCR data (Extended Data Fig. 9e, f), confirming that WNT activity was decreased only in *Red2-PIK3CA*^{H1047R} mice, whereas BMP activity was increased in both *Red2Onco* models. These findings were further supported by in vitro organoid assays, in which changes in BMP and WNT signalling mimicked the effects of different niche conditions (Extended Data Fig. 9g–m).

In situ hybridization, qRT–PCR and organoid assays also confirmed increased expression of *Bmp2* in mutant epithelial cells of both *Red2Onco* models (Fig. 4a, b, Extended Data Fig. 9n), as well as the functionality of epithelial BMP ligands in vitro (Extended Data Fig. 9o–r).

PDGFRA^{lo}CD81^{+} stromal cells are an important source of R-spondin and Gremlin-1—a WNT agonist and a BMP antagonist, respectively[38]. The STC2 population in our study, which showed high *Rspo3* expression (Extended Data Fig. 8d), expressed high levels of *Grem1* and *Cd81* compared to other mesenchymal cells; consistent with previous findings[38], these cells were located adjacent to the crypt base (Fig. 4c, Extended Data Fig. 10a, b). Consistent with scRNA-seq data (Fig. 3h), in situ hybridization and qRT–PCR analysis showed that STC2 cells expressed higher levels of *Sfrp2* and *Sfrp4*, and lower levels of *Rspo3*, in *Red2-PIK3CA^{H1047R}* mice than in *Red2-Kras^{G12D}* mice and *Confetti* controls (Fig. 4c, d, Extended Data Fig. 10b–h). To test whether oncogene-driven changes in STC2 cells compromise niche function, we co-cultured LGR5–EGFP^{+} ISCs with PDGFRA^{lo}CD81^{-} STC1 cells or PDGFRA^{lo}CD81^{+} STC2 cells isolated from WT and *Red2-PIK3CA^{H1047R}* mice (Fig. 4e). As expected, the STC2 population from the WT mice supported organoid formation in growth factor-deprived conditions (lacking Noggin and with low levels of R-spondin; –Nog Rspo^{low}), whereas the STC2 population sorted from the intestine of *Red2-PIK3CA^{H1047R}* mice showed a significantly reduced capacity to support organoid formation (Fig. 4f, g).

To further test our hypothesis that BMP and WNT pathways are involved in cross-talk, we used inhibitors that block either reception of BMP ligands (LDN193189, BMP type I receptor blocking agent) or secretion of WNT ligands (LGK974, porcupine inhibitor)[29]. Administration of LDN193189 and LGK974 to *Red2Onco* mice suppressed BMP and WNT pathway signalling, respectively (Extended Data Fig. 11a, b). In addition, LDN193189 rescued the rate of clonal drift of WT clones neighbouring mutant crypts, whereas LGK974 accelerated drift still further (Fig. 4h, i, Extended Data Fig. 11c). Notably, fixation of mutant crypts was delayed by treatment with LDN193189, suggesting that unlabelled WT cells in the same crypt compete more efficiently against mutant clones when the BMP effect is abolished (Extended Data Fig. 11d, e). We also confirmed that the phenotypes from both Red2Onco systems are largely recapitulated in *LSL-Kras^{G12D}* and *PIK3CA$^{Lat\text{-}H1047R}$* mice, which express the respective oncogenes under the corresponding endogenous locus (Extended Data Fig. 11f–m).

Conclusions

Local cell displacement provides a mechanism for eliminating 'loser' cells in epithelial tissues. In this case, direct cell-to-cell competition triggers the differentiation or apoptosis of adjacent WT stem cells[39,40]. Our findings show that, through long-range paracrine signals, mutant cells may also affect WT cells in neighbouring domains, either directly or through oncogene-driven alterations of the shared niche environment. Alongside the two oncogenes considered here, mutants with WNT activation caused by *Apc* loss also show paracrine effects that result from secretion of WNT inhibitory factors (Extended Data Fig. 12), which suggests that niche remodelling may be a general feature of tumorigenic mutants.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03605-0.

1. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
2. Medema, J. P. & Vermeulen, L. Microenvironmental regulation of stem cells in intestinal homeostasis and cancer. *Nature* **474**, 318–326 (2011).
3. Quail, D. F. & Joyce, J. A. Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* **19**, 1423–1437 (2013).
4. Clavería, C. & Torres, M. Cell competition: mechanisms and physiological roles. *Annu. Rev. Cell Dev. Biol.* **32**, 411–439 (2016).
5. Di Gregorio, A., Bowling, S. & Rodriguez, T. A. Cell competition and its role in the regulation of cell fitness from development to cancer. *Dev. Cell* **38**, 621–634 (2016).
6. Moreno, E. & Rhiner, C. Darwin's multicellularity: from neurotrophic theories and cell competition to fitness fingerprints. *Curr. Opin. Cell Biol.* **31**, 16–22 (2014).
7. Maruyama, T. & Fujita, Y. Cell competition in mammals — novel homeostatic machinery for embryonic development and cancer prevention. *Curr. Opin. Cell Biol.* **48**, 106–112 (2017).
8. Egeblad, M., Nakasone, E. S. & Werb, Z. Tumors as organs: complex tissues that interface with the entire organism. *Dev. Cell* **18**, 884–901 (2010).
9. Bjerknes, M. & Cheng, H. Colossal crypts bordering colon adenomas in Apc^{Min} mice express full-length Apc. *Am. J. Pathol.* **154**, 1831–1834 (1999).
10. Curtius, K., Wright, N. A. & Graham, T. A. An evolutionary perspective on field cancerization. *Nat. Rev. Cancer* **18**, 19–32 (2018).
11. Vermeulen, L. & Snippert, H. J. Stem cell dynamics in homeostasis and cancer of the intestine. *Nat. Rev. Cancer* **14**, 468–480 (2014).
12. Baron, C. S. & van Oudenaarden, A. Unravelling cellular relationships during development and regeneration using genetic lineage tracing. *Nat. Rev. Mol. Cell Biol.* **20**, 753–765 (2019).
13. Zong, H., Espinosa, J. S., Su, H. H., Muzumdar, M. D. & Luo, L. Mosaic analysis with double markers in mice. *Cell* **121**, 479–492 (2005).
14. Pontes-Quero, S. et al. Dual ifgMosaic: a versatile method for multispectral and combinatorial mosaic gene-function analysis. *Cell* **170**, 800–814 (2017).
15. Snippert, H. J. et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010).
16. Sheng, H., Shao, J., Townsend, C. M., Jr & Evers, B. M. Phosphatidylinositol 3-kinase mediates proliferative signals in intestinal epithelial cells. *Gut* **52**, 1472–1478 (2003).
17. Haigis, K. M. et al. Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. *Nat. Genet.* **40**, 600–608 (2008).
18. Stanger, B. Z., Datar, R., Murtaugh, L. C. & Melton, D. A. Direct regulation of intestinal fate by Notch. *Proc. Natl Acad. Sci. USA* **102**, 12443–12448 (2005).
19. Lopez-Garcia, C., Klein, A. M., Simons, B. D. & Winton, D. J. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330**, 822–825 (2010).
20. Ritsma, L. et al. Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging. *Nature* **507**, 362–365 (2014).
21. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
22. Fearon, E. R. Molecular genetics of colorectal cancer. *Annu. Rev. Pathol.* **6**, 479–507 (2011).
23. Nicholson, A. M. et al. Fixation and spread of somatic mutations in adult human colonic epithelium. *Cell Stem Cell* **22**, 909–918 (2018).
24. Leedham, S. J. et al. Clonality, founder mutations, and field cancerization in human ulcerative colitis-associated neoplasia. *Gastroenterology* **136**, 542–550 (2009).
25. Galandiuk, S. et al. Field cancerization in the intestinal epithelium of patients with Crohn's ileocolitis. *Gastroenterology* **142**, 855–864 (2012).
26. Vermeulen, L. et al. Defining stem cell dynamics in models of intestinal tumor initiation. *Science* **342**, 995–998 (2013).
27. Snippert, H. J., Schepers, A. G., van Es, J. H., Simons, B. D. & Clevers, H. Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion. *EMBO Rep.* **15**, 62–69 (2014).
28. Kozar, S. et al. Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. *Cell Stem Cell* **13**, 626–633 (2013).
29. Huels, D. J. et al. Wnt ligands influence tumour initiation by controlling the number of intestinal stem cells. *Nat. Commun.* **9**, 1132 (2018).
30. Bruens, L., Ellenbroek, S. I. J., van Rheenen, J. & Snippert, H. J. In vivo imaging reveals existence of crypt fission and fusion in adult mouse intestine. *Gastroenterology* **153**, 674–677 (2017).
31. Grün, D. et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**, 251–255 (2015).
32. Haber, A. L. et al. A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).
33. Ayyaz, A. et al. Single-cell transcriptomes of the regenerating intestine reveal a revival stem cell. *Nature* **569**, 121–125 (2019).
34. Cheng, C. W. et al. Ketone body signaling mediates intestinal stem cell homeostasis and adaptation to diet. *Cell* **178**, 1115–1131 (2019).
35. Gehart, H. & Clevers, H. Tales from the crypt: new insights into intestinal stem cells. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 19–34 (2019).
36. Stzepourginski, I. et al. CD34^{+} mesenchymal cells are a major component of the intestinal stem cells niche at homeostasis and after injury. *Proc. Natl Acad. Sci. USA* **114**, E506–E513 (2017).
37. Greicius, G. et al. *PDGFRα*$^{+}$ pericryptal stromal cells are the critical source of Wnts and RSPO3 for murine intestinal stem cells in vivo. *Proc. Natl Acad. Sci. USA* **115**, E3173–E3181 (2018).
38. McCarthy, N. et al. Distinct mesenchymal cell populations generate the essential intestinal BMP signaling gradient. *Cell Stem Cell* **26**, 391–402 (2020).
39. Jin, Z. et al. Differentiation-defective stem cells outcompete normal stem cells for niche occupancy in the *Drosophila* ovary. *Cell Stem Cell* **2**, 39–49 (2008).
40. Ellis, S. J. et al. Distinct modes of cell competition shape mammalian tissue morphogenesis. *Nature* **569**, 497–502 (2019).

Methods

Mice

All inducible Cre lines (*R26R-CreERT2*: JAX008463; *Villin-CreERT2*: JAX020282; *Lgr5-EGFP-IRES-CreERT2*: JAX008875; *Sftpc-CreERT2*: JAX028054; *Krt5-CreERT2:* JAX029155), the *R26R-Confetti* line (JAX017492), the *LSL-Kras*^{G12D} (JAX008179) line, the *Apc*$^{fl/fl}$ line (JAX009045) and the *Apc*^{Min} line (JAX002020) were obtained from The Jackson Laboratory. The *PIK3CA*$^{Lat\text{-}H1047R}$ line was donated by W. A. Phillips[41]. *Red2Onco* targeting vectors were generated by gene synthesis. Oncogene sequences were obtained from Addgene (*Notch1*^{ICD}*:* Addgene plasmid #15079[42]; *Kras*^{G12D}: Addgene plasmid #11549[43]; *PIK3CA*^{H1047R}: Addgene plasmid #12524[44]). Mice were created by inserting the *Red2Onco* cassette into the *tdimer2* locus in ES cells obtained from *R26R-Confetti* mice using CRISPR–Cas9 nickase-mediated homologous recombination. Insertion of the oncogenic sequence was confirmed by long-range PCR. Specific genotyping primers were designed outside of the homology arms and were used in combination with primers within the knock-in cassette.

Animal treatments

All experiments were approved by the UK Medical Research Council and University of Cambridge local ethical review committees and conducted according to Home Office project license PPL70/8296. To perform lineage tracing, tamoxifen (Sigma, T5648) dissolved in corn oil was injected intraperitoneally into 8–12-week-old mice. As indicated, the dose of tamoxifen used for each experiment was determined on the basis of recombination efficiency. For the clonal analysis, it is crucial that crypts are marked at clonal density so that labelled cells can be identified reliably as the output of an individual labelled cell. Titration experiments indicated that this can be achieved by a tamoxifen dose of 0.2 mg per 20 g mouse body weight. Tamoxifen and chemical inhibitors (LDN193189, BMP type I receptor blocking agent, Selleckchem, S2618; LGK974, porcupine inhibitor, Cayman Chemical, 14072) were concomitantly administered as indicated (Fig. 4h). LGK974 and LDN193189 were administered every 48 h through oral gavage, at a concentration of 5 mg/kg (LGK974) or 3mg/kg (LDN193189) in a vehicle of 0.5% Tween-80/0.5% methylcellulose. Chemical-treated mice did not show any loss of body weight or physical activity, demonstrating that there were no severe toxic side effects.

All mice were group housed under specific pathogen-free conditions in individually ventilated cages always with companion mice, and cages were placed under a 12-h light–dark cycle. Food and water were provided ad libitum. Room temperature was maintained at 22 °C ± 1 °C with 30–70% humidity. None of the mice had been involved in any previous procedures before the study.

Experiments were carried out with male and female animals, except for single-cell transcriptomic analysis, for which only females were used. No gender-specific differences were observed.

Intestine preparation

Mice were killed by cervical dislocation and the intestine collected by dissection. The small intestine was cut longitudinally and subsequently placed on a piece of cold PBS-soaked 3M paper, using forceps to flatten the tissue before fixation in 4% paraformaldehyde at 4 °C overnight (~18 h) with shaking. After fixation the intestine was washed for 3 × 6 h with PBS at 4 °C with shaking.

Tissue clearing and immunofluorescence

CUBIC clearings were performed as previously described[45]. In brief, fixed tissue was dissected into small fragments (~2 cm) and transferred into 10 ml CUBIC R1a solution (10% urea, 5% N,N,N′,N′-tetraki s(2-hydroxypropyl) ethyl-enediamine, 10% Triton X-100 and 25 mM NaCl in distilled water) in a 15-ml conical tube and incubated for 2 days at 37 °C with shaking. All subsequent incubation steps were then performed at 4 °C on a rotor. The fragments were incubated with blocking and permeabilization solution consisting of 5% DMSO, 0.5% Triton-X-100 and 2% normal donkey serum (NDS) in PBS overnight. The following day, the solution was replaced with primary antibody for cleaved caspase-3 (1:200; Cell Signaling Technology, 9661), OLFM4 (1:100; Cell Signaling Technology, 39141), p-ERK (1:100; Cell Signaling Technology, 4370), p-AKT (1:100; Cell Signaling Technology, 4060) or KRAS(G12D) (1:100; Cell Signaling Technology, 14429) diluted in blocking solution (1% DMSO, 0.5% Triton-X-100 and 2% NDS in PBS), and the whole mount was incubated for 3 days. The samples were then washed 6 times with fresh PBS over a 24-h period before incubation with secondary antibody (1:500; donkey anti-rabbit or anti-mouse, Invitrogen) diluted in blocking solution for 3 days. The samples were then washed 6 times with PBS and incubated with 2 µg/ml DAPI in PBS for 24 h. The samples were then transferred into 10 ml CUBIC R2 solution (50% sucrose, 25% urea, 10% 2,20,20′-nitrilotrie-thanol and 0.1% Triton X-100 in distilled water) in a 15-ml conical tube and incubated for 2 days at room temperature with shaking. To match the refractive index, samples were transferred into an Eppendorf tube containing 1 ml RapiClear 1.52 (Sunjin Lab) and incubated overnight at 4 °C. Samples were then mounted in a 0.25 mm i-spacer (Sunjin Lab) for confocal imaging.

Proliferation assay

To measure cell proliferation in vivo, 1 mg of 5-ethynyl-2′-deoxyuridine (EdU) (Life Technologies, A10044) was dissolved in 200 µL of PBS and injected into each mouse. Tissues were collected after 2 h and a Click-iT EdU Alexa Fluor 647 Imaging Kit (Thermo Fisher, C10340) was used to detect EdU+ cells.

Immunohistochemistry on near-native sections

Organs were dissected, fixed and embedded in 4% low-melt agarose as previously described[46] for generation of 100-µm-thick near-native sections using a LAICA VT 1000S Vibratome. The sections were removed from any remaining agarose using forceps and subsequently transferred to a 12-well plate into wells containing blocking and permeabilization solution (5% DMSO, 0.5% Triton-X-100 and 2% NDS in PBS) to be incubated overnight (~18 h) at 4 °C with shaking. The following day, the blocking and permeabilization solution was replaced with primary antibody for CPA1 (1:200; R&D systems, AF2765), β-catenin (1:200; Santa Cruz, sc-7199), SPC (1:300; Millipore, AB3786), LYZ (1:200; DAKO, A009902-2) or MUC2 (1:200; Abcam, ab90007) diluted in blocking solution (1% DMSO, 0.5% Triton-X-100 and 2% NDS in PBS) and the section incubated for 72 h at 4 °C. Sections were subsequently washed and incubated with secondary antibody (1:500; donkey anti-rabbit or goat AF647, Invitrogen) and DAPI diluted in blocking solution (1% DMSO, 0.5% Triton-X-100 and 2% NDS in PBS) for 48 h at room temperature with shaking. After being washed for 3 × 2 h with PBS, sections were carefully transferred from wells to microscope slides using a brush before mounting in RapiClear 1.52 (Sunjin Lab).

Immunohistochemistry on paraffin sections

Immunohistochemistry was performed according to standard protocols. In brief, the intestine was dissected and fixed in 4% PFA overnight at 4 °C before paraffin embedding. Paraffin-embedded sections (5 µm) were rehydrated, and the epitopes were exposed using Tris/EDTA (EDTA) buffer. Sections were then incubated in blocking solution (2% donkey or goat serum, 5% DMSO and 0.5% Triton-X-100 in PBS) at room temperature for 2 h. Primary antibody against NOTCH1 (1:100; Abcam, ab52627), β-catenin (1:50; Sigma, 05-665), or OLFM4 (1:100; Cell Signaling Technology, 39141) was diluted in blocking solution (1% DMSO, 0.5% Triton-X-100 and 2% normal goat serum in PBS) and the sections were incubated in this for 24 h at 4 °C. Sections were subsequently washed and incubated with secondary antibody (1:500; goat anti-rabbit HRP, Perkin Elmer) and DAPI diluted in blocking solution (1% DMSO, 0.5% Triton-X-100 and 2% normal goat serum in PBS) for

2 h at room temperature with shaking. The TSA Plus Cyanine 5 System (Perkin Elmer, NEL752001KT) was used for visualization.

In situ hybridization in cryo-sections

RNA staining was performed with an RNAscope Multiplex Reagent Kit V2 (ACDBio, 323100) according to the manufacturer's protocols. Collected tissues were fixed with 4% PFA, cryoprotected in OCT and sectioned at 16 µm. In brief, the sections were pre-treated with protease IV for 15 min. ACD-designed and -synthesized probes were used to detect transcripts of interest. Up to two probes were hybridized simultaneously. The RNAscope assay was followed by a regular immunostaining protocol with antibodies against RFP (1:50; SICGEN, AB8181-200), KRAS(G12D) (1:250; Cell Signaling Technology, 14429), p-AKT (1:50; Cell Signaling Technology, 4060) or β-catenin (1:50; Sigma, 05-665). Sections were subsequently washed and incubated with secondary antibody (1:500; horse anti-rabbit, Vector Lab or donkey anti-goat HRP, GeneTex) and mounted with Prolong Gold antifade reagent supplemented with DAPI. Probes for *Fabp1* (Cat No. 562831), *Axin2* (Cat No. 400331), *Id1* (Cat No. 312221), *Bmp2* (Cat No. 406661-C2), *Grem1* (Cat No. 314741-C3), *Sfrp2* (Cat No. 576891-C2), *Cd81* (Cat No. 556971), *Rspo3* (Cat No. 402011-C3), *Wif1* (Cat No. 412361-C3), *Lgr5* (Cat No. 312171-C2), and *Notum* (Cat No. 428981) were purchased from ACDBio. The TSA kits TSA Plus Cyanine 5 (Perkin Elmer, NEL745001KT), TSA Plus Fluorescein (Perkin Elmer, NEL741001KT) and TSA Plus TMR (Perkin Elmer, NEL742001KT) were used for visualization. The multiplexed staining was performed with the help of the Wellcome Trust–Medical Research Council Cambridge Stem Cell Institute Histology facility.

Single-cell sorting for sequencing

To collect mutant and wild-type cells from neighbouring crypts for scRNA-seq, we developed a strategy based on mosaic labelling. In this approach, *Villin-CreERT2*; *R26R-Red2Onco* mice (*Kras*^{G12D} or *PIK3CA*^{H1047R} mutation) were induced using a high dosage of tamoxifen (4 mg per 20 g body weight). The majority of crypts (>56% for *Red2-Kras*^{G12D} and >54% for *Red2-PIK3CA*^{H1047R}) were found to be monoclonal for mutant oncogene expression by 2 weeks post-induction at this level, while almost all of the remaining (wild-type) crypts, 94% for *Red2-Kras*^{G12D} and 92% for *Red2-PIK3CA*^{H1047R}, neighboured mutant crypts. After the dissection of these tissues, intestinal cell dissociation was performed as previously described with a few modifications to deplete villi further and enrich for intestinal crypts[32,47]. In brief, the proximal halves of small intestines were carefully washed with cold PBS, and villi were scraped off using a cover slip. The remaining tissue with crypts and lamina propria was then cut into 2-mm fragments and incubated in 10 ml Gentle Cell Dissociation Reagent (STEMCELL technologies) at 4 °C for 30 min. The fragments were then shaken vigorously and the supernatant was collected as the first fraction in a new conical tube. The remaining fragments were further incubated in fresh Gentle Cell Dissociation Reagent and a new fraction was collected every 30 min. The first and second fractions were discarded because they might contain some debris and villi. The third to fifth fractions contained the majority of the intestinal crypts. The crypt-enriched fractions were then washed with 10 ml of cold PBS and filtered through a 100-µm cell strainer (Falcon). The crypts were then further dissociated into single cells by incubation with TrypLE Express enzyme (GIBCO) at 37 °C for 5 min. The cells were filtered through a 40-µm cell strainer (Falcon). Remaining fragments with lamina propria were further digested with dissociation solution (2.5 mg/ml Liberase TL, Sigma; 10 U/ml DNase, Promega) at 37 °C for 1 h. To aid dissociation, we gently mixed tissue pieces by pipetting up and down every 10 min. After 20 min, supernatants were harvested and 1 volume of DMEM containing 10% fetal bovine serum (FBS, Sigma Aldrich) was added while adding 5 ml of fresh dissociation mix to the remaining tissue pieces. This step was repeated three times for a total time of 1 h. After completion of the three cycles, the remaining intestinal fragments were mechanically disaggregated on a 100-µm mesh using a syringe plunger. The cell suspension was filtered through a 40-µm cell strainer into a 15-ml tube. The tube and filters were washed twice with 1 ml of 2% FBS in PBS and the cell suspension was then centrifuged at 300*g* for 5 min. The supernatant was removed and the cell pellets were resuspended in 100 µl antibody mix (2% FBS; 1:125 CD45 (30-F11)-APC, BD Biosciences; 1:125 EPCAM (G8.8)-PE-Cy7, BioLegend) and incubated for 1 h on ice. Both epithelial and lamina propria cells were then washed with 3 ml of 2% FBS in PBS and filtered once more if clumps were observed. After centrifugation at 300*g* for 5 min, the cell pellets were resuspended in 1 ml of 2% FBS and 10 U/ml DNase in PBS for sorting. The MOFLO system (Summit software v5.2, Beckman Coulter) or SH800S Cell Sorter (SH800 software v2.1.5, Sony) was used for cell sorting and data were analysed with FlowJo software (v10.6.2, BD).

Library preparation and sequencing of RNA from single cells

scRNA-seq libraries were generated using 10X Genomics kits. As we wished to achieve statistically significant results across the wide range of intestinal cell types, for each biological replicate, cells sorted from gates R5, R6, R7 and R8 (Extended Data Fig. 6a) were pooled in equal ratios (RFP$^+$ epithelial cell:YFP$^+$ epithelial cell:mesenchymal cell:immune cells, 1:1:1:1) and loaded into one channel of a 10X Chromium microfluidics chip to package them into one library. Thus, as a result of this enrichment, the relative proportions of epithelial cells, immune cells and mesenchymal cells are not expected to reflect the in vivo ratios found in the small intestine. In our experiments, seven biological replicates (two for *Confetti*, three for *Red2-Kras*^{G12D} and two for *Red2-PIK3CA*^{H1047R}) were used to make seven libraries in total. The libraries were sequenced on an Illumina HiSeq 4000.

Intestinal organoid culture and imaging

Intestinal epithelial organoids were established as previously described[48]. In brief, we freshly isolated crypts from mouse small intestine and mixed the crypts with 20 µl Matrigel (Corning). After Matrigel polymerization, the crypts were cultured in ENR medium composed of advanced Dulbecco's modified Eagle's medium/F12 supplemented with penicillin/streptomycin, 10 mM HEPES, glutamax, N2 (Life Technologies), B27 (Life Technologies) and 1 mM *N*-acetylcysteine (Sigma), 50 ng/ml murine recombinant epidermal growth factor (EGF; Peprotech), R-spondin1 (conditioned medium from 293T-HA-RspoI-Fc cells, 10% final volume), and 100 ng/ml Noggin (Peprotech) for 3 days to generate organoids. On the basis of the evidence for oncogene-associated changes in the BMP and WNT pathways, we tested two conditions: (1) withdrawal of Noggin from the full culture medium (−Nog), which mimics a BMP-rich environment in *Red2-Kras*^{G12D}, and (2) withdrawal of Noggin from R-spondin-reduced (10% to 1%) medium (−Nog Rspo^{low}), which mimics the BMP-rich and WNT-deprived environment in *Red2-PIK3CA*^{H1047R} (Extended Data Fig. 9h–j). For the organoid formation assay (Extended Data Fig. 9l, m), organoids were collected 2 days after treatment and then dissociated into single cells with TrypLE (Thermo Fisher Scientific) for 5 min at 37 °C. The dissociated cells were filtered through a 40-µm cell strainer (Falcon). We mixed 2×10^4 collected cells with 20 µl Matrigel and seeded them into each well of 48-well plate. After 20 min of solidification at 37 °C, 250 µl of WENR medium was added. ENR was supplemented with WNT3A (conditioned medium (CM) from WNT3A L-cells, 50% final volume), 10 µM Y-27632 (ROCK inhibitor, STEMCELL Technologies) and nicotinamide (Sigma) to make the WENR medium. Organoids were imaged and counted using an EVOS M7000 microscope (EVOS M7000 Revision software v2.0, Thermo Scientific).

To obtain organoids from *Red2Onco* mice, RFP$^+$ cells were sorted from *Villin-CreERT2;Red2Onco* mice two weeks after tamoxifen administration. We pelleted 1×10^5 sorted cells in 20 µl Matrigel. After solidification at 37 °C, organoids were initially formed and treated with the following conditions: WENR: WNT3A CM, EGF, Noggin and R-spondin1 CM; ENR: EGF, Noggin, and R-spondin1 CM; −Nog: withdrawal of Noggin from the ENR medium; −Nog Rspo^{low}: withdrawal of Noggin and lowered

R-spondin1 CM concentration (10% to 1%) from the ENR medium; −EGF + Gefitinib: withdrawal of EGF and addition of Gefitinib (EGFR inhibitor) (Extended Data Fig. 9q). Note that organoids from both mutants are sensitive to growth factor withdrawal while they are resistant to EGF removal.

To produce conditioned medium, 250 µl of WENR medium was added to organoids from *Confetti* or *Red2Onco* mice. After 7 days, organoids were supplemented with Noggin-deprived medium (−Nog) to generate conditioned medium. The conditioned medium was harvested after 7 days of culture. The medium was centrifuged at 1,000*g* for 5 min at 4 °C and the supernatant was then filtered through a 0.22-mm filter (Sartorius). Wild-type organoids were then treated with CM from either *R26R-Confetti* or *Red2Onco* mice for 6 h with or without LDN193189 (BMP type I receptor blocking agent, Selleckchem, S2618, 1 µM) to assess the effect of the CM on activation of BMP signalling.

RNA isolation and quantitative PCR

Isolated intestinal organoids or cells were resuspended in 350 µl of RLT buffer (QIAGEN). Total RNA was isolated using RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. Equivalent quantities of total RNA were reverse-transcribed to synthesize cDNA using a SuperScript cDNA synthesis kit (Life Technology). qPCR was performed using Power SYBR Green PCR Master Mix (Thermo). Primer sequences were as follows:

Lgr5: F, CGGGACCTTGAAGATTTCCT; R, GATTCGGATCAGCCAGCTAC

Olfm4: F, CGAGACTATCGGATTCGCTATG; R, TTGTAGGCAGCCAGA GGGAG

Id1: F, ATCGCATCTTGTGTCGCTGAG; R, AGTCTCTGGAGGCTGAAAGGT

Ttr: F, AGCCCTTTGCCTCTGGGAAGAC; R, TGCGATGGTGTAGTGGCGA TGG

Chga: F, GTCTCCAGACACTCAGGGCT; R, ATGACAAAAGGGGACACCAA

Clca1: F, TCTTGTGTAGATGCCATCATTTTT; R, CCAATGTCACAGCCCT CATA

Lyz: F, ATGGAATGGCTGGCTACTATGGAG; R, CTCACCACCCTCTTTGCA CATTG

Alpi: F, AGGATCCATCTGTCCTTTGG; R, ACGTTGTATGTCTTGGACAG

Aldh1a1: F, TGTCGGATTTAGGAGGCTGC; R, GCATTGGAAAATTCCAGG GGATG

Mki67: F, CCAGCTGCCTGTAGTGTCAA; R, TCTTGAGGCTCGCCTTGATG

Ccnb2: F, GCCAAGAGCCATGTGACTATC; R, CAGAGCTGGTACTTTGGT GTTC

Axin2: F, GCTCCAGAAGATCACAAAGAGC; R, AGCTTTGAGCCTTCAGCATC

Ascl2: F, GCCTGACCAAATGCCAAGTG; R, ATTTCCAAGTCCTGATG CTGC

Bmp2: F, GAAGTTCCTCCACGGCTTC; R, AGATCTGTACCGCAGGCACT

Bmp7: F, ACGGACAGGGCTTCTCCTAC; R, ATGGTGGTATCGAGGGTG GAA

Cd81: F, ACACCTTCTACGTGGGCATC; R, TGCTTCACATCCTTGGCGAT

Grem1: F, TGGAGAGGAGGTGCTTGAGT; R, AACTTCTTGGGCTTGCA GAA

Frzb: F, TGCAAATGTAAGCCTGTCAGAGC; R, TCCACAACGGCGGTCA CATC

Rspo3: F, ATGCACTTGCGACTGATTTCT; R, GCAGCCTTGACTGACAT TAGGAT

Sfrp2: F, CGTGGGCTCTTCCTCTTCG; R, ATGTTCTGGTACTCGATGCCG

Sfrp4: F, GTGGCGTTCAAGGATGATGCTTC; R, TTACTGCGACTGGTGC GACTG

Pdgfra: F, GCAGTTGCCTTACGACTCCAGA; R, GGTTTGAGCATCTTCAC AGCCAC

Foxl1: F, TCATCATGGATCGCTTCCCG; R, CCTCTTCCTGCGCCGATAAT

Dkk2: F, ACTCCATCAAGTCCTCTC; R, TCACATTCCTTATCACTGCTG

Wif1: F, ACCCTGCCGAAATGGAGGT; R, TTGGGTTCGTGGCAGGTTC

Notum: F, CTGCGTGGTACACTCAAGGA; R, CCGTCCAATAGCTCCGT ATG

Actin: F, AAATCGTGCGTGACATCAAA; R, AAGGAAGGCTGGAAAAGAGC

In vitro co-culture

As described above, lamina propria was isolated from mouse small intestine. The cells were resuspended in 100 µl antibody mix (2% FBS; 1:125 CD45 (30-F11)-APC, BD Biosciences; 1:125 EPCAM (G8.8)-PE-Cy7, BioLegend; 1:125 PDGFRA (CD140a) (APA5)-BV421, Biolegend; 1:125 CD81 (Eat-2)-PE, BioLegend) and incubated for 1 h on ice. The cells were then washed with 3 ml of 2% FBS in PBS and filtered once more. After centrifugation at 300*g* for 5 min, the cell pellets were resuspended in 1 ml of 2% FBS and 10 U/ml DNase in PBS for sorting. We cultured 8×10^4 sorted STC1 ($CD81^-$ $PDGFRA^{lo}$ $EPCAM^-$ $CD45^-$) or STC2 cells ($CD81^+$ $PDGFRA^{lo}$ $EPCAM^-$ $CD45^-$) in 48-well plates in 20 µl Matrigel together with 3×10^4 $EGFP^+$ ISCs sorted from *Lgr5-EGFP-IRES-CreERT2* mice. Niche factor-reduced medium (−Nog $Rspo^{low}$, described above) was supplemented to test the supporting function of the mesenchymal cells on organoid formation. Organoids were imaged using an EVOS M7000 microscope (Thermo Scientific).

Confocal imaging and quantification

Mouse small intestine samples were prepared for confocal microscopy as described above. Confocal imaging was performed with a Leica SP5 TCS confocal microscope (LAS software v2.8.0, Leica). The argon laser intensity was set to 20% and Z-step size was set to 2 µm. Laser scanning frequency was set to 400 Hz and the frame average was set to 3. Section images were processed and analysed using ImageJ. To chart the clone dynamics of labelled mutant and wild-type cells, intestinal crypt bottoms were visualized under the confocal microscope and clone sizes were quantified according to their circumferential span at the '+4 row position', representing the border of the niche domain[20] (Z-step 8–10; 16–20 µm upward from the crypt base). Clone sizes at the crypt bottom are denoted as eighths of the crypt circumference: 1 indicates that a fraction of 1/8 of the crypt circumference is occupied by a single clone, and so on. Clone sizes at the crypt neck are determined as the total circumferential angle spanned by the clone. Statistical analysis of the detailed cellular arrangements of clones at the crypt base confirmed that such measures were representative of ISC clone sizes (Extended Data Fig. 2b–e)

To perform reliable and consistent measurements, we set *z*-stack positions for each crypt by considering the first *z*-plane that showed any epithelial cell as the $z = 0$ position of individual crypts. Consistent with studies based on the original *Confetti* construct, yellow (EYFP), red (RFP) and cyan (mCFP) labelling was induced in approximately equal proportions, while green (nuclear GFP) cells were observed only infrequently (Extended Data Fig. 1b). Note the similar degree of induction between models, which enables comparative clonal analysis using the *R26R-Confetti* and *Red2Onco* mice.

The central angle of the clonal circular sector (Extended Data Fig. 2g) was determined by three points (P1: clone starting point, Pc: crypt centre, P3: clone end point) on the crypt circumference. Only RFP^+ clones (wild-type in *Confetti* mice, mutant in *Red2Onco* mice) or YFP^+ clones (wild-type) were quantified. Wild-type clones in *Red2Onco* mice were grouped with respect to their proximity to crypts containing mutant clones (Fig. 1d). Wild-type (YFP^+) clones distant (three crypt diameters or more) from crypts containing mutant clones were considered to be 'remote' while YFP^+ clones neighbouring fixed mutant (RFP^+) crypts were considered to be 'proximate'. In Extended Data Fig. 4, proximate wild-type clonal data from Fig. 1 were reanalysed and partitioned into two groups, 'inner' and 'outer', depending on whether the geometric centre of the clones was positioned in the crypt half nearest to or furthest from the mutant crypt (Extended Data Fig. 4e). For quantification of crypt base columnar cells, either the nuclear morphology (Extended Data Figs. 3i, 4j) or overlap with LGR5–GFP was used (Fig. 2d, f). Three-dimensional representations were created using ImageJ (1.52e). The number of cells per crypt was counted in *z*-stacks.

Crypt size and circularity measurement

As indicated in Extended Data Fig. 5a, for analysis of the circularity and area of individual crypts, thresholded images were generated by ImageJ. The regions of interest were then analysed using the Analyze Particles function in ImageJ to obtain outlined images and measures of location, size and circularity of individual crypts. Circularity was defined as $4\pi \times (\text{area}/\text{perimeter}^2)$ so that unity represents a perfect circle and a figure <1 indicates a degree of distortion (Extended Data Fig. 5d, e). For these analyses, the particle size was set to a lower limit of 400 pixels. The size limit corresponds to circular particles of approximately 20 μm in diameter and was chosen to exclude cells in the lamina propria from the analysis.

Intestinal clone dynamics in Red2Onco mice

To address the influence of oncogene expression on the clone dynamics of mutant and wild-type crypts, we use an established statistical modelling approach used previously to study the stem cell dynamics of the intestinal crypt under normal and perturbed conditions[19,27]. In this approach, quantitative features of the clonal data, including the time evolution of the average clone size (scored as fraction of the crypt circumference) and crypt fixation frequency, are fit against a statistical modelling scheme based on a minimal number of adjustable parameters. Best fit parameters are then obtained using a standard least-squares fitting approach. Details of the modelling framework, its justification, and the implementation of the statistical analysis are presented in the Supplementary Theory.

Single-cell RNA-seq data analysis

Data processing for scRNA-seq. The raw sequencing data from the 10X Genomics platform were processed using CellRanger (v2.1.1). CellRanger aligned reads, filtered empty dropouts and counted unique molecular identifiers (UMIs) to generate a count matrix. We used Ensembl GRCm38/mm10 (release 92) appended with the sequences for *tdimer2* (RFP) and *EYFP* as the reference genome for the read alignment. The filtration of empty droplets was also checked with the R package DropletUtils (v1.2.2). To filter out low quality cells, cells with fewer than 100 genes were removed. In addition, cells with mitochondrial proportions above 15% were discarded from further analysis. Genes expressed in fewer than three cells were removed. Basic statistics and QC metrics for all samples are included in Supplementary Table 1. $EYFP^+$ cells were considered to be wild-type epithelial cells in the intestinal crypts of *Confetti*, *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} animals, whereas $EPCAM^+$ and $EYFP^-$ cells were considered to be mutant epithelial cells in the intestinal crypt. UMIs were normalized by a deconvolution method using the R package scran (v1.12.1)[49].

Dimension reduction and data visualization. Principal component analysis (PCA) combined with technical noise modelling was applied to the normalized data for dimension reduction, which was implemented by the denoisePCA() function in the R package scran. This denoise PCA does not strictly require explicit feature selection, such as highly variable genes. The data were then projected using two-dimensional UMAP or t-SNE with default parameter settings.

The biological replicates for each condition (two for *Confetti*, three for *Red2-Kras*^{G12D} and two for *Red2-PIK3CA*^{H1047R} models) overlapped well with each other, implying reproducibility between replicates and providing confidence in the statistical reliability of our comparative analysis. The reproducibility confirmed that batch effects between different samples and conditions were negligible and batch effect correction was not necessary for further analysis (Extended Data Fig. 6b).

Data clustering and cluster annotation. Based on the expression of marker genes such as *Epcam*, *Vim* and *Ptprc* (*Cd45*), all cells were split into three major categories: epithelial cells, mesenchymal cells and immune cells. For each category, we then performed clustering using a graph-based method. First, a shared nearest-neighbour graph was constructed using *k* nearest neighbours of each cell (buildSNNGraph function in the R package scran). *k* was set to 6 for epithelial cells and to 10 for mesenchymal and immune cells. In this graph, two cells were connected by an edge if they shared nearest neighbours, with the edge weight determined by the highest average rank of the shared neighbours. Then the Walktrap method from the R package igraph (v1.2.4.1) (with steps = 4 as the default option) was used to identify densely connected communities that were considered to be cell clusters.

Cell clusters were annotated on the basis of differentially expressed genes and known marker genes for cell types. To annotate epithelial cells of the intestinal crypts, we referred to marker genes for cell types used in Extended Data Fig. 1 of a previous study[32]. If a few neighbouring clusters in the dimension reduction spaces shared key expression patterns, they were merged into one cell type manually. Thus, we classified 33 clusters into 8 epithelial cell types (Extended Data Fig. 6c–e). As a result, we confirmed that the fractions of individual cell types in the *Confetti* control were comparable to those in previous reports[32–34]. The fractions were as follows: 29.6% for stem cells; 30.2% for transit-amplifying cells, 11% for enterocyte progenitors; 5.3% for enterocytes; 3.6% for Paneth cells; 14.7% for goblet cells; 2.4% for tuft cells; 3.2% for enteroendocrine cells. We note that our stringent removal of the villus fraction during sample preparation is likely to explain the smaller fraction of enterocyte lineage cells that are found compared to other reports[32–34]. While annotating epithelial, mesenchymal, and immune cells we found that, out of a total of 21,183 cells, 1,400 cells (6.6%) had a very low number of genes, and showed ambiguous and promiscuous expression. We considered them to be pseudo-cells, possibly contaminated with ambient RNA floating in single-cell suspension, and they were removed from further analysis.

Statistical analysis of cell-type composition. To obtain an overview of cell-type composition, we calculated the fraction of each cell type by taking the number of a cell type normalized by the total number of cells for each animal. To detect changes in cell-type composition, we modelled the number of each cell type as a random count variable using a Poisson process. The rate of detection was modelled using the total number of cells profiled in each condition (RFP^+ or YFP^+) of an animal as an offset variable, with the condition of each animal (*Confetti* or *Red2Onco*) used as a covariate. We fitted the model using the glm function in the R package stats (v3.6.0). The *P* value for the significance of the effect due to oncogene expression was estimated using a likelihood ratio test on the regression coefficient.

To interpret the fractional change of stem and progenitor cells in Extended Data Fig. 6h, it is important to note that the mutant stem cell fraction may not correlate straightforwardly with the increased division rate observed in *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice (Extended Data Fig. 3h, i). Although both mutant and neighbouring WT crypts show a fractional decrease in stem cell number, the absolute number of mutant stem cells in both *Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R} mice remained comparable to the *Confetti* and remote WT controls (Extended Data Fig. 6k, l). This implies that although oncogene expression increases the division rate of mutant stem cells, it may result in more rapid production of progenies rather than an increase in stem cell number. If oncogene expression drives an increased rate of differentiation[50], it is more likely that the number of mutant stem cells may be outweighed by an increase in the abundance of their differentiating progenies, leading to a net reduction in the mutant stem cell fraction. It is likely that such changes in differentiation bias are responsible for the significant drop in the fraction of mutant stem cells, as observed in Extended Data Fig. 6h, l, m.

As described above, changes in the composition of immune and mesenchymal cells were also analysed. The results are shown in Extended Data Fig. 8g, h, except for stromal cells 1, 3, B cells, macrophages 1 and

plasma cells, which did not show significant changes in fractions (data not shown).

Analysis of priming towards differentiation. To understand the origin of the altered relative proportions of mutant and wild-type epithelial cell types in *Red2Onco* mice, we calculated the degree of fate priming of stem cells and their progenitors using the R package FateID (v0.1.9)[51]. We first identified marker genes for each differentiated sub-lineage, including enterocytes, Paneth cells, goblet cells, tuft cells and enteroendocrine cells. To this end, we analysed scRNA-seq data generated from small intestinal crypts in a previous study[32]. We identified highly expressed genes in each of the differentiated sub-lineages using the findMarkers function from the R package scran (FDR < 0.05). As a result, the numbers of the marker genes for differentiated sub-lineages were as follows: 2,497 genes for enterocytes; 4,941 genes for goblet cells; 2,494 genes for Paneth cells; 3,416 genes for tuft cells; 4,660 genes for enteroendocrine cells. Using the marker genes for differentiated sub-lineages, we then calculated the fate bias scores of all epithelial cells for each differentiated sub-lineage using two functions (reclassify and fateBias) of the FateID package with default parameter values. We considered the fate bias scores to be the degree of fate priming. Then the distribution of the priming scores for stem cells and their progenitors in the *Red2Onco* mice was compared with that from the *Confetti* control using the Kolmogorov–Smirnov test in GraphPad Prism 8.

Gene set enrichment analysis (GSEA). Based on prior knowledge, we selected three major signalling pathways (WNT, BMP and NOTCH) that substantially affect the stem cell and differentiation potential of intestinal stem cells[2,35,52]. To define gene sets for these pathways, we curated and referred to previous studies where the pathways of interest were considered to be specifically altered in intestinal crypts. Specifically, as the gene set for the BMP pathway, we used 293 genes that were highly expressed in BMP4-treated LGR5^{+} intestinal organoids[53]. As the gene set for the WNT pathway, we used 113 genes (MSigDB ID M1428) that were highly expressed in the intestinal crypts of *Apc* KO mice[54]. As the gene set for the NOTCH pathway, we used 315 genes that were highly expressed in the intestinal crypts of *Atoh1* KO mice as compared to DBZ-treated mice[55].

Using the manually curated gene sets, we performed GSEA using the R package AUCell (v1.6.1). To identify 'active' cells with high enrichment scores, we fitted the distribution of the enrichment scores using the AUCell_exploreThresholds function of the R package AUCell and the cut-off ('Global_k1' value) for the high enrichment scores was selected among those suggested by the AUCell_exploreThresholds function. The fraction of active cells above the cut-off for each pathway is as shown in Fig. 3e and Extended Data Fig. 7m.

See Supplementary Table 2 for the details of the gene sets used in Fig. 3 and Extended Data Fig. 7.

Estimating the degree of transcriptomic change. To estimate the degree of transcriptomic change for mesenchymal and immune cells, for each cell type we tested the statistical significance of the distance between the cell clusters in the *Confetti* and *Red2Onco* mice using two different methods.

In the first method, statistical significance was estimated based on how much larger inter-variability between cell clusters was than intra-variability. To this end, we defined cell-to-cell variability as 1 − (Pearson correlation) for any pair of cells. If two cells from the same model (either the *Confetti* control or *Red2Onco* mice) were selected, the cell-to-cell variability was considered to be the intra-variability. On the other hand, if one cell was selected from the *Confetti* control and the other from a *Red2Onco* mice to form a pair, the cell-to-cell variability was considered to be the inter-variability. The dispersion and mean of the distributions of the inter-variability and intra-variability were summarized in the form of a *t*-statistic. To make a null hypothesis distribution for this statistic, we randomly sampled cells for each cell type and formed 'pseudo-confetti' and 'pseudo-Red2' samples, before calculating the *t*-statistics for the variability; this was repeated 20,000 times and generated an empirical null hypothesis distribution of the statistics. The significance of larger inter-variability than intra-variability was tested against this null hypothesis distribution, generating a *P* value, P_{VAR}.

In the second approach, we used Augur (v1.0.2), a method to rank cell types based on their degree of response to biological perturbations in single-cell data[56]. Augur uses a machine-learning framework to quantify the separability of perturbed and unperturbed cells within a high-dimensional space of single-cell measurements. To feed our data set to the Augur workflow, we randomly selected *Confetti* cells and split them into two groups (*Confetti* set A and B) for each cell type. Then we calculated condition-specific AUC values for each cell type by comparing *Confetti* set A with *Red2Onco* mice and *Confetti* set A with B using the calculate_auc function with subsample_size = 6 and default values for other parameters. To calculate the null distribution of AUCs for each cell type and condition, the calculate_auc function was executed again for *Confetti* set A versus *Red2Onco* mice and for *Confetti* set A versus B with subsample_size = 6, augur_mode = 'permute', and default values for other parameters. Using the condition-specific AUCs and their null distributions, the statistical significance of differential prioritization between *Confetti* set A and *Red2Onco* mice and *Confetti* sets A and B was calculated by running calculate_differential_prioritization function with default parameter values. To guarantee the robustness of our results, we repeated the random splitting of the *Confetti* cells followed by the calculation of the statistical significance of the differential prioritisation 50 times, and chose the median *P* values as the representative ones (P_{AUGUR}) for each cell type and condition. P_{VAR} and P_{AUGUR} for each cell type were compared, as shown in Fig. 3i for mesenchymal cells and Extended Data Fig. 8i for immune cells.

Gene ontology analysis. Genes that were differentially expressed in the stromal cell STC2 cluster between *Red2-PIK3CA*^{H1047R} and *Confetti* mice were identified using the findMarkers function from the R package scran. We selected 377 genes with *P* value <0.05 and absolute value of $\log_2$(fold change) > α(= 0.259) for gene ontology analysis, where α is the 95th percentile of absolute values of all $\log_2$(fold change). The gene ontology analysis was performed using DAVID (v6.8)[57].

Statistical analysis

GraphPad Prism 8 software was used to perform statistical analyses. Unless otherwise specified, statistical significance was determined by applying Student's *t*-test or analysis of variance (ANOVA) to raw values from at least three independent experiments.

Sample size, randomization and blinding

The sample size was chosen based on previous experience in the laboratory and the literature[15,19]. No statistical methods were used to predetermine sample size. No method of randomization was followed and no animals were excluded from this study. The investigators were not blinded to sample allocation during the experiments and assessment of results.

Extended Data Figure 1. In a, images are representative of two independent experiments; b, *n* = 3 mice per group, 10 fields of images were analysed per mouse; c–e, *n* = 5 mice per group, 100 crypts analysed per mouse; f, images are representative of two independent experiments; g, images are representative of two independent experiments.

Extended Data Figure 2. In c–e, *n* = 3 mice per group (*R26R-Confetti*: 211 clones; *Red2-Kras*^{G12D}: 171 clones; *Red2-PIK3CA*^{H1047R}: 142 clones); f, *n* = 3 mice per group, 10 fields of images (2.41 mm^2) were analysed per mouse; i, *n* = 6 mice per group and time point; in i–k, 112, 103, 80, 169 clones

were scored for Conf, 76, 78, 87, 128 for Red2-wild-type R2N1, 80, 93, 129, 178 for Red2-Wild-type R2KR, 81, 106, 94, 187 for Red2-Wild-type R2P3, 93, 91, 74, 118 for Red2-Mutant R2N1, 107, 63, 89, 120 for Red2-Mutant R2KR, and 121, 61, 98, 111 for Red2-Mutant R2P3 at 4 d, 1 w, 2 w, 3 w after induction, respectively; j, k, $n = 6$ mice per group.

Extended Data Figure 3. In b, c and e, f, source experimental data, mice cohort numbers and clone numbers are as in Fig. 1f and Extended Data Fig. 2i, respectively; g, images are representative of three independent experiments; h, i, $n = 5$ mice per group, 100 clones analysed per mouse.

Extended Data Figure 4. In b, $n = 3$ mice per group, 10 images analysed per mouse; c, d, $n = 6$ mice per group and time point; *R26R-Confetti* and remote WT control data from Extended Data Fig. 2j, k are reproduced for comparison in c, d; f, $n = 6$ mice per group and time point; 103, 80, 169 clones scored for Conf, 49, 52, 140 for 'inner' proximate WT R2KR, 41, 47, 101 for 'outer' proximate WT R2KR, 33, 68, 38 for 'inner' proximate WT R2P3, and 31, 69, 50 for 'outer' proximate WT R2P3, at 1, 2, 3 w after induction, respectively; g, $n = 6$ mice per group and time point (here we reproduced *R26R-Confetti* control data (Extended Data Fig. 2j) for comparison, and we regrouped proximate WT data (Fig. 1f) into inner and outer clones (f, g)); h, $n = 6$ mice per group and time point; 103, 80, 169 clones scored for Conf, 90, 99, 241 for proximate WT R2KR, 64, 137, 88 for proximate WT R2P3, and 78, 87, 128 for R2N1-prox at 1, 2, 3 w after induction, respectively; analysis based on experimental data presented in Fig. 1f; j, $n = 5$ mice per group, 100 crypts analysed per mouse.

Extended Data Figure 5. In b, c, 265 crypts for R26R-*Confetti*, 326 crypts for *Red2-Kras*^{G12D}, and 305 crypts for *Red2-PIK3CA*^{H1047R} were analysed; in c, data from b are regrouped and displayed as a violin plot (the *n* for each group is shown); d, e, 265 crypts for *R26R-Confetti*, 326 crypts for *Red2-Kras*^{G12D}, and 305 crypts for *Red2-PIK3CA*^{H1047R} were analysed; in e, data from d are regrouped and displayed as a violin plot (the *n* for each group is shown); h, $n = 3$ mice per group and time point; 146 clones scored for Conf, 60 neighbouring 1 mutant crypt and 58 neighbouring >1 mutant crypt for proximate WT R2KR, and 58 neighbouring 1 mutant crypt and 63 neighbouring >1 mutant crypt for proximate WT R2P3; i, $n = 3$ mice per group; j, k, $n = 5$ mice per group, 100 crypts analysed per mouse; n, $n = 3$ mice per group, 50 images analysed per group; p, 670 crypts for *R26R-Confetti*, 452 crypts for *Red2-Kras*^{G12D}, and 457 crypts for *Red2-PIK3CA*^{H1047R} were analysed; $n = 3$ mice per group.

Extended Data Figure 6. In b, *n* for 'within condition': 5 except MF ($n = 4$) and IC ($n = 2$); *n* for 'between condition': 16 except MF ($n = 11$) and IC ($n = 8$); see Source Data for statistical significance; h, i, $n = 2$ for *Confetti* and *Red2-PIK3CA*^{H1047R}, $n = 3$ for *Red2-Kras*^{G12D}; j–l, $n = 4$ mice per group, 100 crypts analysed per mouse; m, n, $n = 6$ mice per group.

Extended Data Figure 7. In c, $n = 5$ mice per group; d, e, $n = 5$ mice per group, 100 crypts analysed per mouse; h–j, $n = 3$ mice per group, 50 crypts analysed per group; m, $n = 2$ for Conf and R2P3, $n = 3$ for R2KR.

Extended Data Figure 8. In g, h, $n = 2$ for Conf and R2P3, $n = 3$ for R2KR.

Extended Data Figure 9. In a–d, $n = 3$ mice per group, 50 crypts analysed per group; e, f, $n = 5$ mice per group; h, images are representative of three independent experiments; i, $n = 3$ independent experiments; j, images are representative of three independent experiments; k, experimental set-up for l, m; l, m, images are representative of organoids quantified in m, $n = 3$ independent experiments; n, $n = 5$ mice per group; p, $n = 3$ independent experiments; q, images are representative of three independent experiments; r, $n = 3$ independent experiments.

Extended Data Figure 10. In b, c, $n = 3$ mice per group, 50 crypt pairs analysed per group; g, h, $n = 3$ mice per group.

Extended Data Figure 11. In a, b, $n = 3$ mice per group; c, $n = 3$ mice per group, 212 clones scored for Conf + Veh, Veh: 85, LDN: 78, LGK: 96 for remote WT R2KR, Veh: 91, LDN: 90, LGK: 113 for remote WT R2P3, Veh: 130, LDN: 108, LGK: 75 for proximate WT R2KR, and Veh: 132, LDN: 81, LGK: 65 for proximate WT R2P3; d, $n = 3$ mice per group, Veh: 65, LDN: 116, LGK: 96 clones scored for mutant R2KR, and Veh: 82, LDN: 99, LGK: 92 for mutant R2P3; e, $n = 3$ mice per group, Veh: 65, LDN: 116, LGK: 96 clones scored for mutant R2KR, and Veh: 82, LDN: 99, LGK: 92 for mutant R2P3; f, g, $n = 3$ mice per group, 100 crypts analysed per mouse; i, $n = 3$ mice per group, 180 crypts for *Lgr5-EGFP-IRES-CreERT2* control, 142 crypts for *Lgr5-EGFP-IRES-CreERT2;LSL-Kras*^{G12D}, and 160 crypts for *Lgr5-EGFP-IRES-CreERT2;PIK3CA*$^{Lat\text{-}H1047R}$ mice; j, k, $n = 3$ mice per group, 50 crypts analysed per group; l, m, $n = 3$ mice per group, 50 crypt pairs analysed per group.

Reporting summary

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

Data availability

The scRNA-seq data generated for this study have been deposited in ArrayExpress under E-MTAB-8656. The reference genome sequence was downloaded from Ensembl (http://www.ensembl.org/Mus_musculus) and used for alignment of the scRNA-seq data. To evaluate stem cell priming, scRNA-seq data were obtained from the Single Cell Portal (https://portals.broadinstitute.org/single_cell/study/small-intestinal-epithelium) and used to define gene sets for differentiated sub-lineages of epithelial cells. The lists of marker genes used to annotate types of epithelial, mesenchymal and immune cells in Fig. 3b, f and Extended Data Figs. 6d, 8a–c are given in Supplementary Tables 1, 3. Gene sets used in Fig. 3d, e and Extended Data Fig. 7a, b, k–m are provided in Supplementary Table 2. Source data are provided with this paper.

Code availability

The statistical analysis of the clone fate data, based on a fit to the established modelling scheme, was performed using a FORTRAN (G95 compiler) code developed for this study. The scRNA-seq data were analysed using publicly available R packages. The codes and data used for clonal analysis and scRNA-seq data analysis have been deposited in GitHub (available at https://github.com/BenSimonsLab/Yum_Nature_2021).

41. Kinross, K. M. et al. An activating Pik3ca mutation coupled with Pten loss is sufficient to initiate ovarian tumorigenesis in mice. *J. Clin. Invest.* **122**, 553–557 (2012).
42. Murtaugh, L. C., Stanger, B. Z., Kwan, K. M. & Melton, D. A. Notch signaling controls multiple steps of pancreatic differentiation. *Proc. Natl Acad. Sci. USA* **100**, 14920–14925 (2003).
43. Orsulic, S. et al. Induction of ovarian cancer by defined multiple genetic changes in a mouse model system. *Cancer Cell* **1**, 53–62 (2002).
44. Zhao, J. J. et al. The oncogenic properties of mutant p110α and p110β phosphatidylinositol 3-kinases in human mammary epithelial cells. *Proc. Natl Acad. Sci. USA* **102**, 18443–18448 (2005).
45. Susaki, E. A. et al. Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nat. Protocols* **10**, 1709–1727 (2015).
46. Snippert, H. J., Schepers, A. G., Delconte, G., Siersema, P. D. & Clevers, H. Slide preparation for single-cell-resolution imaging of fluorescent proteins in their three-dimensional near-native environment. *Nat. Protocols* **6**, 1221–1228 (2011).
47. Stzepourginski, I., Eberl, G. & Peduto, L. An optimized protocol for isolating lymphoid stromal cells from the intestinal lamina propria. *J. Immunol. Methods* **421**, 14–19 (2015).
48. Sato, T. et al. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
49. Lun, A. T., Bach, K. & Marioni, J. C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol.* **17**, 75 (2016).
50. Feng, Y. et al. Mutant *Kras* promotes hyperplasia and alters differentiation in the colon epithelium but does not expand the presumptive stem cell pool. *Gastroenterology* **141**, 1003–1013 (2011).
51. Herman, J. S., Sagar, & Grün, D. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat. Methods* **15**, 379–386 (2018).

52. Pinto, D., Gregorieff, A., Begthel, H. & Clevers, H. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes Dev*. **17**, 1709–1713 (2003).
53. Qi, Z. et al. BMP restricts stemness of intestinal Lgr5^{+} stem cells by directly suppressing their signature genes. *Nat. Commun*. **8**, 13824 (2017).
54. Sansom, O. J. et al. Myc deletion rescues Apc deficiency in the small intestine. *Nature* **446**, 676–679 (2007).
55. Kim, T. H. et al. Broadly permissive intestinal chromatin underlies lateral inhibition and cell plasticity. *Nature* **506**, 511–515 (2014).
56. Skinnider, M. A. et al. Cell type prioritization in single-cell data. *Nat. Biotechnol*. **39**, 30–34 (2021).
57. Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* **4**, 44–57 (2009).

Acknowledgements We thank S. Alexandrova and members of the Simons laboratory for discussions, C. J. Hindley for reading the manuscript, S. Bae for assistance with graphic illustrations, the Gurdon Institute Imaging Facility for microscopy and image analysis support, and the Cancer Research UK Cambridge Institute Genomics Core Facility for sequencing and library generation. This work was supported by a Cancer Research UK Multidisciplinary Project Award to B.-K.K. and B.D.S. (C52767/A23363), and grants to A.P. (Cancer Research UK, A25636), J.-H.L. (Wellcome Trust, 107633/Z/15/Z; ERC Starting Grant, 679411), B.-K.K. (ERC Starting Grant, 639050; Human Frontier Science Program, RGY0071/2018; Interpark Bio-Convergence Center Grant Program, IBCC-IMBA-2020-01) and B.D.S. (Wellcome Trust, 098357/Z/12/Z and 219478/Z/19/Z), as well as by studentships and fellowships to J.F. (Wellcome Trust), S.-H.S.W. (DOC Fellowship of the Austrian Academy of Sciences), L.C. (Herchel Smith Fund), S.H. (Human Frontier Science Program, LT000092/2016-L; Basic Science Research Program NRF-2014R1A6A3A01005675) and B.D.S. (Royal Society EP Abraham Research Professorship, RP\R1\180165). The research team also acknowledges core grant support from the Austrian Academy of Sciences to the Institute of Molecular Biotechnology, the Wellcome Trust and MRC to the Wellcome–MRC Cambridge Stem Cell Institute, and the Wellcome Trust (092096) and CRUK (C6946/A14492) to the Gurdon Institute.

Author contributions M.K.Y., S.H., J.F., B.-K.K. and B.D.S. planned and designed the experiments. M.K.Y. performed lineage tracing, tissue imaging and quantification, flow cytometry and cell sorting, which were supervised by A.P., J.-H.L., B.-K.K. and B.D.S. B.-K.K. conceived the Red2Onco design and the Red2Onco mouse models were established by M.K.Y., S.H., J.F., C.D., T.T., R.M., J.-H.L. and B.-K.K. M.K.Y. and B.D.S. performed the quantitative statistical analysis of the clone size data. M.K.Y. and S.-H.S.W. performed organoid experiments. M.K.Y., L.C. and I.P. performed in situ hybridization experiments. C.D., L.C., R.A. and F.E. performed lineage tracing with oesophagus, stomach corpus, pancreas and lung tissue, respectively. E.L. and J.K.K. aligned raw sequencing data from scRNA-seq experiments. M.K.Y. and S.H. analysed scRNA-seq data, supervised by J.K.K., B.-K.K. and B.D.S. S.H. devised the algorithm for statistical analysis of transcriptomic changes with scRNA-seq data. S.-H.S.W. performed lineage tracing and tissue preparation of *LSL-Kras*^{G12D}, *PIK3CA*$^{Lat\text{-}H1047R}$ and *Apc*$^{fl/fl}$ mice, which were provided by D.E.S. M.K.Y., S.H., B.-K.K. and B.D.S. wrote the manuscript with input from all authors. These authors contributed equally: J.F., S.-H.S.W., C.D., T.T.

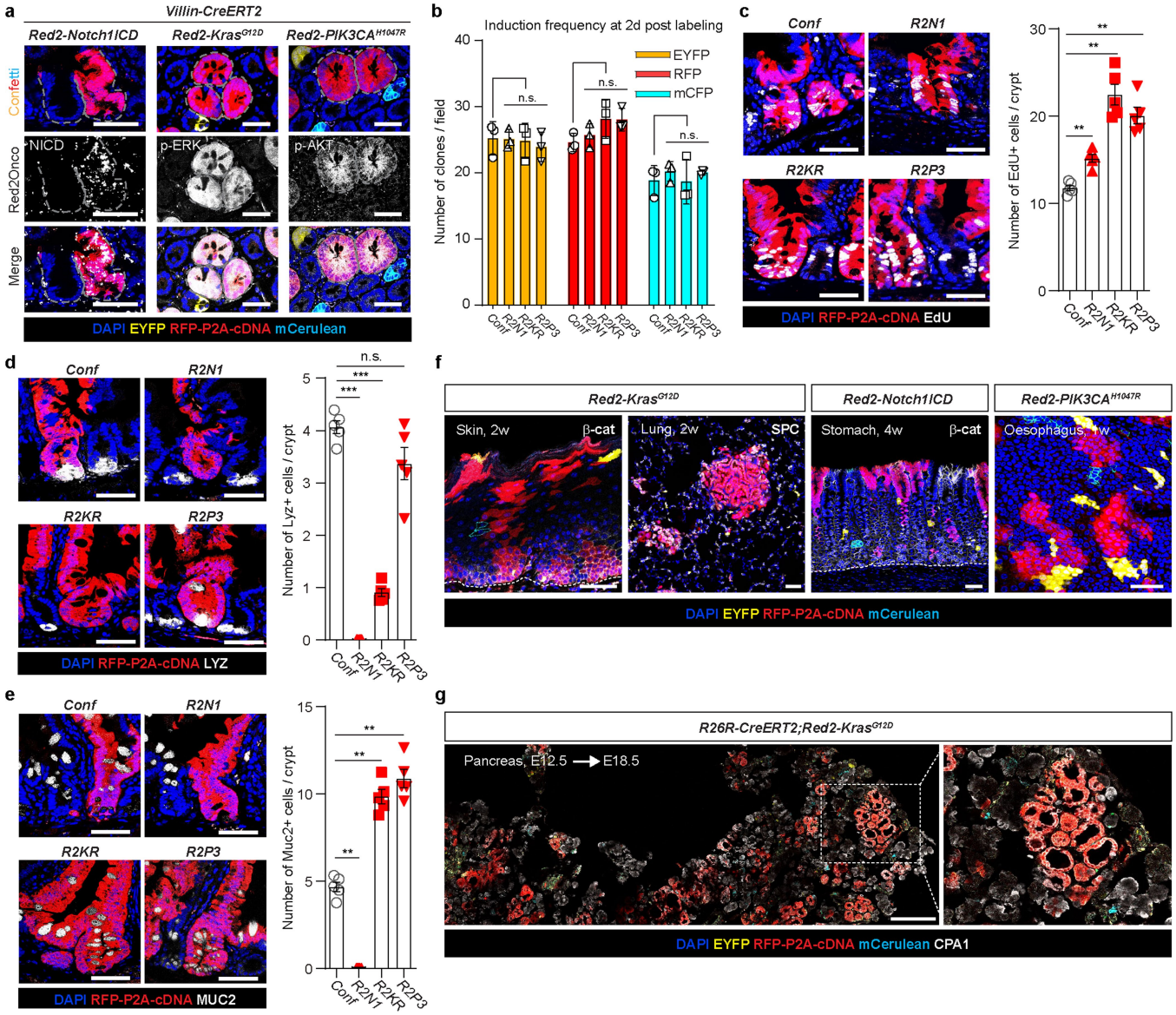
Competing interests The authors declare no competing interests.

Additional information

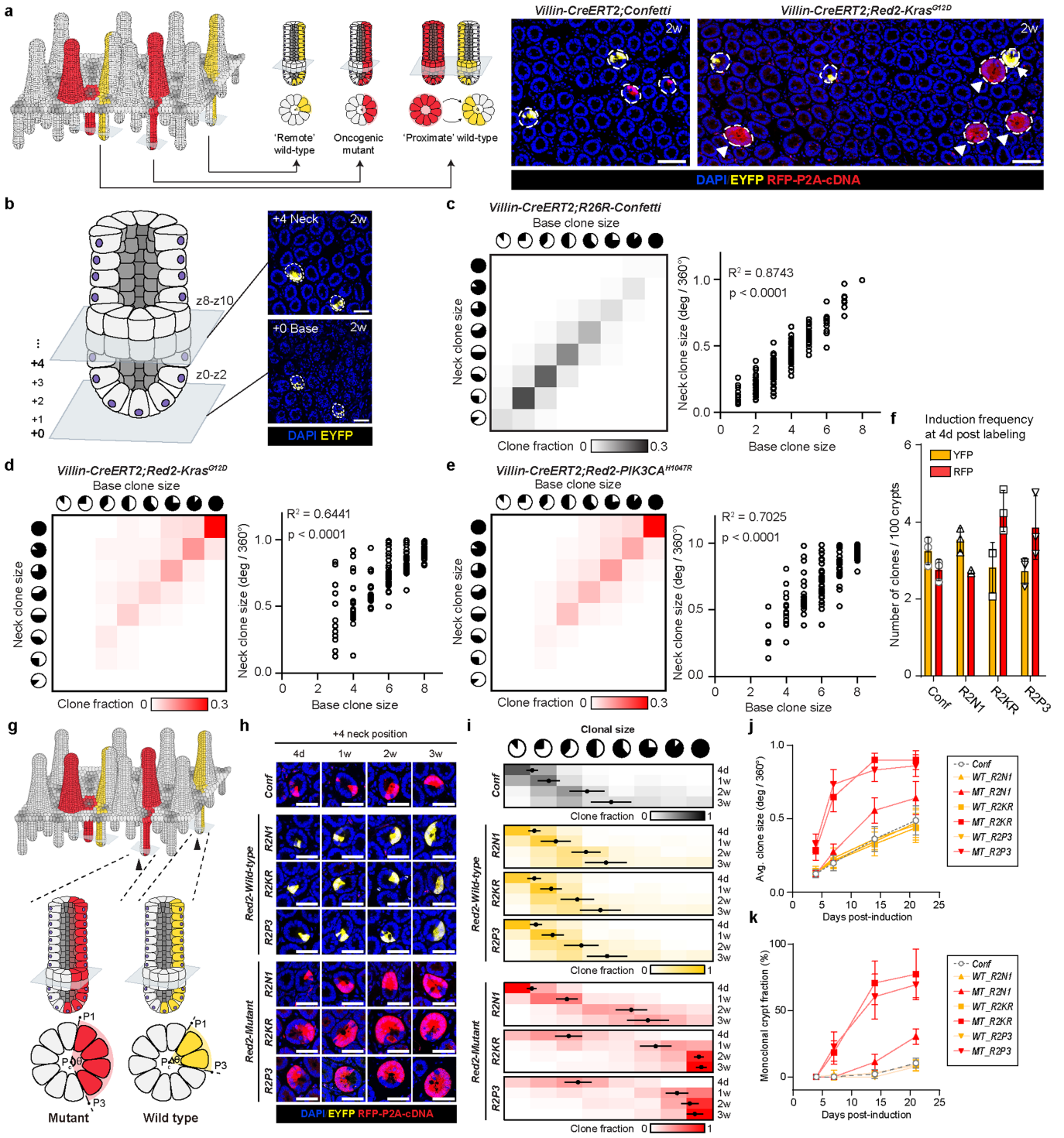
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03605-0.

Correspondence and requests for materials should be addressed to B.-K.K. or B.D.S.

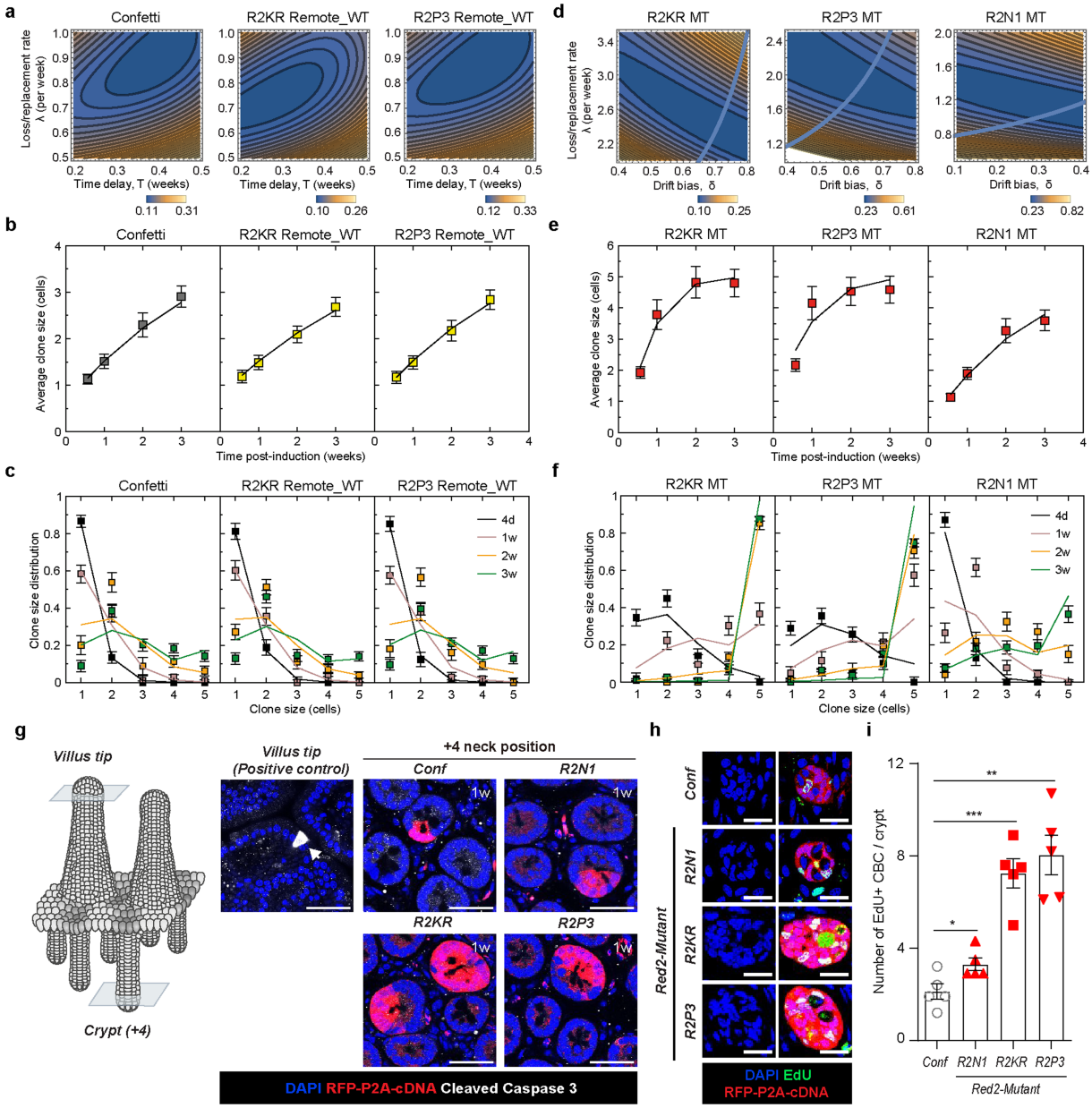
Peer review information *Nature* thanks Hans Clevers, James DeGregori, Dominic Grün and Toshiro Sato for their contribution to the peer review of this work. Peer reviewer reports are available.



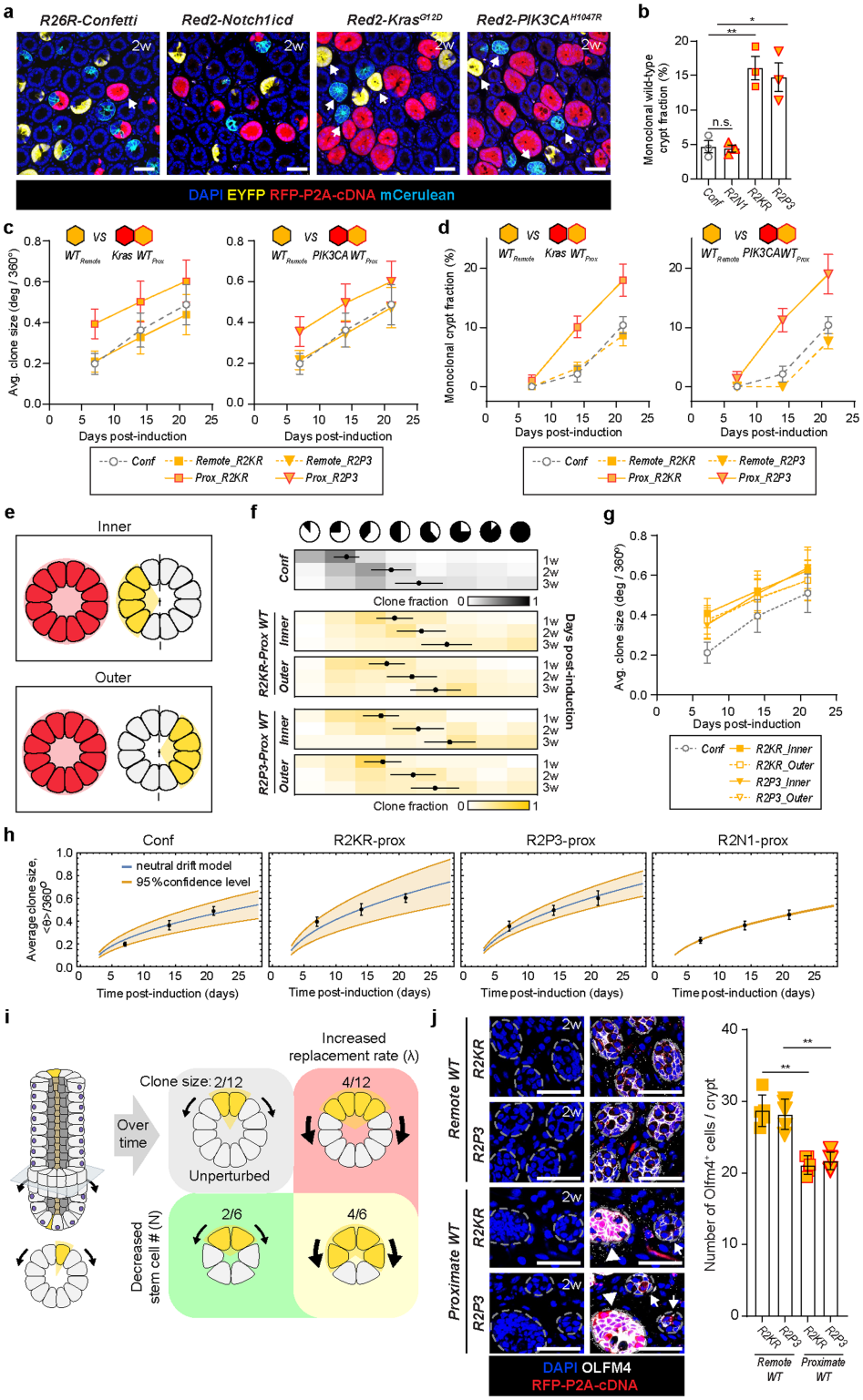
Extended Data Fig. 1 | Red2Onco system: an oncogene-associated multicolour reporter. a, Representative confocal images of mutant clones from sections (*Red2-Notch1*^{ICD}) or whole mounts (*Red2-Kras*^{G12D} and *Red2-PIK3CA*^{H1047R}) of small intestine from *Villin-CreERT2;Red2Onco* mice 2 w after tamoxifen administration. Crypt borders are marked with a grey dashed outline. **b**, Average clone numbers collected from a single field of image (0.15 mm^2) of whole-mount small intestine from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice 2 d after tamoxifen administration. **c**–**e**, Representative confocal images (left) and quantification (right) of EdU$^+$ proliferating crypt cells (**c**), LYZ$^+$ Paneth cells (**d**) and MUC2$^+$ goblet cells (**e**) from sections of small intestine from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice 4 w after tamoxifen administration. **f**, Representative confocal images of 100-µm-thick sections or whole mounts of tissues from adult *R26R-CreERT2;Red2Onco* mice (skin and stomach corpus), *Sftpc-CreERT2;Red2Onco* mice (lung) or *Krt5-CreERT2;Red2Onco* mice (oesophagus) at 1, 2 and 4 w after tamoxifen administration. White dashed line, epithelial lining. β-catenin stained as a cell membrane marker. SPC marks alveolar type II cells in lung. **g**, Representative confocal images of sectioned mouse embryonic pancreas tissue from the *R26R-CreERT2;Red2-Kras*^{G12D} mouse on embryonic day (E)18.5, 6 d after tamoxifen administration. Magnified panel to the right shows an example of acinar cell expansion in developing pancreas. CPA1 marks acinar cells. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$ by one-way ANOVA with Games-Howell's multiple comparisons test (**b**, **c**) or unpaired two-tailed *t*-test (**d**, **e**) from biological replicates. Data are mean ± s.d. (**b**) or mean ± s.e.m. (**c**–**e**). Exact *P* values are presented in Source Data. Scale bars: 50 µm (**a**, **c**–**f**) or 200 µm (**g**).



Extended Data Fig. 2 | Oncogenes drive non-neutral clone expansion in the mouse intestinal crypt. a, Schematic illustration of clonal events within the Red2Onco system (left) and representative tile scan images (right). Images are representative of tissues quantified in **h**–**k**. Arrow, WT crypt; arrowheads, fixed mutant crypts. **b**, Representative confocal images of the base and neck of crypts. Images are representative of tissues quantified in **c**. **c**–**e**, Strong correlation between clone size at the base and neck of a crypt from *Villin-CreERT2;R26R-Confetti* (2 w after tamoxifen administration) or *Red2Onco* mice (1 w after tamoxifen administration). **f**, Average number of clones per 100 crypts. **g**, Schematic illustration of mutant (RFP$^+$) and WT (YFP$^+$) clones in crypts, remote from each other. **h**, Representative confocal images at 4 d, 1 w, 2 w and 3 w after tamoxifen administration. Images are representative of tissues quantified in **i**–**k**. Red2-Wild-type: remote YFP$^+$ clones; Red2-Mutant: RFP$^+$ clones. **i**, Heat maps indicate the relative clone fractions of the indicated sizes (columns) at various time points after induction (rows). Data are mean ± s.e.m. **j**, **k**, Average clone size (**j**) and percentage of monoclonal crypts (**k**) at different time points after tamoxifen administration. Asterisks for statistical significance omitted in graphs in **j**, **k** for better visualization. Confocal images of small intestine are from *Villin-CreERT2;R26R-Confetti* (**a**, **b**, **h**) or *Red2Onco* mice (**a**, **h**). Crypt borders are marked with a white dashed outline (**a**, **b**). Significance (*P < 0.05; **P < 0.01; ***P < 0.0001) was determined by unpaired two-tailed Pearson's correlation test (**c**–**e**), one-way ANOVA with Games-Howell's multiple comparisons test (**j**) or unpaired two-tailed *t*-test (**k**) from biological replicates. Data are mean ± s.d. (**f**) or mean ± s.e.m. (**j**, **k**). For exact *P* values, see Source Data. Scale bars: 100 µm (**a**, **b**) or 50 µm (**h**).

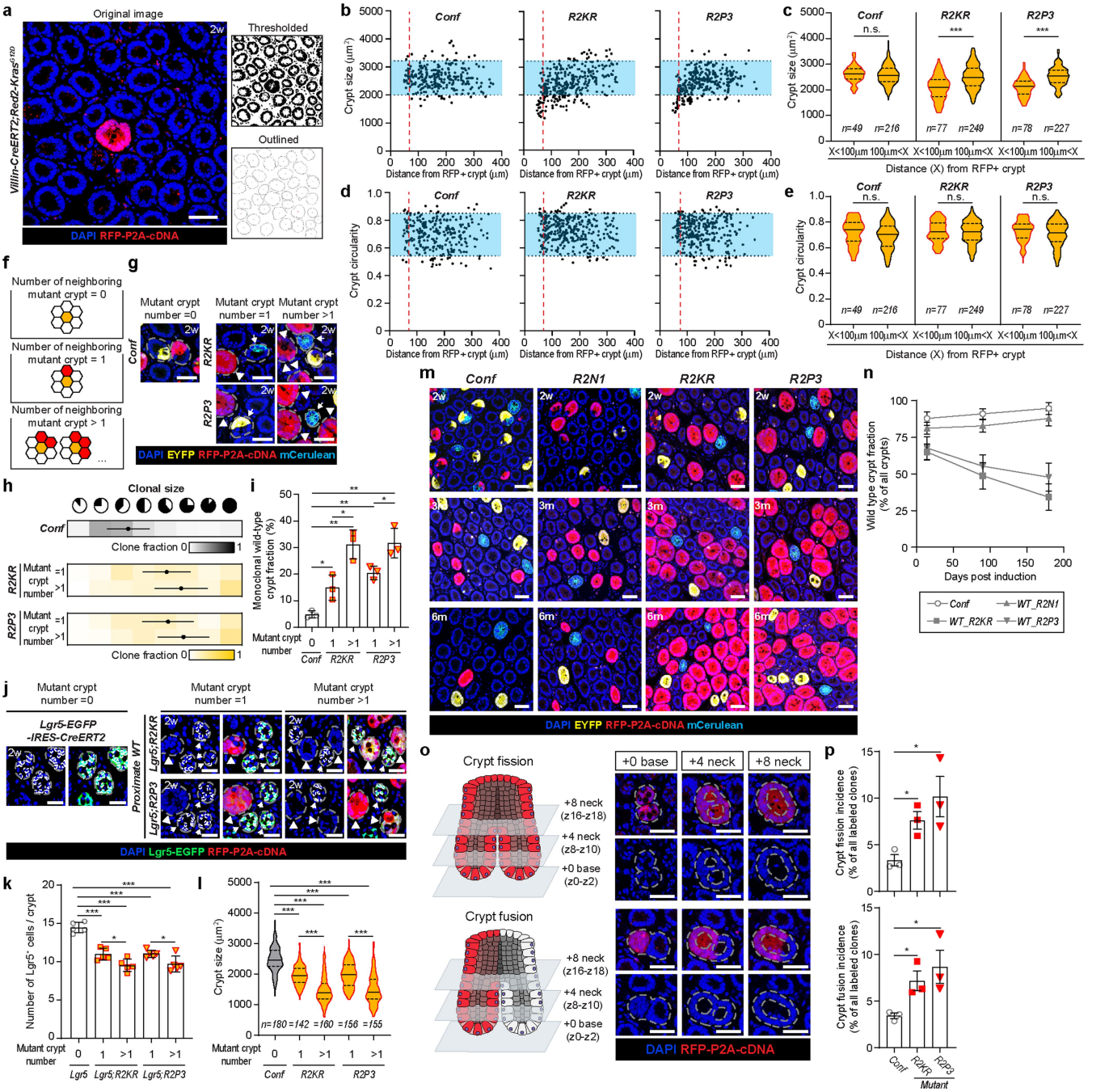


Extended Data Fig. 3 | Biophysical modelling of mutant clone expansion. **a**, **d**, Contour plots showing mean-square differences of clone size distribution between neutral drift model and YFP clone data from *Confetti* (left), WT crypts remote from mutant crypts in R2KR (middle) and R2P3 (right) in **a**; and between biased drift model and RFP mutant (MT) clone data from R2KR (left), R2P3 (middle) and R2N1 (right) in **d**. Plots: scan of loss/replacement rate λ against time delay between injection and induction in **a**; and drift bias δ with time delay of 0.29 w (R2KR), 0 w (R2P3), and 0.43 w (R2N1) in **d**. Blue lines in **d**, constraint $\lambda(1-\delta)=\lambda_{WT}$, where λ_{WT} is the loss/replacement rate inferred from *Confetti* (Supplementary Theory). Analysis in **a**, **d** based on data in **c**, **f**, respectively. **b**, **e**, Average clone size (effective stem cell number) from **a**, **d**, respectively. Points show data, lines show model prediction at optimal parameter values. In each case, total effective stem cell number $N=5$, so that an average clone size of, for example, 2 corresponds to circumferential angle of $360° \times 2/5$. **c**, **f**, Distribution of clone sizes for models from **a**, **d**, respectively. Points show data; lines show model prediction at optimal parameter values. **g**, Representative confocal images of cleaved caspase-3$^+$ apoptotic cells. A single cleaved caspase-3$^+$ apoptotic cell in the villus tip is indicated by the white arrow as a positive control. **h**, **i**, Representative confocal images (**h**) and quantification (**i**) of EdU$^+$ proliferating crypt base columnar cells. Whole mount of small intestine from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* 1 w (**g**) or 2 w (**h**, **i**) after tamoxifen administration. $^*P<0.05$, $^{**}P<0.01$, $^{***}P<0.0001$; unpaired two-tailed *t*-test (**i**) from biological replicates. Data are mean ± s.e.m. (**b**, **c**, **e**, **f**, **i**). Exact *P* values in Source Data. Scale bars: 50 µm (**g**) or 25 µm (**h**).

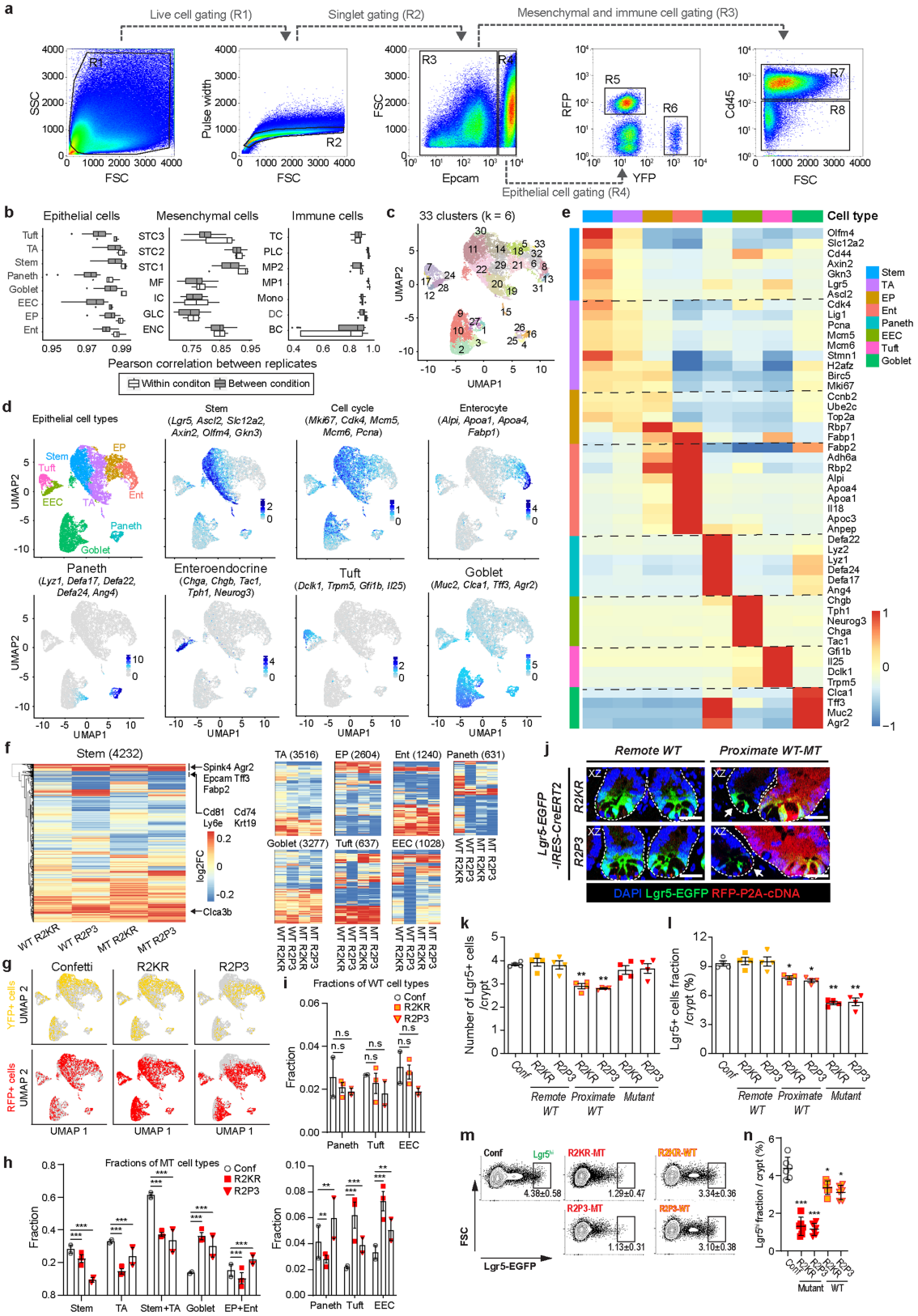


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Mutant crypts perturb clonal dynamics of WT cells in neighbouring crypts. a, Representative confocal images of tissues quantified in **b**. Fixed (monoclonal) WT crypts are indicated by white arrows. **b**, Percentage of monoclonal WT small intestinal crypts. **c**, **d**, Average clone size (**c**) and percentage of monoclonal crypts (**d**) of remote and proximate WT (YFP^+) clones at different time points after tamoxifen administration. **e**, Schematic illustration of proximate WT clones in relation to fixed mutant crypts. **f**, Heat maps indicate the relative clone fractions of the indicated sizes (columns) at various time points after induction (rows). Data are mean ± s.e.m. **g**, Average clone size of proximate (inner and outer) WT (YFP^+) clones at different time points after tamoxifen administration. **h**, Average clone size $\langle\theta\rangle/360°$ of WT (YFP^+) clones in crypts neighbouring fixed mutant crypts as a function of time t after induction. Data are mean ± s.e.m. Blue line shows a fit to the square root dependence predicted by the neutral drift model (Supplementary Theory). Orange line shows the 95% confidence interval. **i**, Schematic illustration of factors that affect rate of clonal drift (Supplementary Theory). **j**, Representative images (left) and quantification (right) of $OLFM4^+$ ISCs. Arrows, proximate WT crypts; arrowheads, fixed mutant crypts; grey dashed outlines, crypt borders. Confocal images of whole-mount small intestine are from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice (**a**) and *Lgr5-EGFP-IRES-CreERT2;Red2Onco* mice (**j**) 2 w after tamoxifen administration. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$; one-way ANOVA with Games-Howell's multiple comparisons test (**c**, **g**) and unpaired two-tailed t-test (**b**, **d**, **j**) from biological replicates. Data are mean ± s.e.m. (**b**–**d**, **f**–**h**) or mean ± s.d. (**j**). For exact P values, see Source Data. Asterisks for statistical significance omitted in graphs in **c**, **d**, **g** for better visualization. Scale bars, 50 µm (**a**, **j**).

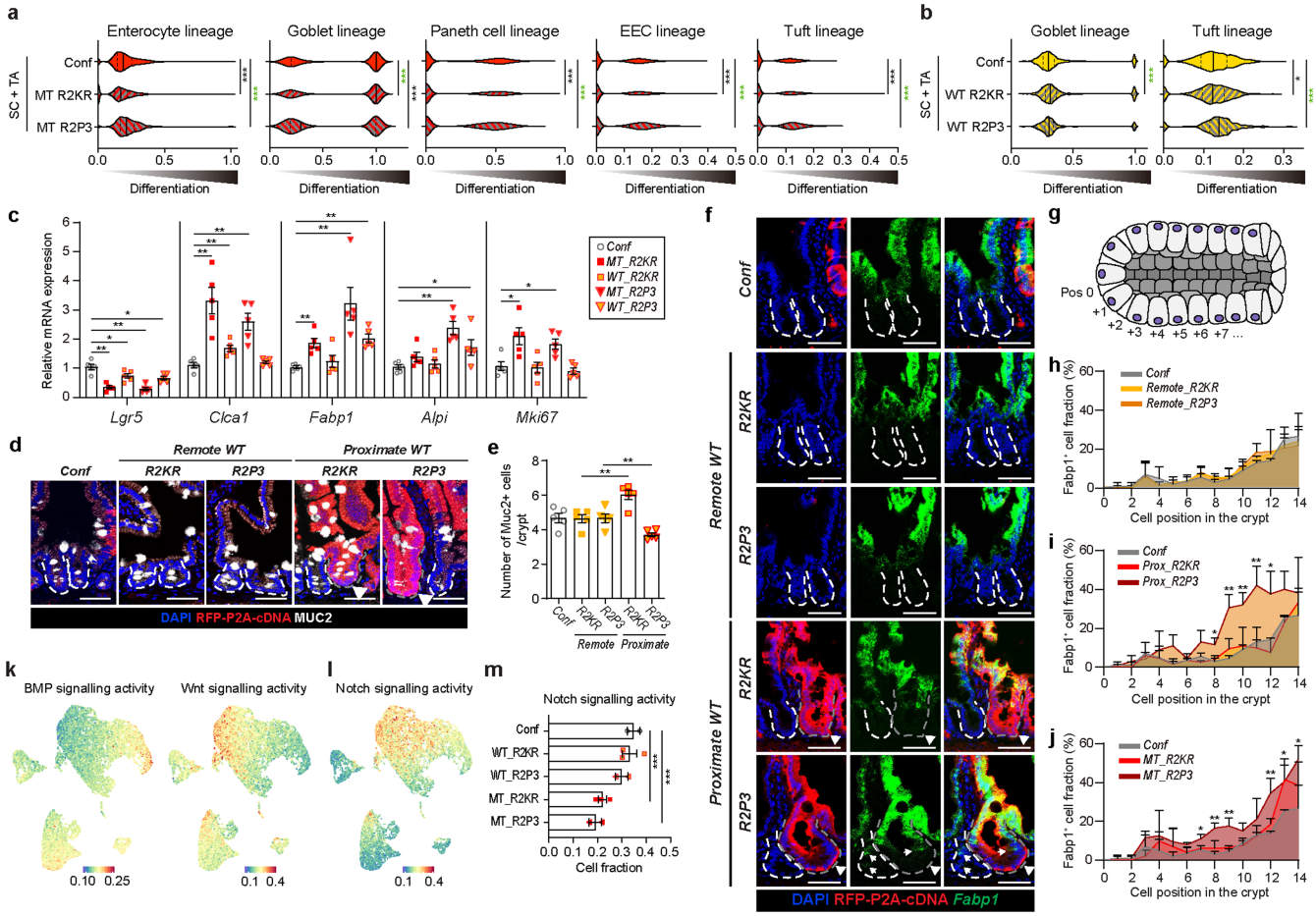


Extended Data Fig. 5 | Reduced effective stem cell number leads to accelerated drift dynamics. a, Original image (left) was thresholded (top right) and outlined (bottom right) to measure crypt size and circularity. Image representative of tissues quantified in **b**–**e**. **b**–**e**, Scatter and violin plots display size (**b**, **c**) and circularity (**d**, **e**) of WT crypts against distance from nearest fixed mutant (RFP$^+$) crypt. **f**, **g**, Illustration (**f**) and confocal images (**g**) of clones representative of tissues quantified in **h**, **i**. **h**, Heat maps indicate relative clone fractions of given sizes. Data are mean ± s.e.m. **i**, Percentage of monoclonal crypts of proximate WT (YFP$^+$) clones. **j**, **k**, Confocal images (**j**) and quantification (**k**) of EGFP$^+$ (LGR5$^+$) ISCs. Images representative of tissues quantified in **k**, **l**. White dashed lines, EGFP$^+$ cells in WT crypts. **l**, Violin plots display size of WT crypts in relation to multiplicity of neighbouring mutant crypts. *n* for each group is shown. **m**, **n**, Representative confocal images of *Red2Onco* intestine (**m**) and fractions of WT crypts from single field (0.15 mm^2; **n**). **o**, Illustration (left) and representative images (right) of crypt fission and fusion event in '8-shaped crypts'[30]. Images representative of tissues quantified in **p**. **p**, Percentage of crypts undergoing crypt fission (upper) or fusion (lower). Whole mount of small intestine from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice (**a**–**i**, **m**, **o**, **p**), and *Lgr5-EGFP-IRES-CreERT2;Red2Onco* mice (**j**, **k**) at indicated time points. Proximate WT crypts and fixed mutant crypts indicated by white arrows and arrowheads, respectively (**g**, **j**). Crypt borders marked by dashed grey outlines (**g**, **j**, **o**). In **b**, **d**, blue shaded area and red dashed line indicate 95% confidence interval of *R26R-Confetti* controls and average distance between the centre of fixed mutant crypt and proximate WT crypts, respectively. *$P < 0.05$, **$P < 0.01$, ***$P < 0.0001$; unpaired two-tailed *t*-test (**c**, **e**, **i**, **k**, **l**, **p**). Data are mean ± s.d. (**i**, **k**, **n**) or mean ± s.e.m. (**h**, **p**). For exact *P* values, see Source Data. Scale bars, 50 μm (**a**, **g**, **j**, **m**, **o**).

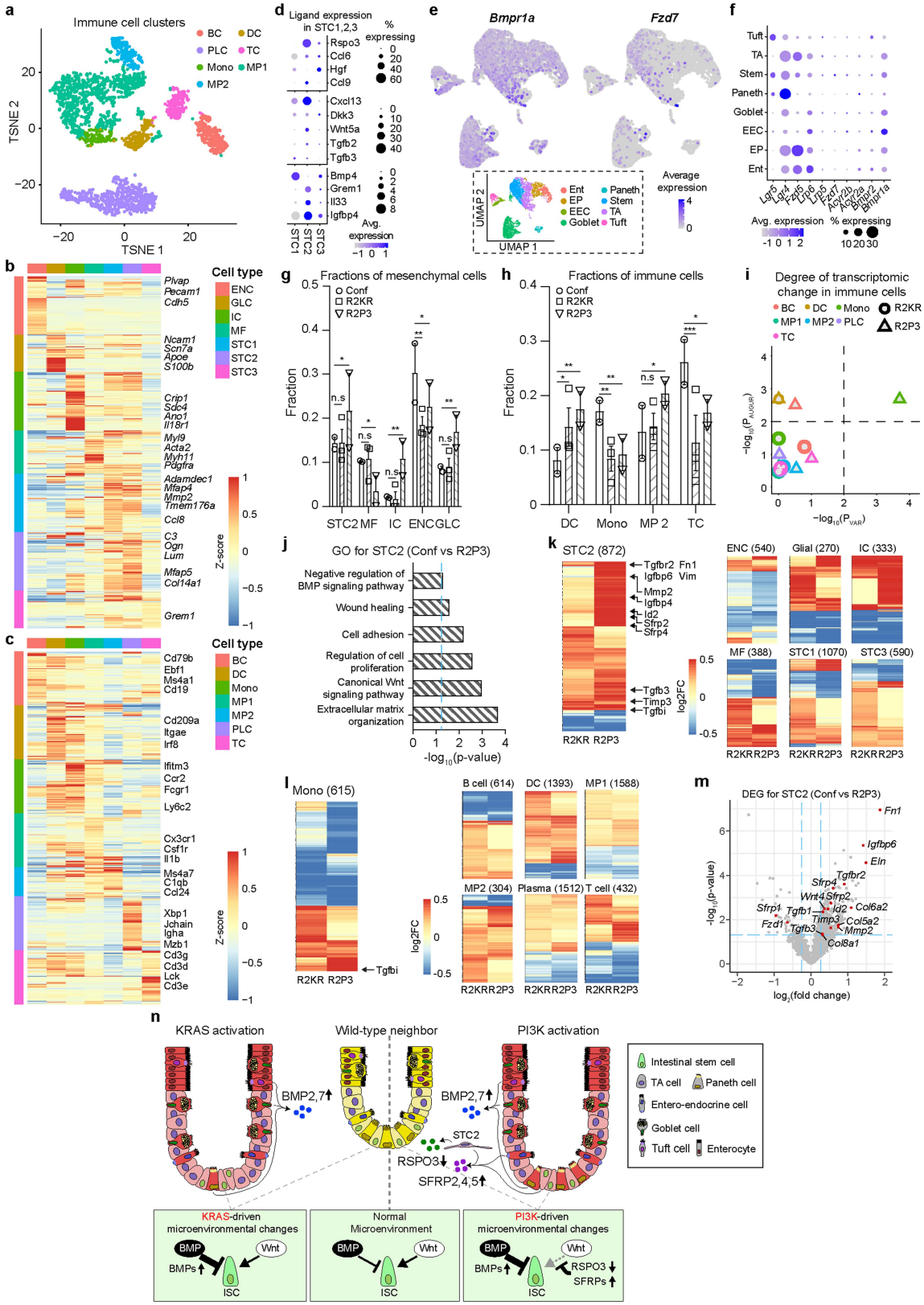


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Oncogene-driven signalling changes. a, FACS sorting strategy to isolate cells from *Confetti* and *Red2Onco* tissue. R1, live; R2, singlet; R3, mesenchymal/immune ($EPCAM^-$); R4, epithelial ($EPCAM^+$); R5, mutant-epithelial (RFP^+); R6, WT epithelial (YFP^+); R7, immune ($CD45^+$); R8, mesenchymal ($CD45^-$). **b**, Box plots showing distributions of Pearson correlation coefficients in averaged $\log_2$-transformed normalized UMIs for cell types across all pairs of mice from the same (white) and between different (grey) conditions. **c**, UMAP of epithelial cells detected by Louvain. *k*, *k* nearest-neighbour value. **d**, UMAPs showing distribution of averaged expression of marker genes. Colour bars, average $\log_2$-transformed normalized UMIs. Top left from Fig. 3b. **e**, Heat map representing marker expression for epithelial cells. Coloured panel (left) groups marker genes (right) for cell types. Colour bar, auto-scaled $\log_2$-transformed normalized UMIs. **f**, Heat maps representing differential gene expression for epithelial cells in *Red2Onco* compared to *Confetti*. Parentheses, number of differentially expressed genes. Colour bar, $\log_2$(fold change) (Supplementary Table 1). **g**, UMAPs showing distributions of mutant (RFP^+) and WT (YFP^+) epithelial cells for *Confetti* and *Red2Onco* mice. **h**, **i**, Fractions of mutant (**h**) and WT (**i**) epithelial cells in *Red2Onco* and *Confetti* mice. See Fig. 3c for other WT data. **j**–**l**, Confocal images (**j**) of $EGFP^+$ cells, representative of tissues quantified for stem cell number (**k**) and fraction (**l**). In **j**, white arrows indicate WT crypts proximate to mutant (MT) crypts. White dashed lines mark crypts. Scale bars, 25 µm. **m**, **n**, FACS plots (**m**) and quantification (**n**) of $EGFP^{hi}$ stem cell fractions from R5 or R6 (**a**). Small intestine from *Lgr5-EGFP-IRES-CreERT2;Red2Onco* 2 w after induction (clonal dosage (0.2 mg per 20 g body weight) for **j**–**l**, mosaic dosage (4 mg per 20 g body weight) for **m**, **n**). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$; n.s., statistically not significant, $P > 0.05$; two-sided Kolmogorov–Smirnov test (**b**), two-sided likelihood ratio test (**h**, **i**) and one-way ANOVA with Games-Howell's multiple comparisons test (**k**, **l**, **n**). Data are mean ± s.e.m. (**h**, **i**, **k**, **l**) or mean ± s.d. (**n**). For exact *P* values, see Source Data.

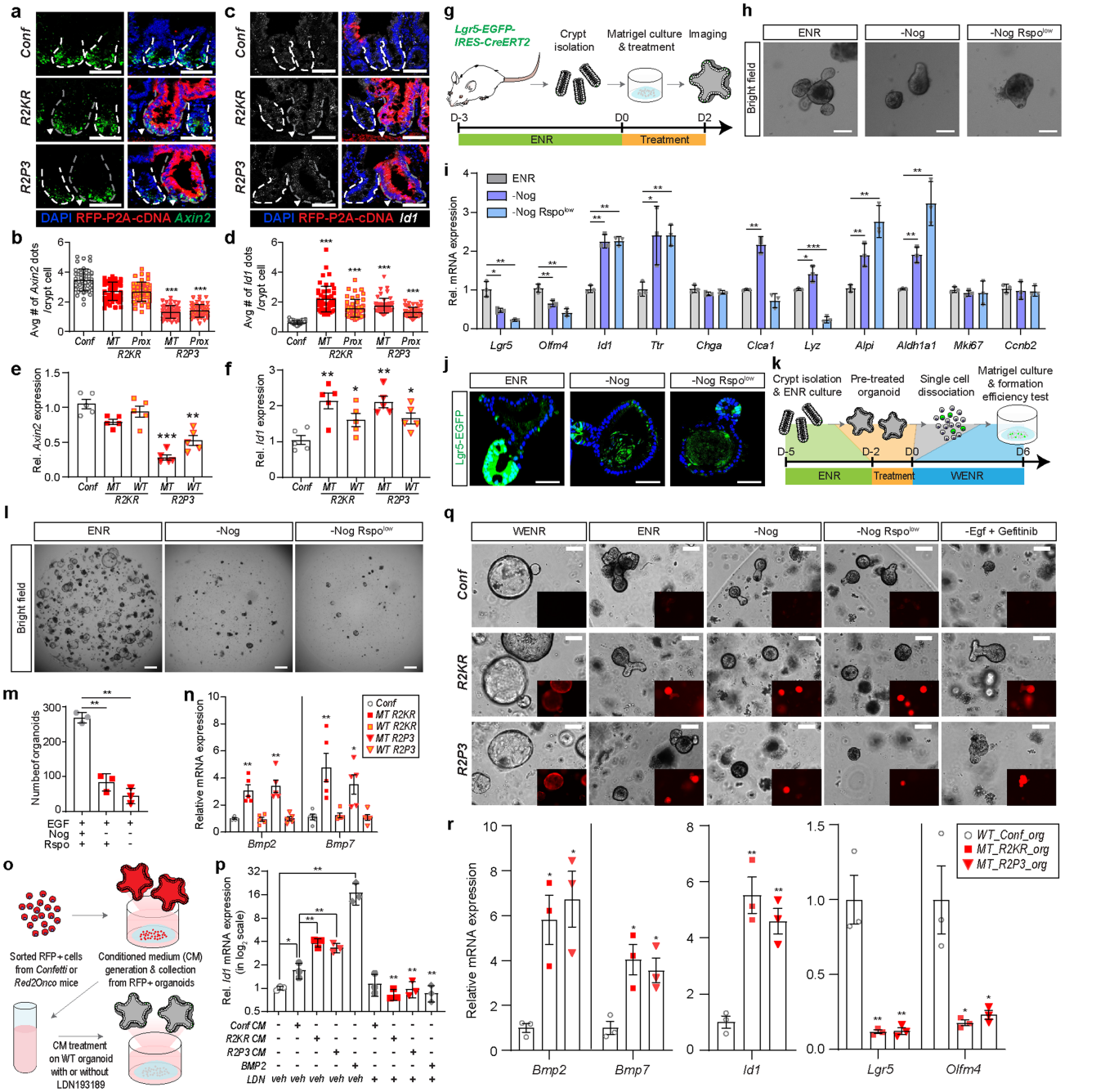


Extended Data Fig. 7 | Mutant crypt induces primed differentiation. **a**, **b**, Priming scores of stem (SC) and transit-amplifying (TA) cells of mutant (**a**) and WT (**b**) crypts towards secretory and enterocyte lineages in *Red2Onco* and *Confetti* mice. Black line, 50th percentile; dashed lines, 25th and 75th percentiles. Green and black asterisks, higher and lower in *Red2Onco* than in *Confetti*, respectively. **c**, qPCR of lineage markers (*Lgr5*, ISC; *Clca1*, goblet cell; *Fabp1*, *Alpi*, enterocyte; *Mki67*, proliferation) using sorted RFP$^+$ and YFP$^+$ cells from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice 2 w after tamoxifen administration. **d**, **e**, Confocal images (**d**) and quantification (**e**) of MUC2$^+$ goblet cells. **f**, **h**–**j**, Images (**f**) from RNA in situ hybridization of enterocyte marker *Fabp1* and quantification in remote WT (Remote_R2KR, Remote_R2P3; **h**), proximate WT (Prox_R2KR, Prox_R2P3; **i**) and mutant crypts (MT_R2KR or MT_R2P3; **j**) along crypt axis. In **f**, *Fabp1*$^+$ cells in lower crypts (below +8) marked by white arrow. **g**, Illustration of cellular localization along crypt axis. Position 0 is crypt base cell. **k**, **l**, UMAPs showing distributions of enrichment scores for BMP (**k**, left), WNT (**k**, right), and NOTCH (**l**) pathways in epithelial cells of *Red2Onco* and *Confetti* mice. Colour bars, enrichment scores. **m**, Fractions of 'active' cells with high enrichment scores for NOTCH pathway in mutant (MT) and WT epithelial cells from *Red2Onco* and *Confetti* mice. Small intestine sections from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* 2 w after tamoxifen administration (**d**, **f**). WT and mutant crypts marked with white and grey dashed outlines, respectively (**d**, **f**). Remote WT in crypts separated by >3 crypt diameters from mutant crypts. Proximate WT in crypts neighbouring fixed mutant crypts. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$; two-sided Kolmogorov–Smirnov test (**a**, **b**), unpaired two-tailed *t*-test (**c**, **e**, **h**–**j**) and two-sided likelihood ratio test (**m**). Data are mean ± s.e.m. (**c**, **e**, **m**) or mean ± s.d. (**h**–**j**). For exact *P* values, see Source Data. Scale bars, 50 µm (**d**, **f**).

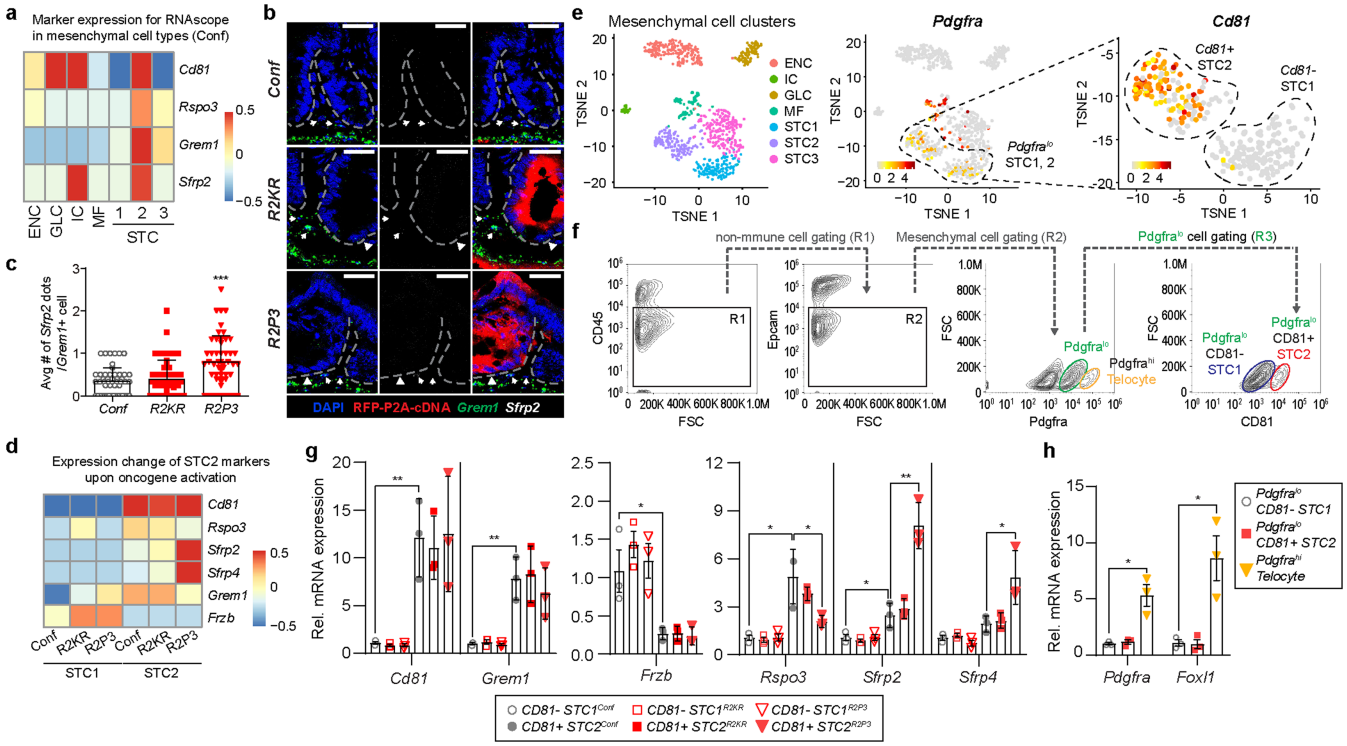


Extended Data Fig. 8 | See next page for caption.

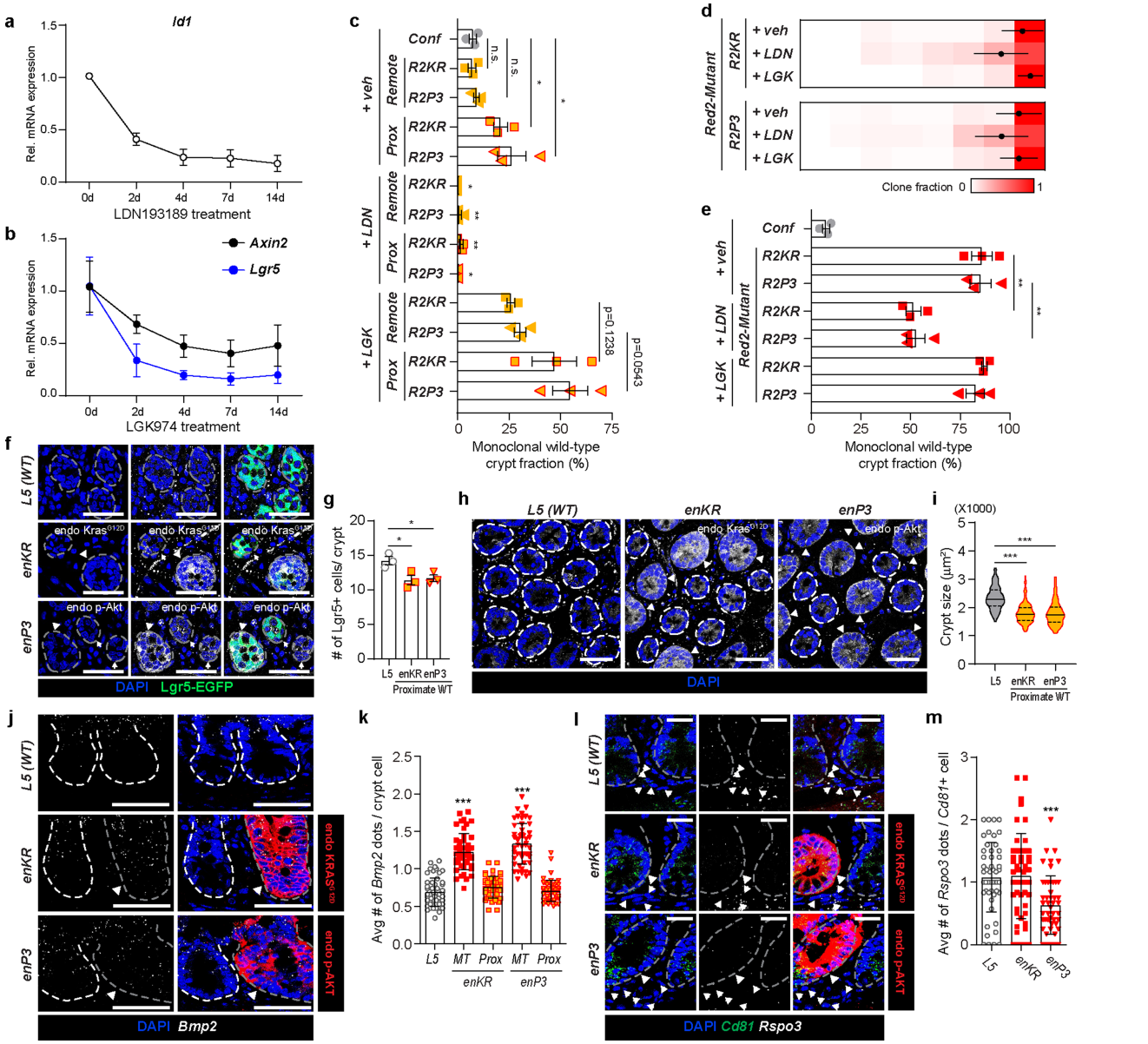
Extended Data Fig. 8 | Mutation-induced environmental changes. a, t-SNE representing immune cells from *Confetti* and *Red2Onco* mice. **b**, **c**, Heat maps representing differential expression (DE) patterns for mesenchymal (**b**) and immune (**c**) cells from *Red2Onco* and *Confetti* mice: top 300 genes or less (FDR < 0.05, pairwise *t*-test). Colour bar, averaged *Z*-scores of $\log_2$-transformed normalized UMIs. **d**, Secretion factor expression in stromal clusters for *Confetti*. In **d**, **f**, dot size denotes percentage of cells expressing gene; colour shows average expression. **e**, UMAPs showing expression of *Bmpr1a* and *Fzd7* in epithelial cells. Colour bar, $\log_2$-transformed normalized UMIs. Inset from Fig. 3b. **f**, Dot plots showing expression of receptors upstream of BMP and WNT pathways for epithelial cells. **g**, **h**, Fractions of mesenchymal (**g**) and immune (**h**) cells in *Red2Onco* and *Confetti* mice. Data are mean ± s.e.m. Two-sided likelihood ratio test: $^*P < 0.05$, $^{**}P < 0.01$; n.s., statistically not significant ($P > 0.05$). **i**, Degree of transcriptomic change for immune cells estimated by cell-to-cell variability (P_{VAR}) or separability of perturbed and unperturbed cells (P_{AUGUR}). Dotted lines show $-\log_{10}(0.01)$. Dot colour denotes cell type; dot shape shows *Red2Onco*. **j**, Enriched biological processes from gene ontology (GO) analysis of DE genes in STC2 of *Red2-PIK3CA*^{H1047R} mice relative to *Confetti*. One-sided Fisher's exact test. Dotted line, $-\log_{10}(0.05)$. **k**, **l**, Heat maps representing DE genes and their numbers (in parentheses) for mesenchymal (**k**) and immune (**l**) cells in *Red2Onco* mice compared to *Confetti*. Colour bar, $\log_2$(fold change) (Supplementary Table 3). **m**, Volcano plot representing DE genes in STC2 of *Red2-PIK3CA*^{H1047R} mice relative to *Confetti*. Two-sided pairwise *t*-test. Red dots, genes for biological processes in **j**. Vertical dotted lines, absolute value of $\log_2$(fold change) = 0.259; horizontal, $-\log_{10}(0.05)$. **n**, Model of direct and indirect cross-talk between mutant and WT crypts in *Red2Onco* mice. BC, B cell; DC, dendritic cell; Mono, monocyte; MP1, 2, macrophage 1, 2; PLC, plasma cell; TC, T cell. For exact *P* values, see Source Data.



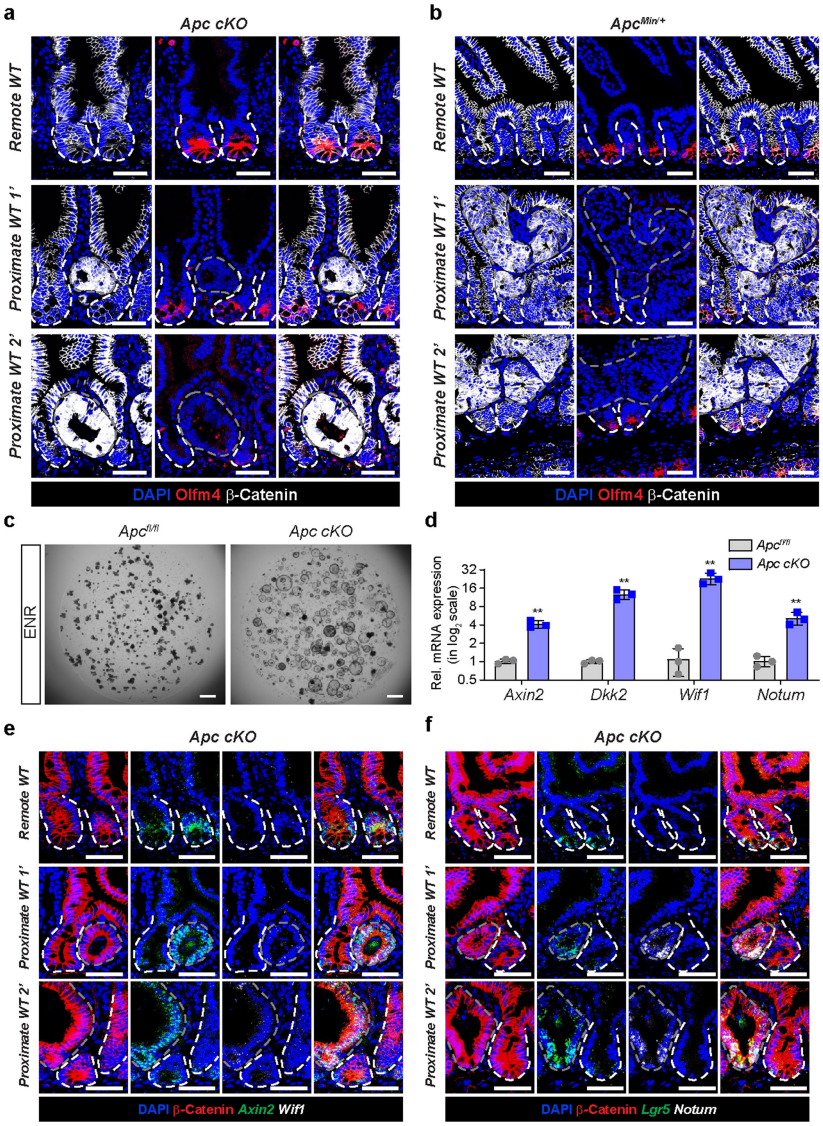
Extended Data Fig. 9 | Mutant clones secrete functional BMP ligands. **a**–**d**, Representative in situ hybridization images and quantification of *Axin2* (**a**, **b**) and *Id1* (**c**, **d**) on sections of small intestine from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice 2 w after tamoxifen administration. Arrowheads, fixed mutant crypts. Crypts are marked with dashed outlines. **e**, **f**, qPCR analysis of *Axin2* (**e**), and *Id1* (**f**). **g**, Experimental set-up for **h**–**j**. **h**, Bright-field images of intestinal organoids after 2 days of treatment. The number and size of crypt-like budding structures are reduced in treated organoids. **i**, qPCR analysis of lineage markers. **j**, Representative images of LGR5–EGFP organoids show that the number of LGR5$^+$ cells decreases following treatment. **k**, Experimental set-up for **l**, **m**. **l**, **m**, Bright-field images (**l**) and quantification (**m**) of intestinal organoids after 6 days of culture in WENR medium. **n**, qPCR analysis of BMP ligands (*Bmp2* and *Bmp7*). **o**, Experimental set-up for **p**. **p**, qPCR analysis of *Id1* using WT organoids after the CM treatment. **q**, Bright-field images of intestinal organoids from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice 1 month after tamoxifen administration. Insets, RFP expression in the mutant organoids. **r**, qPCR analysis of WT and mutant organoids cultured in ENR medium. In **e**, **f**, **n**, sorted RFP$^+$ or YFP$^+$ cells from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice 2 w after tamoxifen administration (4 mg per 20 g body weight, mosaic dosage) were analysed. *$P < 0.05$, **$P < 0.01$, ***$P < 0.0001$; one-way ANOVA with Games-Howell's multiple comparisons test (**b**, **d**) and unpaired two-tailed *t*-test (**e**, **f**, **i**, **m**, **n**, **p**, **r**). Quantification graphs show data from three independent experiments (**i**, **m**, **p**, **r**). Data are mean ± s.d. (**b**, **d**, **i**, **m**, **p**) or mean ± s.e.m. (**e**, **f**, **n**, **r**). For exact *P* values, see Source Data. Scale bars: 50 µm (**a**, **c**, **h**), 100 µm (**j**, **q**) and 500 µm (**l**).



Extended Data Fig. 10 | Mutant clones drive niche stromal remodelling. **a**, Heat map showing marker gene expression for STC2 among mesenchymal cells. Colour bar, averaged *Z*-scores of $\log_2$-transformed normalized UMIs over all cells within a cell type in *Confetti* mice. **b**, **c**, Representative multiplexed in situ hybridization images (**b**) and quantification (**c**) of *Sfrp2* in *Grem1*$^+$ cells on small intestine sections from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice 2 w after tamoxifen administration. Arrowheads, fixed mutant crypts; grey dashed outlines indicate crypts; arrows, *Grem1*$^+$ STC2 cells. **d**, Heat map showing expression of marker genes and secreted factors in STC1, 2 from *Red2Onco* and *Confetti* mice. Colour bar, averaged *Z*-scores of $\log_2$-transformed normalized UMIs over all cells within a cell type and condition. **e**, Projection of *Pdgfra* expression (middle) onto UMAP from Fig. 3f (left) for comparison. Projection of *Cd81* expression onto *Pdgfra*^{lo} cell clusters (STC1, 2) (right). Colour bar, $\log_2$-transformed normalized UMIs. **f**, Sorting strategy to isolate STC2 from intestinal mesenchymal cells by FACS. R1, non-immune cells (CD45$^-$); R2, mesenchymal cells (EPCAM$^-$); R3, PDGFRA^{lo} population. **g**, qPCR of STC2 markers (*Cd81*, *Grem1*), STC1 marker (*Frzb*) and secreted WNT modulators (*Rspo3*, *Sfrp2*, *Sfrp4*) using sorted CD45$^-$EPCAM$^-$PDGFRA^{lo}CD81$^-$ cells (STC1) or CD45$^-$EPCAM$^-$PDGFRA^{lo}CD81$^+$ cells (STC2) from *Villin-CreERT2;R26R-Confetti* or *Red2Onco* mice 2 w after tamoxifen administration. **h**, qPCR of telocyte markers (*Pdgfra*, *Foxl1*) using sorted STC1, 2 and PDGFRA^{hi} telocytes. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$; one-way ANOVA with Games-Howell's multiple comparisons test (**c**) and unpaired two-tailed *t*-test (**g**, **h**). Data are mean ± s.d. (**c**, **g**) or mean ± s.e.m. (**h**). For exact *P* values, see Source Data. Scale bars, 25 µm (**b**).



Extended Data Fig. 11 | Functional validation of oncogene-driven niche remodelling. a, **b**, qPCR analysis of *Id1* (**a**), *Axin2* and *Lgr5* (**b**) after administration of indicated inhibitor. **c**, Fraction of monoclonal WT (YFP$^+$) crypts remote from (Remote) or proximate to (Prox) mutant crypts in *Red2Onco* mice. **d**, Heat maps indicate relative clone fractions of the indicated sizes. Data are mean ± s.e.m. **e**, Fraction of monoclonal (RFP$^+$) mutant crypts in *Red2Onco* mice. **f**, **g**, Representative confocal images (**f**) and quantification (**g**) of EGFP$^+$ (LGR5$^+$) ISCs. Images are representative of tissues quantified in **g**. Arrows, proximate WT crypts; arrowheads, fixed mutant crypts. **h**, Representative confocal images of whole-mount small intestine. Images are representative of tissues quantified in **i**. Arrowheads, fixed mutant crypts. **i**, Violin plots of proximate WT crypt size. **j**, **k**, RNA in situ hybridization (**j**) and quantification (**k**) of *Bmp2*. Arrowheads, fixed mutant crypts. **l**, **m**, Representative multiplexed in situ hybridization images (**l**) and quantification (**m**) of *Rspo3* in *Cd81*$^+$ cells. Arrow: *Cd81*-positive STC2 cells; Arrowheads, fixed mutant crypts. Whole mount (**f**–**i**) and sections (**j**–**m**) of small intestine from *Lgr5-EGFP-IRES-CreERT2* control (L5), *Lgr5-EGFP-IRES-CreERT2;LSL-Kras^{G12D}* (enKR) or *PIK3CA$^{Lat\text{-}H1047R}$* (enP3) mice 2 w after tamoxifen administration. In **c**–**e**, graphs show data collected 2 w after concomitant administration of indicated drug and tamoxifen. Crypt borders are marked by dashed outlines (**f**, **h**, **j**, **l**). In **f**, **h**, **j**, **l**, white (**f**, **h**) or red (**j**, **l**): immunostaining for mutant KRAS(G12D) in enKR, or p-AKT in enP3. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$; one-way ANOVA with Games-Howell's multiple comparisons test (**k**, **m**) and unpaired two-tailed *t*-test (**c**, **e**, **g**, **i**). Data are mean ± s.d. (**a**, **b**, **g**, **k**, **m**) or mean ± s.e.m. (**c**, **e**). For exact *P* values, see Source Data. Scale bars: 50 µm (**f**, **h**, **j**) and 25 µm (**l**).



Extended Data Fig. 12 | *Apc* mutation induces reduction of stem cells in neighbouring wild-type crypts. a, **b**, Representative confocal images of small intestine from *Villin-CreERT2;Apc*$^{f/f}$ mice at 2 w after tamoxifen administration (**a**) and from *Apc*$^{Min/+}$ mice at 12 weeks of age (**b**). Images are representative of two independent experiments. OLFM4 staining shows a reduced number of stem cells in wild-type crypts neighbouring mutant crypts. Grey dashed outlines, *Apc* mutant foci (*Villin-CreERT2;Apc*$^{f/f}$) or polyps (*Apc*$^{Min/+}$); white dashed outlines, crypt borders. Scale bars, 50 µm. **c**, Bright-field images of intestinal organoids after 7 days of culture in ENR medium. Images are representative of three independent experiments. Note that organoids from *Villin-CreERT2;Apc*$^{f/f}$ mice form spheroids in ENR medium. Scale bars, 500 µm. **d**, qPCR analysis of WNT target gene (*Axin2*) and secreted WNT inhibitory factors (*Dkk2*, *Wif1* and *Notum*) following *Apc* deletion. Data are mean ± s.d. $n = 3$ independent experiments. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$; unpaired two-tailed *t*-test. **e**, **f**, Representative multiplexed in situ hybridization images of *Axin2* and *Wif1* (**e**) and *Lgr5* and *Notum* (**f**) on sections of small intestine from *Villin-CreERT2;Apc*$^{f/f}$ mice 2 w after tamoxifen administration. Images are representative of two independent experiments. *Axin2* (**e**) and *Lgr5* (**f**) staining shows a reduced number of stem cells in wild-type crypts neighbouring *Apc* mutant crypts. Grey dashed outlines, *Apc* mutant foci (*Villin-CreERT2;Apc*$^{f/f}$); white dashed outlines, crypt borders. Scale bars, 50 µm.

nature research

Corresponding author(s): Benjamin D. Simons

Last updated by author(s): Apr 2, 2021

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 our Editorial Policies 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	The confocal (Leica SP5) images generated were merged and displayed using the LAS software (v2.8.0, Leica). Flow cytometry cell sorting was performed on a MOFLO system using the Summit software (v5.2, Beckman Coulter) or on a SH800S Cell Sorter using SH800 software (v2.1.5, Sony). Organoids were imaged and counted using an EVOS M7000 microscope using EVOS M7000 Revision software (v2.0, Thermo Scientific). Realtime PCR was performed on StepOnePlus™ Real-Time PCR System (Thermofisher).
Data analysis	Confocal images were processed and analyzed using ImageJ (1.52e). Statistical analysis were performed using GraphPad Prism (8.00) unless otherwise specified. Flow cytometry data analysis were performed using a FACSAria sorter using FlowJo software (v10.6.2, BD). The mathematical modelling and simulation based on clonal data was carried out using FORTRAN (G95 compiler). The analysis of single-cell RNA-seq data was performed using publicly available softwares and R packages as follows. The raw data from single-cell RNA-seq experiment using the 10X Genomics platform were processed using CellRanger (v2.1.1, https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest). CellRanger aligned reads, filtered empty dropouts and counted unique molecular identifiers (UMIs) to generate a count matrix. Ensembl GRCm38/mm10 (release 92) appended with the sequences for tdimer2 (RFP) and EYFP was used as the reference genome for the read alignment. The filtration of empty droplets was checked with R package DropletUtils (v1.2.2, https://bioconductor.org/packages/release/bioc/html/DropletUtils.html). UMIs were normalised using a deconvolution method in R package scran (v1.12.1, https://bioconductor.org/packages/release/bioc/html/scran.html). Dimension reduction of normalised UMI data was performed by PCA combined with technical noise modeling which was implemented in denoisePCA() function in R package scran. The data were projected using two-dimensional Uniform Manifold Approximation and Projection (UMAP) or t-Distributed Stochastic Neighbor Embedding (t-SNE) with default parameter setting. Based on gene expression, cells were clustered using a graph-based method which was implemented by buildSNNGraph function in R package scran and walktrap method from R package igraph (v1.2.4.1, https://igraph.org/r/) with default option (steps = 4). After annotating cell types based on marker gene expression, the fraction of cell types for all conditions were calculated. The fractional changes between conditions were tested using R package stat (v3.6.0, http://www.R-project.org/). To evaluate the degree of priming of stem cells toward sub-lineages, the fate bias of stem cells and their progenitors was calculated using R package FateID (v0.1.9, https://github.com/dgrun/FateID). The distributions of priming cores were compared between conditions using Kolmogorov-Smirnov test (ks.test function) in R package stats (v3.6.0). To detect signaling pathways contributing stem cell priming, gene set enrichment analysis was performed using R

package AUCell (v1.6.1, https://bioconductor.org/packages/release/bioc/html/AUCell.html). To calculate the degree of transcriptome change for each cell type, Augur (v1.0.2, https://github.com/neurorestore/Augur) was used. To understand transcriptional change caused by oncogene expression, gene ontology analysis was performed using DAVID (v6.8, https://david.ncifcrf.gov/). The custom codes and data used for clonal analysis and single-cell RNA-seq analysis are deposited in GitHub (https://github.com/BenSimonsLab/Yum_Nature_2021).

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 and 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 single-cell RNA-seq data we generated for this study has been deposited in ArrayExpress under E-MTAB-8656. The reference genome sequence was downloaded from the Ensembl (http://www.ensembl.org/Mus_musculus) and used for alignment of the single-cell RNA-seq data. To evaluate stem cell priming, single-cell RNA-seq data was downloaded from the Single Cell Portal (https://portals.broadinstitute.org/single_cell/study/small-intestinal-epithelium) and used to define gene sets for differentiated sub-lineages of epithelial cells. The lists of marker genes used to annotate types of epithelial, mesenchymal, and immune cells in Fig. 3b, f and Extended Data Fig. 6d, 8a, b, c are given in Supplementary Tables 1 and 3. Gene sets used in Fig. 3d, e, and Extended Data Fig. 7a, b, k, l, m are provided in Supplementary Table 2. All source data supporting the findings of this study are provided together with corresponding figures.

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

Life sciences study design

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

Sample size	No statistical methods were used to predetermine sample size for experimentation. Sample size was determined based on our experience and the known literatures (n=3-6). (DOI: 10.1126/science.1196236. and DOI: 10.1016/j.cell.2010.09.016)
Data exclusions	No data were excluded from the analysis
Replication	Experimental data was replicated using at least 2 independent biological samples (Precisely stated in the figure legends) with similar results. For organoid experiments, indepedndent experiments refer to fully independent cultures starting from different animal. Single cell mRNAseq experiments were replicated using at least 2 independent biological samples per condition.
Randomization	Based on their genotype, randomized cohorts including both male and female animals were distributed in an unblinded manner into the experimental time points for analysis
Blinding	The investigators were not blinded to allocation during experiments and outcome assessment. Blinding was not possible as the mice had to be genotyped by PCR before the analysis. Computational analysis required proper identification of sources, and this was also not suitable for blinding.

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 and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used

Goat polyclonal anti-CPA1 R and D Systems Cat# AF2765 1:200
Rabbit monoclonal anti-OLFM4 Cell Signaling Cat# 39141 1:100
Rabbit monoclonal anti-Phospho-p44/42 MAPK(Erk1/2) Cell Signaling Cat# 4370 1:100
Rabbit monoclonal anti-Phospho-Akt (Ser473) Cell Signaling Cat# 4060 1:100 or 1:50
Rabbit polyclonal anti-Cleaved Caspase 3 Cell Signaling Cat# 9661 1:200
Rabbit polyclonal anti-Ras (G12D Mutant specific) Cell Signaling Cat# 14429 1:100 or 1:250
Rabbit polyclonal anti-beta-catenin Santa Cruz Cat# sc-7199 1:200
Mouse monoclonal anti-beta-catenin Sigma cat# 05-665 1:50
Goat polyclonal anti-RFP SICGEN Cat# AB8181-200 1:50
Rabbit polyclonal anti-SPC Merck Millipore Cat# AB3786 1:300
Rabbit polyclonal anti-LYZ DAKO Cat# A009902-2 1:200
Rabbit polyclonal anti-MUC2 Abcam Cat# ab90007 1:200
Rabbit monoclonal anti-NOTCH1 Abcam Cat# ab52627 1:100
Rat monoclonal anti-CD45 APC conjugated BD Pharmingen Cat# 559864 1:125
Rat monoclonal anti-CD326 (EpCAM) PE/Cy7 conjugated Biolegend Cat# 118216 1:125
Armenian Hamster monoclonal anti-CD81 PE conjugated Biolegend Cat# 104906 1:125
Rat monoclonal anti-CD140a (PDGFRA) BV421 conjugated Biolegend Cat# 135923 1:125
Goat anti rabbit IgG HRP conjugated Perkin Elmer Cat# NEF812001EA 1:500
Donkey Anti-Rabbit IgG Alexa Fluor 647 conjugated Invitrogen Cat# A32795 1:500
Donkey Anti-Mouse IgG Alexa Fluor 647 conjugated Invitrogen Cat# A32787 1:500
Donkey Anti-Goat IgG Alexa Fluor 647 conjugated Invitrogen Cat# A32849 1:500
Horse Anti-Rabbit IgG HRP Vectorlab Cat# MP-7401-50 1:500
Donkey Anti-Goat IgG HRP Genetex Cat# GTX232040-01 1:500

Concentration of each antibody is detailed in the manuscript as well.

Validation

Primary antibodies for immunostaining:

Goat polyclonal anti-CPA1 R and D Systems Cat# AF2765: The antibody guarantee covers the use of detecting mouse Carboxypeptidase A1/CPA1 in direct ELISAs and Western blots.
https://www.rndsystems.com/products/mouse-carboxypeptidase-a1-cpa1-antibody_af2765

Rabbit monoclonal anti-OLFM4 Cell Signaling Cat# 39141: The antibody guarantee covers the use of detecting mouse OLFM4 in Western blot, Immunoprecipitation, Immunohistochemistry and Immunofluorescence.
https://www.cellsignal.com/products/primary-antibodies/olfm4-d6y5a-xp-rabbit-mab-mouse-specific/39141

Rabbit monoclonal anti-Phospho-p44/42 MAPK(Erk1/2) Cell Signaling Cat# 4370: The antibody guarantee covers the use of detecting phospho-p44/42 MAPK(Erk1/2) from Human, Mouse, Rat, Hamster, Monkey, Mink, D. melanogaster, Zebrafish, Bovine, Dog, Pig and S. cerevisiae in Western blot, Immunoprecipitation, Immunohistochemistry, Immunofluorescence and Flow Cytometry.
https://www.cellsignal.com/products/primary-antibodies/phospho-p44-42-mapk-erk1-2-thr202-tyr204-d13-14-4e-xp-rabbit-mab/4370

Rabbit monoclonal anti-Phospho-Akt (Ser473) Cell Signaling Cat# 4060: The antibody guarantee covers the use of detecting pPhospho-Akt (Ser473) from Human, Mouse, Rat, Hamster, Monkey, Mink, D. melanogaster, Zebrafish and Bovine in Western blot, Immunoprecipitation, Immunohistochemistry, Immunofluorescence and Flow Cytometry.
https://www.cellsignal.com/products/primary-antibodies/phospho-akt-ser473-d9e-xp-rabbit-mab/4060

Rabbit polyclonal anti-Cleaved Caspase 3 Cell Signaling Cat# 9661: The antibody guarantee covers the use of detecting cleaved Caspase 3 from Human, Mouse, Rat and Monkey in Western blot, Immunoprecipitation, Immunohistochemistry, Immunofluorescence and Flow Cytometry.
https://www.cellsignal.com/products/primary-antibodies/cleaved-caspase-3-asp175-antibody/9661

Rabbit polyclonal anti-Ras (G12D Mutant specific) Cell Signaling Cat# 14429: The antibody guarantee covers the use of detecting G12D Mutant Ras from Human in Western blot.
https://www.cellsignal.com/products/primary-antibodies/ras-g12d-mutant-specific-d8h7-rabbit-mab/14429

Rabbit polyclonal anti-beta-catenin Santa Cruz Cat# sc-7199: The antibody guarantee covers the use of detecting beta-catenin from Mouse, Rat, Human, Xenopus and Zebrafish in Western blot, Immunoprecipitation, Immunohistochemistry and Immunofluorescence.
https://www.scbt.com/ko/p/beta-catenin-antibody-h-102

Mouse monoclonal anti-beta-catenin Sigma cat# 05-665: The antibody guarantee covers the use of detecting beta-catenin from Human, Mouse and Rat in Western blot, Immunohistochemistry, Immunofluorescence and Flow Cytometry.
https://www.sigmaaldrich.com/catalog/product/mm/05665?lang=en&region=GB&gclid=CjwKCAjw3pWDBhB3EiwAV1c5rFm_3Q5WYQVp_-lzKXdflPe3RAjV7GloHqf_UeeJN2Nfv383Z1O-bxoClCcQAvD_BwE

Goat polyclonal anti-RFP SICGEN Cat# AB8181-200: The antibody guarantee covers the use of detecting RFP in Western blot and Immunofluorescence.
https://www.origene.com/catalog/antibodies/primary-antibodies/ab8181-200/tdtomato-goat-polyclonal-antibody

Rabbit polyclonal anti-SPC Merck Millipore Cat# AB3786: The antibody guarantee covers the use of detecting SPC from Human, Mouse and Rat in Western blot, Immunoprecipitation, Immunohistochemistry, Immunofluorescence and ELISA.

https://www.merckmillipore.com/GB/en/product/Anti-Prosurfactant-Protein-C-proSP-C-Antibody,MM_NF-AB3786?ReferrerURL=https%3A%2F%2Fwww.google.com%2F

Rabbit polyclonal anti-LYZ DAKO Cat# A009902-2: The antibody guarantee covers the use of detecting Lysozyme from Human in Immunohistochemistry.
https://www.agilent.com/en/product/immunohistochemistry/antibodies-controls/primary-antibodies/lysozyme-ec-3-2-1-17-(concentrate)-76124

Rabbit polyclonal anti-MUC2 Abcam Cat# ab90007: The antibody guarantee covers the use of detecting pMUC2 from Human in Immunohistochemistry and Immunofluorescence.
https://www.abcam.com/muc2-antibody-ab90007.html

Rabbit monoclonal anti-NOTCH1 Abcam Cat# ab52627: The antibody guarantee covers the use of detecting NOTCH1 from Human and Mouse in Western blot,Immunohistochemistry, Immunofluorescence and Flow Cytometry.
https://www.abcam.com/notch1-antibody-ep1238y-ab52627.html

Primary antibodies for flow cytometry:

Rat monoclonal anti-CD45 APC conjugated BD Pharmingen Cat# 559864: The antibody guarantee covers the use of detecting CD45 from Mouse in Flow Cytometry.
https://www.bdbiosciences.com/eu/applications/research/stem-cell-research/cancer-research/mouse/apc-rat-anti-mouse-cd45-30-f11/p/559864

Rat monoclonal anti-CD326 (EpCAM) PE/Cy7 conjugated Biolegend Cat# 118216: The antibody guarantee covers the use of detecting CD326 (EpCAM) from Mouse in Flow Cytometry.
https://www.biolegend.com/en-us/products/pe-cyanine7-anti-mouse-cd326-ep-cam-antibody-5303

Armenian Hamster monoclonal anti-CD81 PE conjugated Biolegend Cat# 104906: The antibody guarantee covers the use of detecting CD81 from Mouse and Rat in Flow Cytometry.
https://www.biolegend.com/en-us/products/pe-anti-mouse-rat-cd81-antibody-238?GroupID=GROUP20

Rat monoclonal anti-CD140a (PDGFRA) BV421 conjugated Biolegend Cat# 135923: The antibody guarantee covers the use of detecting CD140a (PDGFRA) from Mouse and Rat in Flow Cytometry.
https://www.biolegend.com/en-gb/products/brilliant-violet-421-anti-mouse-cd140a-antibody-17921?GroupID=ImportedGROUP1

All antibodies have been validated by the manufacturer as above for the species and the application to be used. As positive control for immunostaining, tissues with known expression of the marker were used on mouse sections. As negative control, staining omitting the primary antibody was performed.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	All inducible Cre lines (R26R-CreERT2: JAX006965, Villin-CreERT2: JAX020282, Lgr5-EGFP-IRES-CreERT2: JAX008875, Sftpc-CreERT2: JAX028054, Krt5-CreERT2: JAX02915), the R26R-Confetti line (JAX017492), LSL-KrasG12D (JAX008179) line, Apcfl/fl line (JAX009045) and ApcMin line (JAX002020) were obtained from The Jackson Laboratory. Pik3caLat-H1047R line was kindly provided by Wayne.A Phillips. The generation of Red2Onco mice (R26R-Red2-Notch1icd, R26R-Red2-KrasG12D, R26R-Red2-PIK3CAH1047R) is described in this publication. Both male and female mice were used. Experiments were performed with 8-12 week-old mice unless otherwise stated in the figure legends. All mice were group housed under specific pathogen-free conditions in individually ventilated cages always with companion mice, and cages were placed under a 12hr light-dark cycle. Food and water were provided ad libitum. Room temperature was maintained at 22°C ± 1°C with 30–70% humidity. None of the mice were involved in any previous procedures before the study.
Wild animals	No wild animals were used in this study.
Field-collected samples	No field-collected samples were used in this study.
Ethics oversight	All mice were group housed under specific pathogen-free conditions. All procedures were performed according to UK Home Office regulations and local animal welfare committee guidelines. All experiments were approved by the UK Medical Research Council and University of Cambridge local ethical review committees and conducted according to Home Office project license PPL70/8296.

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	Small intestines were washed with cold PBS and villi were scraped off using a cover slip. The remaining tissue with crypts and lamina propria were then cut into 2 mm fragments and incubated in 10 mL Gentle Cell Dissociation Reagent (STEMCELL technologies) at 4 °C for 30 min. The fragments were then shaken vigorously and the supernatant was collected as the first fraction in a new conical tube. The remaining fragments were further incubated in fresh Gentle Cell Dissociation Reagent and a new fraction was collected every 30 min. The first and second fractions were discarded since they might contain some debris and villi. The third to fifth fractions contained the majority of the intestinal crypts. The crypt-enriched fractions were then washed with 10 mL of cold PBS and filtered through a 100 µm cell strainer (Falcon). The crypts were then further dissociated into single cells by incubation with TrypLE Express enzyme (GIBCO) at 37 °C for 5 min. The cells were filtered through a 40 µm cell strainer (Falcon). Remaining fragments with lamina propria were further digested with dissociation solution (2.5 mg/mL Liberase TL, Sigma; 10 U/mL DNAse, Promega) at 37 °C for 1 h. To aid dissociation, we gently mixed tissue pieces by pipetting up and down every 10 min. After 20 min, supernatants were harvested and 1 volume of DMEM containing 10% fetal bovine serum (FBS, Sigma Aldrich) was added while adding 5 mL of fresh dissociation mix to the remaining tissue pieces. This step was repeated 3 times for a total time of 1 h. After completion of the 3 cycles, the remaining intestinal fragments were mechanically disaggregated on a 100 µm mesh using a syringe plunger. The cell suspension was filtered through a 40 µm cell strainer into a 15 mL tube. The tube and filters were washed twice with 1 mL of 2% FBS in PBS and the cell suspension was then centrifuged at 300 g for 5 min. The supernatant was removed and the cell pellets were resuspended in 100 µL of antibody mix (2% FBS; 1:125 CD45 (30-F11)-APC, BD Biosciences; 1:125 EPCAM (G8.8)-PE-Cy7, BioLegend) and incubated for 1 h on ice. Both epithelial and lamina propria cells were then washed with 3 mL of 2% FBS in PBS and filtered once more if clumps were observed. After centrifugation at 300 g for 5 min, the cell pellets were resuspended in 1 mL of 2% FBS and 10 U/mL DNAse in PBS for sorting. For specific stromal cell sorting, the dissociated lamina propria cells were resuspended in 100 µL of antibody mix (2% FBS; 1:125 CD45 (30-F11)-APC, BD Biosciences; 1:125 EPCAM (G8.8)-PE-Cy7, BioLegend; 1:125 PDGFRA (APA5)-BV421, Biolegend; 1:125 CD81 (Eat-2)-PE, BioLegend) and incubated for 1h on ice. The cells were then washed with 3 mL of 2% FBS in PBS and filtered once more with a 40 µm cell strainer. After centrifugation at 300 g for 5 min, the cell pellets were resuspended in 1 mL of 2% FBS and 10 U/mL DNAse in PBS for sorting. The MOFLO system (Beckman Coulter) or SH800S Cell Sorter (Sony) of the Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute Flow Cytometry Facility was used for sorting and data were analyzed with FlowJo software (BD).
Instrument	MOFLO system (Beckman Coulter), SH800S Cell Sorter (Sony)
Software	Summit v5.2, SH800 software v2.1.5
Cell population abundance	Samples were sorted at 1000-1500 events/sec with a 2-way or 4-way purity mode, and using the 100um nozzle, achieving ~95% purity by FACS analysis. Efficiency was typically around 80-90%.
Gating strategy	FSC/SSC: gating for the live cell population; FSC/Pulse width: singlet gating. FSC/EPCAM: gating for mesenchymal and immune cells (EPCAM-)and gating for intestinal epithelial cells (EPCAM+). RFP: gating for mutant epithelial cells (RFP+). YFP: gating for WT epithelial cells (YFP+). CD45: gating for immune cells (CD45+) and gating for mesenchymal cells (CD45-). The gating strategy of the relevant cell populations is shown in Extended Data Fig.6a and Extended Data Fig.10f.

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