Shared requirement for MYC upstream super-enhancer region in tissue regeneration and cancer

Inderpreet Sur[1], Wenshuo Zhao[1], Jilin Zhang[1], Margareta Kling Pilström[1], Björn Rozell[9], Anna T. Webb[2], Huaitao Cheng[3], Ari Ristimäki[5,6], Pekka Katajisto[2,7,8], Martin Enge[3], Helena Rannikmae[4§], Marc de la Roche[4], Jussi Taipale*[1,4,5]

[1]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Sweden.

[2]Department of Cell and Molecular Biology, Karolinska Institutet, Sweden.

[3]Department of Oncology and Pathology, Karolinska Institutet, Sweden.

[4]Department of Biochemistry, University of Cambridge, United Kingdom.

[5]Applied Tumor Genomics Program, University of Helsinki, Biomedicum, P.O. Box 63 (Haartmaninkatu 8), FIN-00014 University of Helsinki, Finland.

[6]Department of Pathology, HUSLAB, HUS Diagnostic Center, Helsinki University Hospital and University of Helsinki, Finland.

[7]Institute of Biotechnology, HiLIFE, University of Helsinki, Finland.

[8]Molecular and Integrative Bioscience Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Finland.

[9]Department of Laboratory Medicine, Karolinska Institutet, Sweden.

*Corresponding Author: jussi.taipale@ki.se

[§]Present address: Complex In Vitro Models, In vitro/In Vivo Translation, GlaxoSmithKline, United Kingdom.

Abstract

Cancer has been characterized as a wound that does not heal. Malignant cells are morphologically distinct from normal proliferating cells, but have extensive similarities to tissues undergoing wound healing and/or regeneration. The mechanistic basis of this similarity has, however, remained enigmatic. Here we show that the genomic region upstream of Myc, which is required for intestinal tumorigenesis, is also required for intestinal regeneration after radiation damage. The region is also critical for growth of adult intestinal cells in 3D organoid culture, indicating that culture conditions that recapitulate most aspects of normal tissue architecture still reprogram normal cells to proliferate using a mechanism similar to that employed by cancer cells. Our results uncover a genetic link between tissue regeneration and tumorigenesis, and establish that normal tissue renewal and regeneration of tissues after severe damage are mechanistically distinct.

Summary: Myc regulatory region links regeneration and cancer

Main

Out of all genomic regions, the ~500 kb super-enhancer region on 8q24 that regulates MYC expression carries the largest burden of population-level cancer susceptibility[1,2]; it contains several common alleles that increase risk for multiple major forms of human cancer [1,3-6]. For example, cancer-associated G-allele of SNP rs6983267 is found in 50% and >85% of people with European and African ancestry, respectively[7,8]. This allele alone increases the risk for colorectal and prostate cancers more than 20%, causing hundreds of thousands of cancer deaths globally per year (https://www.wcrf.org/cancer-trends/global-cancer-data-by-country/). We have previously deleted regulatory sequences syntenic to the super-enhancer region in mice [9], and demonstrated that the mice are healthy and fertile in spite of greatly reduced levels of MYC expression in epithelia of the breast, prostate and intestine. The mice, however, are highly resistant to intestinal and mammary tumors. Although deletion of the region did not cause major effect on fitness in our assays, the enhancer elements within it are highly conserved between mammalian species [10-16]. This suggests that this region would also have a beneficial biological function that would balance the negative impact of the increased cancer predisposition.

In addition to being aberrantly expressed at high levels in majority of human cancers [17,18] , MYC is also induced by tissue damage[19-21]. MYC has similar functions in both tumorigenesis and tissue repair, driving several biochemical processes including increased ribosomal biogenesis and metabolic flux to support the proliferative state. An acute reduction in MYC level can completely block tumorigenesis[22] while, both a decrease and an increase in MYC level hinders the process of wound healing [23,24]. However, it is not known if the same regulatory elements control MYC levels during wound healing and tumorigenesis. Since MYC is one of the main drivers of cell proliferation, we hypothesized that the upstream super-enhancer region would be beneficial under some environmental condition that requires rapid cell proliferation, but which the mice fed well and housed in a relatively clean laboratory environment would not encounter. For example, the super-enhancer region could be important for responses to chronic infections, wounding, or other forms of tissue damage[25,26].

To test this, we subjected the $Myc^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice to ionizing radiation, which induces a uniform and highly reproducible damage to the intestinal lining [27,28]. Wild-type mice γ–irradiated for a total dose of 12 Grays rapidly developed intestinal damage and lost 7.0 ± 0.8% of their weight in 3 days (n=19); the weight loss stabilized at that time, and the damage was largely repaired by day 5. By contrast, the $Myc^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice appeared unable to repair the damage, as they continued

to lose weight beyond 3 days, losing on average 17% by 5 days (n=10) (**Fig. 1a-b**). Morphological analysis of the small intestines confirmed that the damage was repaired in the wild-type mice but, sustained in the *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice due to failure to regenerate the intestinal crypts (**Fig. 1c, Extended data Fig. 1**). The wild-type mice displayed robust *Myc* induction at 3 days, and a strong proliferative response at 5 days, whereas at the same time points, in the *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice *Myc* expression was undetectable and there was almost complete absence of proliferating cells (**Fig. 2a-b**, **Extended data Fig. 2**). Peak induction of *Myc* transcripts occurred 3 days post irradiation (dpi, **Fig. 2c, Extended data Table 1**). The loss of proliferation was specific to the repair/regenerative stimulus since the cellular architecture and expression profile of cells from the small intestine of the *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice and wildtype were similar prior to irradiation (**Extended data Fig. 3**). To decipher changes in cellular transcriptome when MYC levels are increasing after tissue damage, we performed single-cell RNA-seq (scRNA-seq) analysis at day 2 after irradiation. Analysis of scRNA-seq data revealed that the *Myc* expression was limited to stem and transit amplifying cells of wild-type mice. Despite the irradiation, which robustly induced MYC in wild-type mice, stem cells or transit amplifying cells from *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice expressed almost no *Myc*, the signal being even lower than that from irradiated mature enterocytes of wild-type mice (**Fig. 2d**). The *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice were, however, fully capable of mounting a normal DNA damage response as seen by the induction of p53 target genes *Cdkn1a* (p21) and *Bbc3* (PUMA), which facilitate p53 dependent cell cycle arrest and apoptosis, respectively (**Fig. 2e**).

To determine if the loss of regenerative potential was due to MYC, we further analyzed the expression of all genes within 5 Mb of *Myc*. By far, the strongest effect seen was for *Myc* itself, whose expression was almost completely abolished in the mutant intestine. Transcripts of some genes telomeric of *Myc* are also downregulated, but to a much lesser extent. These include the non-coding RNA *Pvt1*, and the gasdermin C genes *Gsdmc2*, *3* and *4* (**Extended data Table 2**). However, only the expression of *Myc* and *Pvt1* was specifically upregulated in wildtype cells upon irradiation. Because *Pvt1* lncRNA is thought to respond to the same enhancer elements and elicit its effect via interaction with *Myc* [29,30], the results indicate that MYC is the primary effector of the *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ phenotype. Consistent with this, unbiased gene set enrichment analysis revealed that a set of MYC targets (V1) failed to be upregulated by irradiation in stem cells of the *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice (**Fig. 3a**, **Extended data Table 3**). We also specifically analyzed expression of a set of 126 functionally conserved MYC targets from Zielke et al. [31], most of which are involved in metabolism and ribosome biogenesis. The regulation of this set of genes is conserved between

Drosophila and humans. The functionally conserved MYC targets were also strongly upregulated in the stem-cells of irradiated wild-type mice compared to *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice (**Fig. 3b**). These results establish that induction of MYC drives a robust program of increased ribosome biogenesis and anabolic activity that facilitates growth of new tissue to repair the damage caused by the irradiation. However, among the MYC regulated genes, we also observed induction of the negative translational regulator eukaryotic initiation factor 4E binding protein 1 (*Eif4ebp1*) [32] (**Fig. 3d**), suggesting that the induction of MYC is necessary but not sufficient to induce the regenerative response. 4EBP1 is inactivated by phosphorylation by the mTOR complex (mTORC) kinase [33], and it has been previously shown that mTORC signalling is required for regeneration of intestine after radiation damage [34]. Staining of the intestines before radiation damage revealed strong signal for phosphorylated 4EBP1 (Thr37/46) while after radiation staining was more prominent in the regenerating crypts of the wildtype (**Fig. 3e**), consistent with a role for both MYC and mTORC in the regenerative response.

Analysis of cell types and proliferation prior to damage revealed no major differences between wild-type and *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice (**Extended data Fig. 3**). Furthermore, in the normal non-irradiated *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice, the functionally conserved MYC targets were only modestly downregulated within the stem and transit amplifying cells compared to wildtype mice (**Fig. 3b**, **Extended data Table 4**). By contrast, there was no downregulation of the MYC target genes in mature enterocytes, which do not express MYC. We observed an increase in *L-Myc* expression within the stem and transit amplifying cell clusters of *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice suggesting compensation of MYC function during homeostasis and explaining why the loss of the Myc 2-540 kb region is well tolerated in mice under normal conditions (**Fig. 3c, Extended data Fig. 4**). It is noteworthy that the functionally conserved MYC target gene set that we find enriched after tissue damage was originally detected from tumor tissues [31], suggesting that increased ribosomal biogenesis and metabolic flux drive both cancer and tissue repair. These results indicate that the failure of the *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ is most likely due to failure to respond to damage, as opposed to differences in tissue structure, morphology, or gene activity prior to the damage.

To further study the mechanism underpinning the failure of *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ intestines to regenerate, we determined whether the loss of MYC affected the mobilization of reserve stem/progenitor cells which previously have been shown to be associated with intestinal regeneration [35,36]. For this purpose, we checked expression of *Ly6a* which codes for the stem cell marker Sca1. Expression of Sca1 is induced by colitis and radiation induced damage as well as

damage caused by helminth infections [35,36]. Analysis of our scRNA-seq data showed that irradiation of the wildtype intestine mobilized a robust population of Ly6a/Sca1 expressing cells. These cells were distributed amongst the stem, transit amplifying and enterocyte progenitor cell clusters, confirming their stem cell potential (**Fig. 3f-g**). By contrast, the *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice almost completely failed to mobilize the Ly6a/Sca1^{+} cells after irradiation (**Fig. 3f-g**). These results establish that the 2-540 kb conserved region which is required for tumorigenesis [37] is also required for MYC-dependent repair of intestinal lining after acute large-scale damage.

To further characterize the role of MYC in tissue repair, we decided to test a recently described *ex vivo* tissue regeneration assay, where mouse ileum is decellularized with sodium deoxycholate resulting in a matrix reminiscent of the 3-D intestinal structure *in vivo* with distinct regions where former crypts and villi were located. Overlaying of the decellularized extracellular matrix (dECM) with isolated crypts results in repopulation of the empty crypt pits and regeneration of the entire intestinal epithelium [38]. In this assay, cultures of wild-type crypts repopulated the dECM's and gave rise to crypt and villus structures, whereas crypts from *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice failed to do so (**Fig. 4a**). As repopulation of the bare dECM requires cellular motility, with proliferation and self-renewal restricted to the empty crypt pits, we next placed isolated crypts in standard organoid culture conditions in matrigel. Surprisingly, adult intestinal organoids from *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice failed to grow also in such culture, where self-renewal and proliferation promoting signals are provided evenly (**Fig. 4b**).

Taken together, these results establish that although the *Myc* upstream super-enhancer region is not required for normal tissue renewal in mouse intestine, it is required for tissue repair after radiation damage and for growth of adult intestinal cells under organoid culture conditions.

Discussion

We show here that the same genetic element, the *Myc* upstream super-enhancer region, is required for both tissue renewal after radiation damage, and tumorigenesis. The super-enhancer region is critical for the induction of *Myc* after intestinal damage caused by irradiation. Mice deficient of the region fail to induce expression of MYC target genes important for tissue repair, and are unable to mobilize Ly6a/Sca1^{+} regenerative stem/progenitor cells. These results establish the long-sought mechanistic link between wound healing and cancer, and show that proliferation during normal homeostasis and tissue regeneration use regulatory mechanisms that are genetically

distinct. Given that upregulation of ribosome biogenesis is required for MYC-driven proliferation [31,39,40], the simplest model consistent with our findings is that different levels of protein synthesis are needed for distinct proliferative cell fate choices. In this model, relatively low levels of protein synthesis and MYC are sufficient for proliferation that maintains stem cell number in tissue renewal, whereas processes that require exponential increase in the number of somatic stem cells, such as cell proliferation *in vitro,* regeneration, and tumorigenesis (**Fig. 4c**) require high levels of MYC and protein synthesis. Interestingly, the model has similarities to growth control in unicellular species, where nutrient availability specifies the protein synthetic rate and controls the balance between non-proliferative and proliferative states [41]. Multicellularity could have evolved simply by the addition of a hierarchy of stem cells with new proliferative fates, which remain controlled by the same input variable, the total metabolic activity per genome. Consistently with this model, ES cells lacking MYC enter into a diapause-like non-proliferative state *in vitro* [42], and we also observe here that adult intestinal cells from $Myc^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice enter a non-proliferative state under full mitogen stimulation (organoid culture).

Our results also show that cell culture models that are often considered to resemble normal *in vivo* tissue renewal actually drive cells to adopt an exponential cell proliferation program associated with regeneration and tumorigenesis. This has important consequences for drug development. In particular, compounds that inhibit proliferation of normal cells in culture have commonly been discarded during the development process due to toxicity, even when a subset of them may actually target the very mechanism that is distinct between normal proliferative cells and cancer cells.

Methods

Animals

Mouse lines used in this study have been described previously [9] . The conditional knockout mice were generated on a C57BL/6N genetic background. After Cre-mediated deletion of 538 kb region, mice were backcrossed to C57Bl6/J for two generations and subsequently intercrossed. All mice were kept on a 12:12 h light cycle in standard housing conditions and provided with food (Teklad global diet 2918; Envigo) and water *ad libitum*. Animal experiments were done in accordance with the regulations and with approval of the local ethical committee (Stockholms djurförsöksetiska nämnd). Mice were irradiated with Gammacell® 40 Exactor that uses Caesium-137 as a source for γ-rays. Mice were given 2 doses of 6 Gy with a 6 hr interval in between the doses. After irradiation mice were provided with soft food. For measuring survival, the general health of the animal was monitored and mice were sacrificed when they reached the humane end-point according to the ethical permit.

Immunohistochemistry (IHC) and RNA *In Situ* hybridization (ISH)

PFA-fixed and paraffin embedded (FFPE) tissues were sectioned at 5 µm thickness and subjected to IHC using standard methods. For antigen retrieval, the sections were boiled in microwave for 20 minutes in 10 mM Citrate buffer pH 6.0 and subsequently processed for respective antibody stainings. Following antibodies and working dilutions were used: anti-Ki67 (Abcam, 1:200), anti-Lysozyme (Abcam, 1:500), anti-Muc2 (Abcam, 1:500), anti-Chga (Abcam, 1:1000), anti-phospho 4EBP1 (Cell signalling, 1:2500). ISH was also performed on 5 µm thick sections using the RNAscope technology and according to manufacturer's protocol. Probes against mouse *Myc* (catalog nr: 413451), *L-Myc* (catalog nr: 552711) and *Lgr5* (catalog nr: 312171) were used. RNA integrity was confirmed using the control probe for *Ubc* (catalog nr:310771).

qPCR

Approximately 1 cm of the mouse distal ileum was cut into small pieces and stored in RNAlater (Qiagen). RNA was prepared using a combination of Trizol (Ambion) and RNeasy mini kit (Qiagen). cDNA was synthesized from either 1 µg (tissue) or 100 ng (organoid) of total RNA using

High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). qPCR was performed using SYBR™ Green master mix (Thermo Fisher) on a LightCycler® 480 instrument (Roche). Primers for *Myc* and *Actin* were as described previously [9].

Single-cell RNA-seq (scRNA-seq)

Mouse intestinal crypts were isolated according to Sato and Clevers [43] , dissociated with TrypLE express (37°C/ 20 min, Invitrogen) and single cell suspensions were run on a 10x Chromium controller using the Chromium Next GEM Single Cell 3´kit v3.1 (catalog nr: PN-1000268) and according to manufacturer's protocol (10x Genomics). Cell number targeted was 6000. For non-irradiated samples, crypts were isolated from ileum and single cell suspension of FACS sorted $7AAD^{-}$ (Life Technologies), $CD45^{-}$ (eBioscience), $CD31^{-}$ (eBioscience), TER-119^{-} (eBioscience), $EpCAM^{+}$ (eBioscience) cells were used. For irradiated samples, crypts from the whole small intestine were isolated and $7AAD^{-}$ FACS sorted cells were used. Libraries were sequenced on illumina sequencing platform NovaSeq 6000 or NextSeq 2000. Sequencing data was processed using Cell Ranger (v6.0.1) (10x Genomics). Quality control (QC) analysis of the Cell Ranger output was done in Scanpy (v1.9.1) [44] as follows: Mean and Standard deviation (SD) of genes in all cells [Mean (n_{genes}), SD(n_{genes})] and Mean, Median, Median absolute deviation (MAD) and SD of total counts in all cells was calculated. Cells were kept with the number of expressed genes larger than [Median(n_{genes})-MAD(n_{genes})], total read counts in the range of [Mean(n_{counts})-1.5*MAD(n_{counts}), Mean(n_{counts})+3*SD(n_{counts}) and percentage of expressed mitochondrial genes larger than its median+3*MAD. Additionally, genes expressed in less than 0.1% of cells were filtered out. Subsequently for dimensional reduction, top 2000 highly variable genes were selected based on the single-cell RNA raw read count matrix (scanpy.pp.highly_variable_genes; flavor= 'seurat_v3'). There were no mitochondrial genes in the top 2000 highly variable genes. The raw read count matrix with top 2000 highly variable genes (22,108 x 2,000) was subjected to the Bayesian variational inference model scVI (v0.16.4) [45] to infer a latent space for all single cells. The batch effect among different samples was automatically removed for all samples within the scVI. After that, two-dimensional visualizations were calculated by using Uniform Manifold Approximation and Projection (UMAP). Clustering analysis was performed using Leiden algorithm. First a k-nearest-neighbor graph from the scVI latent space was constructed and then sc.tl.leiden function in scanpy was used with the resolution of 1. This resulted in 19 clusters. The small cluster expressing immunocyte markers was not used for further analysis (excluded 34 cells

and 22,074 cells were kept). Remaining clusters were annotated into 7 intestinal cell types based on previously published marker genes [46].

For further analysis, the single cell raw read count matrix of all the genes was input into R package Scran (v1.24.1) to calculate the size factor for each cell [47]. To normalize the gene expression across cells, each gene's read count in a cell was divided by that cell's size factor in Scanpy. The normalized gene read counts were $\log_2$ transformed, where 1 pseudo count was added to gene's read count to avoid negative infinity value in the analysis (log1p value). The z-score for the log1p value for each gene was calculated and used for heatmap visualization of cell-type specific marker genes.

Differentially expressed genes between the genotypes were determined using scanpy's function sc.tl.rank_genes_groups. The analysis was done specifically in the stem cell (SC), transit amplifying cell (TA) and enterocyte clusters (EC). Within this function, Wilcoxon test was used to calculate p values and the false discovery rate (FDR) by applying the benjamini-hochberg approach. The log fold changes of the 126 known MYC targets [31] were retrieved from the differential gene expression results and plotted using a dot plot. Polr1f was not detected in our dataset and was excluded from analysis (**Extended data Table 4**). The ranking of the genes on the X-axis was based on the median log fold change values of each MYC target gene among stem cell (SC), transit amplifying cell (TA) and enterocyte (EC) clusters combined.

Gene Set Enrichment Analysis was done using R package clusterProfiler (v4.4.4). Gene signatures were obtained from Gene Ontology Biological Process gene sets [48] and the MSigDB 'Hallmark gene sets' [49]. All the expressed genes in the SC cluster were pre-ranked by their log fold changes which were determined as described in the previous section. GSEA analysis was then performed using ClusterProfiler (ClusterProfiler::GSEA; pvalueCutoff=1, pAdjustMethod=´BH´). The full list of GSEA results is provided in Extended data Table 2.

Single cell ATAC-seq (scATAC-seq)

scATAC-seq was performed according to Cheng *et al 2021* [50]. Single cell suspensions of mouse intestinal crypts were prepared as in [51] and single cells were FACS sorted into 384 well plate with 3 µl lysis buffer in each well (0.03 µl 1M Tris-pH7.4, 0.0078 µl 5M NaCl, 0.075 µl 10% IGEPAL, 0.075 µl RNases Inhibitor, 0.075 µl 1:1.2M ERCC, 2.7372 µl H_2O). After cell lysis, 1µl of the lysis buffer containing the nuclei was used for scATAC library preparation, while the remaining 2 µl

were discarded. The scATAC *in situ* tagmentation was performed with 2µl of the Tn5 tagmentation mix (0.06 µl 1M Tris-pH 8.0, 0.0405 µl 1M $MgCl_2$, 0.2 µl Tn5) at 37°C for 30 mins. After tagmentation, 2 µl of the supernatant was aspirated and the nucleus was then washed once with 10 µl ice-cold washing buffer (0.1 µl 1M Tris-pH7.4, 0.02 µl 5M NaCl, 0.03µl 1M $MgCl_2$, 9.85 µl H_2O). The remaining Tn5 was inactivated by adding 2 µl 0.2% SDS-Triton X-100 and incubating at room temperature for 15 mins followed by 55°C for 10 mins. Before barcoding PCR, genomic DNA was extracted from chromatin by adding 2µl Proteinase K (0.0107 au/mL) and incubation at 55°C for 1h followed by heat inactivation at 70°C for 15 mins. Barcoding PCR was done using KAPA HiFi PCR Kit (Roche) in a final volume of 25 µl. The PCR condition was 72°C/15 min, 95°C/45 s, [98°C/15 s, 67°C/30 s, 72°C/1 min] x 22 cycles, 72°C/5 min, and then 4°C hold. After the PCR, 2µl reaction from each well was pooled and cleaned-up twice using SPRI- beads (at 1.3X volume) and the library was sequenced on Illumina NextSeq 550 as paired-end, dual index (37+37+8+8). For scATAC-seq data, raw sequence reads were trimmed based on quality (phred-scaled value of >20) and the presence of illumina adapters, and then aligned to the mm10 genome build using BWA [52]. Reads that were not mapped, not with primary alignment, missing a mate, mapq <10, or overlapping ENCODEs blacklist regions [53] were removed. A custom script was used to summarize the paired-end reads into a de-duplicated fragment file suitable for downstream analysis. scATAC-seq data was analyzed using ArchR. Cells with fewer than 1000 ATAC-seq reads mapping to transcription start site, or with TSS enrichment < 4, were excluded from analysis. The "iterative LSI" function with maxClusters=6 was used for cluster analysis. Clusters with high expression of 'Muc2' were annotated as goblet cells and discarded. Clusters high in 'Ascl2' and 'Lgr5' were annotated as stem cells, while remaining intermediate/low expressing clusters were annotated as Transit amplifying cells. Each of these superclusters were split by genotype and merged to in-silico bulk experiments.

Intestinal crypt isolation and Organoid culture

Crypts were isolated from adult mouse small intestine (approx. 3 month old) according to Sato and Clevers 2013 [43]. The crypts were seeded in Matrigel (Corning) and cultured in ENR medium (Advanced DMEM/F12 (Thermo Fisher Scientific), 1x penicillin/streptomycin, (Thermo Fisher Scientific), 1x Glutamax (Thermo Fisher Scientific), 10mM HEPES (Thermo Fisher Scientific), 1x B27 and 1x N2 (Thermo Fisher Scientific), 1.25mM N-acetylcysteine (Sigma), 50 ng/ml of

murine recombinant Epidermal growth factor (R&D), 100 ng/ml recombinant murine Noggin (R&D), 1 mg/ml recombinant murine R-Spondin (R&D).

Re-epithelialization of intestinal scaffolds with crypts

The ileum from the small intestine of mice, was decellularized as previously described by Iqbal *et al* 2021 [38]. Pieces of decellularized intestines (dECMs) were primed at 37˚C with 50 µl of media containing Advanced DMEM/F12, 1x penicillin/streptomycin, 1x Glutamax (Thermo Fisher Scientific), 10mM HEPES (Thermo Fisher Scientific), 1x B27 (Life Technologies), 1x N2 (Life Technologies), 1mM N-acetylcysteine (Sigma), 50 ng/ml of murine recombinant Epidermal growth factor (R&D), 100 ng/ml recombinant murine Noggin (Peprotech), 500 ng/ml recombinant murine R-Spondin (R&D), 10µM Y-27632 (Sigma), for 15-60 mins. This was followed by aspirating media carefully and plating around 100 crypts on each scaffold piece and incubating at 37˚C for 10 min. Next, 10 µl of media was added to each scaffold piece and incubated at 37˚C for 10 min, followed by an additional 20 µl per scaffold piece and incubation at 37˚C for 20 min. Finally, 260 µl of media was added to each scaffold piece and put in the incubator at 37˚C. The day after seeding, media was changed, and each scaffold piece was lifted to make it float. The second day after seeding, the media was changed to media without Y-27632. Media was subsequently replaced every other day. Re-epithelialized dECMs were fixed with 4% PFA at day 16 after plating. The re-epithelialized dECMs were analyzed using immunofluorescence. Primary antibody Pan-laminin (1:300, Abcam, 11575) and secondary antibody Alexa-647 conjugated anti-rabbit (1:1000, Life) were used. For nuclei detection, DAPI (Life, 1:1000) was used. Images were obtained with a Zeiss LSM980- Airy2 microscope using a 20x objective, NA=08.

Statistics and reproducibility

Sample size was determined based on our previous experience and pilot experiments. Information on statistical significance is provided in the figure legend.

Data availability: Sequencing data will be uploaded to European Nucleotide Archive (ENA, EMBL-EBI) under project accession number PRJEB57218.

References

1 Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* **47**, D1005-D1012, doi:10.1093/nar/gky1120 (2019).

2 Sur, I., Tuupanen, S., Whitington, T., Aaltonen, L. A. & Taipale, J. Lessons from functional analysis of genome-wide association studies. *Cancer Res* **73**, 4180-4184, doi:10.1158/0008-5472.CAN-13-0789 (2013).

3 Gudmundsson, J. *et al.* Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet* **39**, 631-637, doi:10.1038/ng1999 (2007).

4 Gudmundsson, J. *et al.* Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. *Nat Genet* **41**, 1122-1126, doi:10.1038/ng.448 (2009).

5 Gudmundsson, J. *et al.* A study based on whole-genome sequencing yields a rare variant at 8q24 associated with prostate cancer. *Nat Genet* **44**, 1326-1329, doi:10.1038/ng.2437 (2012).

6 Yeager, M. *et al.* Identification of a new prostate cancer susceptibility locus on chromosome 8q24. *Nat Genet* **41**, 1055-1057, doi:10.1038/ng.444 (2009).

7 Haiman, C. A. *et al.* A common genetic risk factor for colorectal and prostate cancer. *Nat Genet* **39**, 954-956, doi:10.1038/ng2098 (2007).

8 Yeager, M. *et al.* Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* **39**, 645-649, doi:10.1038/ng2022 (2007).

9 Dave, K. *et al.* Mice deficient of Myc super-enhancer region reveal differential control mechanism between normal and pathological growth. *Elife* **6**, doi:10.7554/eLife.23382 (2017).

10 Ahmadiyeh, N. *et al.* 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. *Proc Natl Acad Sci U S A* **107**, 9742-9746, doi:10.1073/pnas.0910668107 (2010).

11 Hallikas, O. *et al.* Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* **124**, 47-59, doi:10.1016/j.cell.2005.10.042 (2006).

12 Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-947, doi:10.1016/j.cell.2013.09.053 (2013).

13 Pomerantz, M. M. *et al.* The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat Genet* **41**, 882-884, doi:10.1038/ng.403 (2009).

14 Sotelo, J. *et al.* Long-range enhancers on 8q24 regulate c-Myc. *Proc Natl Acad Sci U S A* **107**, 3001-3005, doi:10.1073/pnas.0906067107 (2010).

15 Tuupanen, S. *et al.* The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. *Nat Genet* **41**, 885-890, doi:10.1038/ng.406 (2009).

16 Yan, J. *et al.* Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell* **154**, 801-813, doi:10.1016/j.cell.2013.07.034 (2013).

17 Schaub, F. X. *et al.* Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas. *Cell Syst* **6**, 282-300 e282, doi:10.1016/j.cels.2018.03.003 (2018).

18 Vita, M. & Henriksson, M. The Myc oncoprotein as a therapeutic target for human cancer. *Semin Cancer Biol* **16**, 318-330, doi:10.1016/j.semcancer.2006.07.015 (2006).
19 Shin, H. Y. *et al.* Role of Myc Proto-Oncogene as a Transcriptional Hub to Regulate the Expression of Regeneration-Associated Genes following Preconditioning Peripheral Nerve Injury. *J Neurosci* **41**, 446-460, doi:10.1523/JNEUROSCI.1745-20.2020 (2021).
20 Morello, D., Fitzgerald, M. J., Babinet, C. & Fausto, N. c-myc, c-fos, and c-jun regulation in the regenerating livers of normal and H-2K/c-myc transgenic mice. *Mol Cell Biol* **10**, 3185-3193, doi:10.1128/mcb.10.6.3185-3193.1990 (1990).
21 Tsuboi, K. *et al.* Soluble factors including proteinases released from damaged cells may trigger the wound healing process. *Biochem Biophys Res Commun* **168**, 1163-1170, doi:10.1016/0006-291x(90)91151-h (1990).
22 Sodir, N. M. *et al.* Reversible Myc hypomorphism identifies a key Myc-dependency in early cancer evolution. *Nat Commun* **13**, 6782, doi:10.1038/s41467-022-34079-x (2022).
23 Stojadinovic, O. *et al.* Molecular pathogenesis of chronic wounds: the role of beta-catenin and c-myc in the inhibition of epithelialization and wound healing. *Am J Pathol* **167**, 59-69, doi:10.1016/s0002-9440(10)62953-7 (2005).
24 Zanet, J. *et al.* Endogenous Myc controls mammalian epidermal cell size, hyperproliferation, endoreplication and stem cell amplification. *J Cell Sci* **118**, 1693-1704, doi:10.1242/jcs.02298 (2005).
25 Plummer, M. *et al.* Global burden of cancers attributable to infections in 2012: a synthetic analysis. *Lancet Glob Health* **4**, e609-616, doi:10.1016/S2214-109X(16)30143-7 (2016).
26 Schafer, M. & Werner, S. Cancer as an overhealing wound: an old hypothesis revisited. *Nat Rev Mol Cell Biol* **9**, 628-638, doi:10.1038/nrm2455 (2008).
27 Gu, J. *et al.* At What Dose Can Total Body and Whole Abdominal Irradiation Cause Lethal Intestinal Injury Among C57BL/6J Mice? *Dose Response* **18**, 1559325820956783, doi:10.1177/1559325820956783 (2020).
28 Kim, C. K., Yang, V. W. & Bialkowska, A. B. The Role of Intestinal Stem Cells in Epithelial Regeneration Following Radiation-Induced Gut Injury. *Curr Stem Cell Rep* **3**, 320-332, doi:10.1007/s40778-017-0103-7 (2017).
29 Cui, M. *et al.* Long non-coding RNA PVT1 and cancer. *Biochem Biophys Res Commun* **471**, 10-14, doi:10.1016/j.bbrc.2015.12.101 (2016).
30 Tseng, Y. Y. *et al.* PVT1 dependence in cancer with MYC copy-number increase. *Nature* **512**, 82-86, doi:10.1038/nature13311 (2014).
31 Zielke, N., Vaharautio, A., Liu, J., Kivioja, T. & Taipale, J. Upregulation of ribosome biogenesis via canonical E-boxes is required for Myc-driven proliferation. *Dev Cell* **57**, 1024-1036 e1025, doi:10.1016/j.devcel.2022.03.018 (2022).
32 Tameire, F. *et al.* ATF4 couples MYC-dependent translational activity to bioenergetic demands during tumour progression. *Nat Cell Biol* **21**, 889-899, doi:10.1038/s41556-019-0347-9 (2019).
33 Saxton, R. A. & Sabatini, D. M. mTOR Signaling in Growth, Metabolism, and Disease. *Cell* **168**, 960-976, doi:10.1016/j.cell.2017.02.004 (2017).
34 Sampson, L. L., Davis, A. K., Grogg, M. W. & Zheng, Y. mTOR disruption causes intestinal epithelial cell defects and intestinal atrophy postinjury in mice. *FASEB J* **30**, 1263-1275, doi:10.1096/fj.15-278606 (2016).
35 Nusse, Y. M. *et al.* Parasitic helminths induce fetal-like reversion in the intestinal stem cell niche. *Nature* **559**, 109-113, doi:10.1038/s41586-018-0257-1 (2018).

36 Yui, S. *et al.* YAP/TAZ-Dependent Reprogramming of Colonic Epithelium Links ECM Remodeling to Tissue Regeneration. *Cell Stem Cell* **22**, 35-49 e37, doi:10.1016/j.stem.2017.11.001 (2018).
37 Sur, I. K. *et al.* Mice lacking a Myc enhancer that includes human SNP rs6983267 are resistant to intestinal tumors. *Science* **338**, 1360-1363, doi:10.1126/science.1228606 (2012).
38 Iqbal, S. *et al.* Fetal-like reversion in the regenerating intestine is regulated by mesenchymal Asporin. *bioRxiv*, 2021.2006.2024.449590, doi:10.1101/2021.06.24.449590 (2021).
39 van Riggelen, J., Yetil, A. & Felsher, D. W. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer* **10**, 301-309, doi:10.1038/nrc2819 (2010).
40 Pihlajamaa, P., Kauko, O., Sahu, B., Kivioja, T. & Taipale, J. A competitive precision CRISPR method to identify the fitness effects of transcription factor binding sites. *Nat Biotechnol*, doi:10.1038/s41587-022-01444-6 (2022).
41 Jorgensen, P. & Tyers, M. How cells coordinate growth and division. *Curr Biol* **14**, R1014-1027, doi:10.1016/j.cub.2004.11.027 (2004).
42 Scognamiglio, R. *et al.* Myc Depletion Induces a Pluripotent Dormant State Mimicking Diapause. *Cell* **164**, 668-680, doi:10.1016/j.cell.2015.12.033 (2016).
43 Sato, T. & Clevers, H. Primary mouse small intestinal epithelial cell cultures. *Methods Mol Biol* **945**, 319-328, doi:10.1007/978-1-62703-125-7_19 (2013).
44 Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**, 15, doi:10.1186/s13059-017-1382-0 (2018).
45 Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat Methods* **15**, 1053-1058, doi:10.1038/s41592-018-0229-2 (2018).
46 Yum, M. K. *et al.* Tracing oncogene-driven remodelling of the intestinal stem cell niche. *Nature* **594**, 442-447, doi:10.1038/s41586-021-03605-0 (2021).
47 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, doi:10.1186/s13059-016-0947-7 (2016).
48 Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29, doi:10.1038/75556 (2000).
49 Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425, doi:10.1016/j.cels.2015.12.004 (2015).
50 Cheng, H. *et al.* Smart3-ATAC: a highly sensitive method for joint accessibility and full-length transcriptome analysis in single cells. *bioRxiv*, 2021.2012.2002.470912, doi:10.1101/2021.12.02.470912 (2021).
51 Andrews, N. *et al.* An unsupervised method for physical cell interaction profiling of complex tissues. *Nat Methods* **18**, 912-920, doi:10.1038/s41592-021-01196-2 (2021).
52 Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595, doi:10.1093/bioinformatics/btp698 (2010).
53 Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci Rep* **9**, 9354, doi:10.1038/s41598-019-45839-z (2019).

Acknowledgements

The authors wish to thank Drs Rong Yu, Yujiao Wu and Ekaterina Morgunova for their comments on the manuscript. We also wish to thank the core facilities at KI: The unit for morphological phenotype analysis (FENO), Bioinformatics and expression analysis core facility (BEA), Biomedicum imaging core (BIC) and Biomedicum flow cytometry (BFC). The research was funded by grants from the Swedish Research Council (D0815201) and Cancerfonden (20 1116 PjF 02H).

Author contribution

IS and JT conceived and designed the experiment, analysed data and wrote the original draft. IKS, MKP, WZ, ATW, HC and HR performed experiments, BR and AR analysed histological data,WZ and JZ performed scRNA-seq data analysis. ME analysed scRNA-seq/ scATAC-seq data, PK, IKS, MR, JT provided supervision. WZ, JZ, MKP, BR, AW, HC, ME, AR, PK, HR, MR reviewed and edited the manuscript.

Ethics declarations

Competing interests: The authors declare no competing interests.

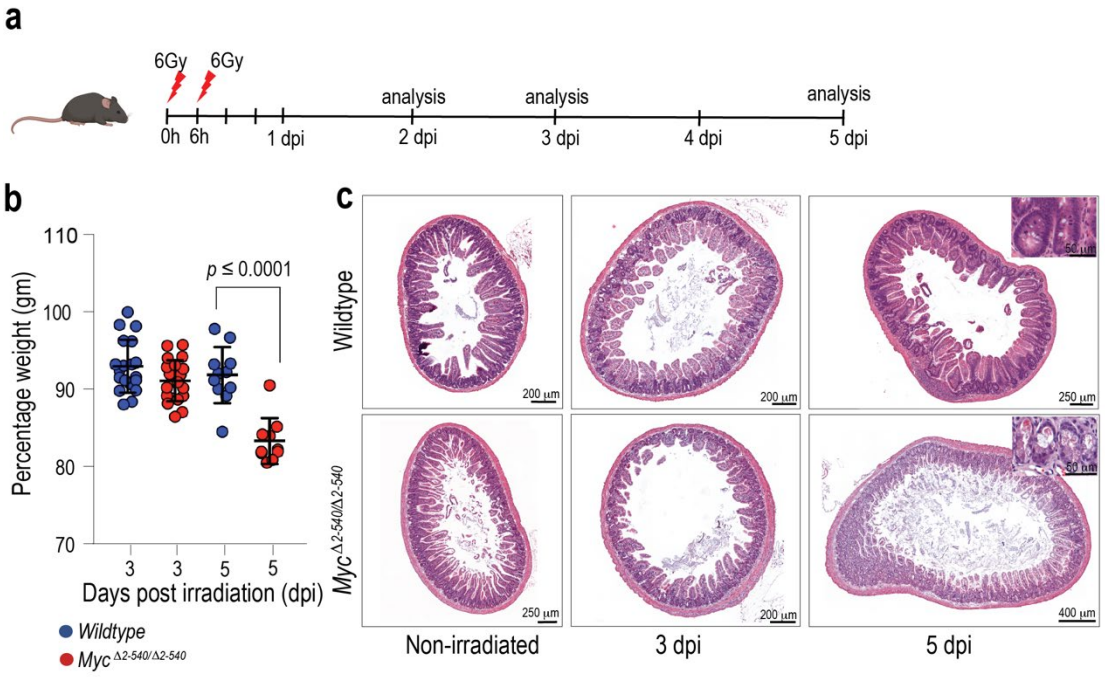

Fig. 1. *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ intestine fails to support regenerative growth *in vivo* post γ-irradiation **a**, Timeline of the irradiation experiment. **b,** *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice (red) are more sensitive to weight loss post irradiation compared to wildtype mice (blue). Wildtype: n = 19 (3dpi), n = 11 (5dpi), *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$: n = 21 (3dpi), n = 10 (5dpi). Statistical significance was calculated using unpaired student's *t-test* (two-tailed). **c,** Hematoxylin and Eosin (H&E) stained sections of small intestine showing *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice fail to regenerate crypts after radiation induced damage.

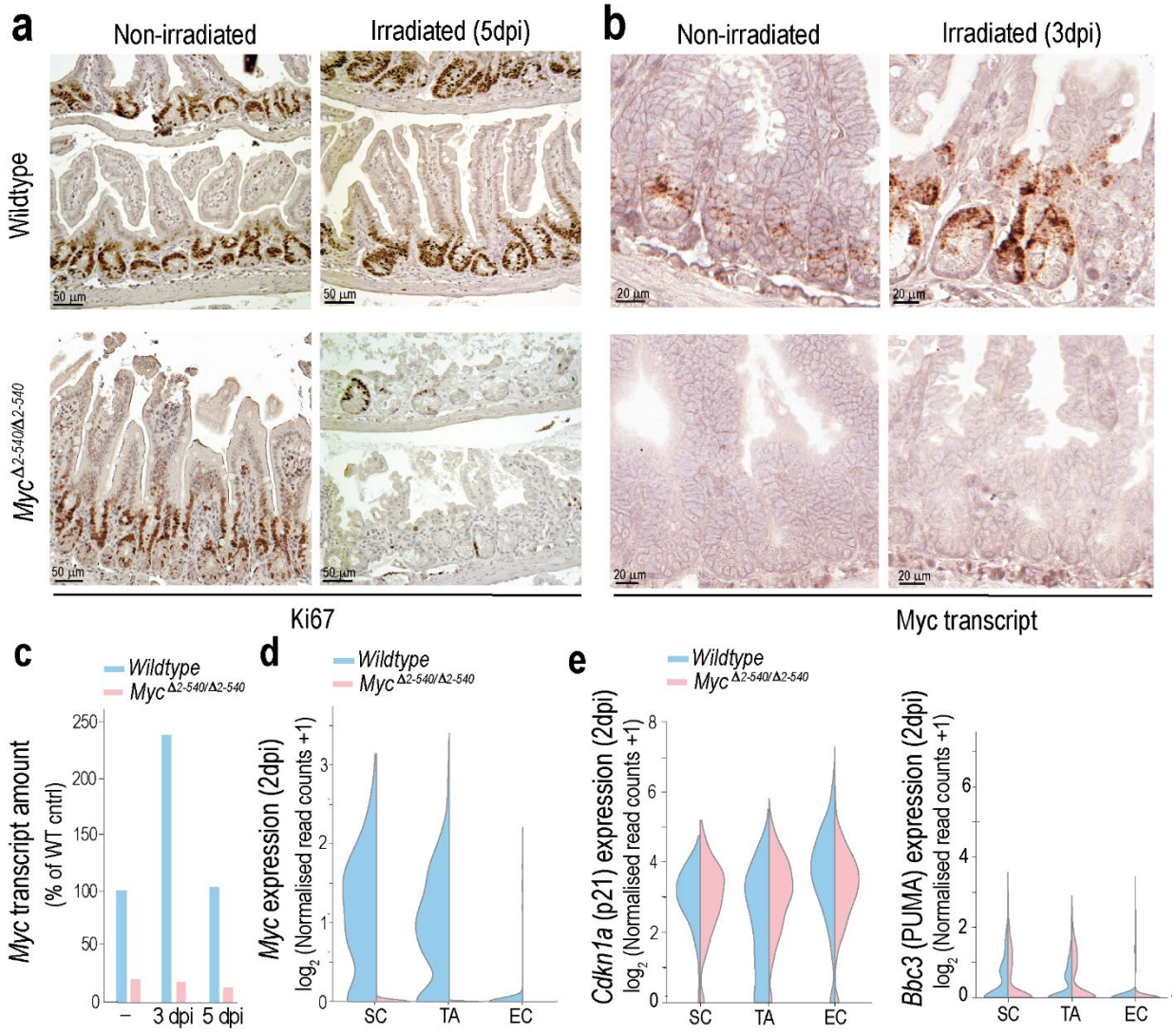

Fig. 2. *Myc$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$* intestine fails to induce *Myc* and regenerative proliferation despite a normal p53-mediated DNA damage response post γ-irradiation. a, Failure to regenerate the crypts is associated with loss of proliferation as shown by IHC for proliferation marker Ki67. **b-d,** *Myc$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$* intestine is unable to upregulate *Myc* expression after irradiation compared to the wildtype (**b**): ISH; (**c**): qPCR, wildtype: n=2 (control), n=3 (3dpi), n=3 (5dpi); *Myc$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$*: n=2 (control), n=2 (3dpi), n=3 (5dpi); (**d**): scRNA-seq, violin plots depicting expression of *Myc* in SC, TA and EC populations are shown. **e,** violin plots showing that unlike *Myc* levels, the levels of p53-mediated DNA damage response indicators *Cdkn1a* (p21) and *Bbc3* (PUMA) were comparable between wildtype and *Myc$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$* intestine at 2dpi (scRNA-seq). SC: stem cell, TA: transit amplifying cell, EC: enterocyte.

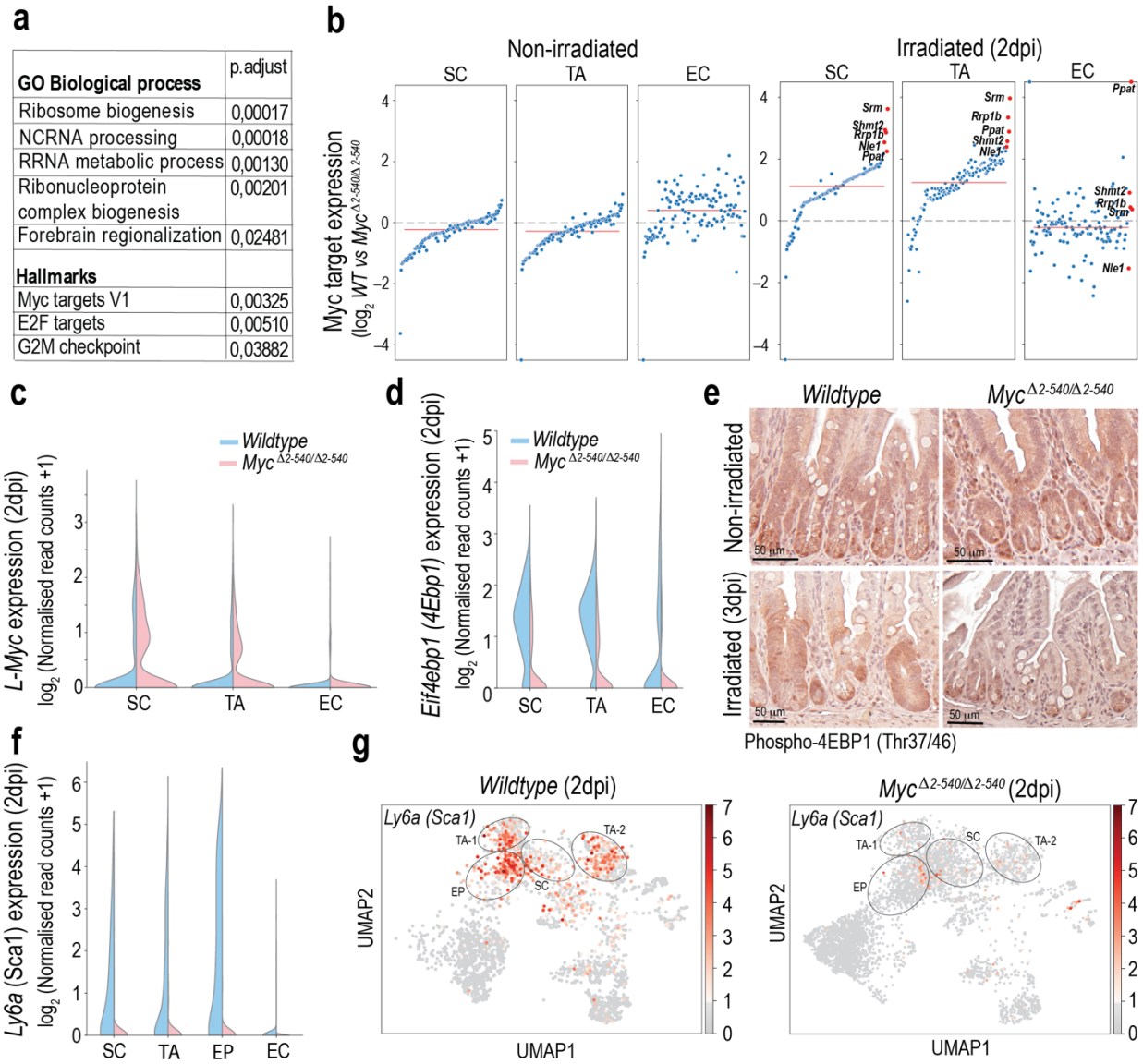

Fig. 3. *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ intestine fails to induce MYC targets after irradiation, compromising mTORC signalling and stem cell mobilization. a, GSEA analysis showing the most enriched gene ontology and hallmarks gene sets in the intestinal SC cluster of irradiated wildtype mice compared to *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice. **b,** Deletion of Myc$^{2\text{-}540}$ region causes a modest decrease in conserved MYC targets expression whereas after irradiation strong induction of MYC targets is seen in wildtype compared to *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice. Red line denotes the median expression of MYC targets. Red dots denote the top 5 differentially induced MYC targets in the stem cell compartment. **c,** Modest decrease of MYC targets during homeostasis is associated with increased expression of *L-Myc* in the intestinal stem and transit amplifying cells of *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice compared to wildtypes (also see Extended data Fig. 4). **d-e,** Altered regulation of MYC and mTORC target Eif4ebp1 (4EBP1) (d) scRNA-seq data (e) IHC showing amount of phosphorylated 4EBP1 before and after radiation. **f-g,** scRNA-seq data: violin plots showing loss of Ly6a/Sca1 induction in *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ mice compared to wildtype after irradiation.
SC: stem cell; TA: transit amplifying cell, EP: enterocyte progenitor; EC: enterocyte.

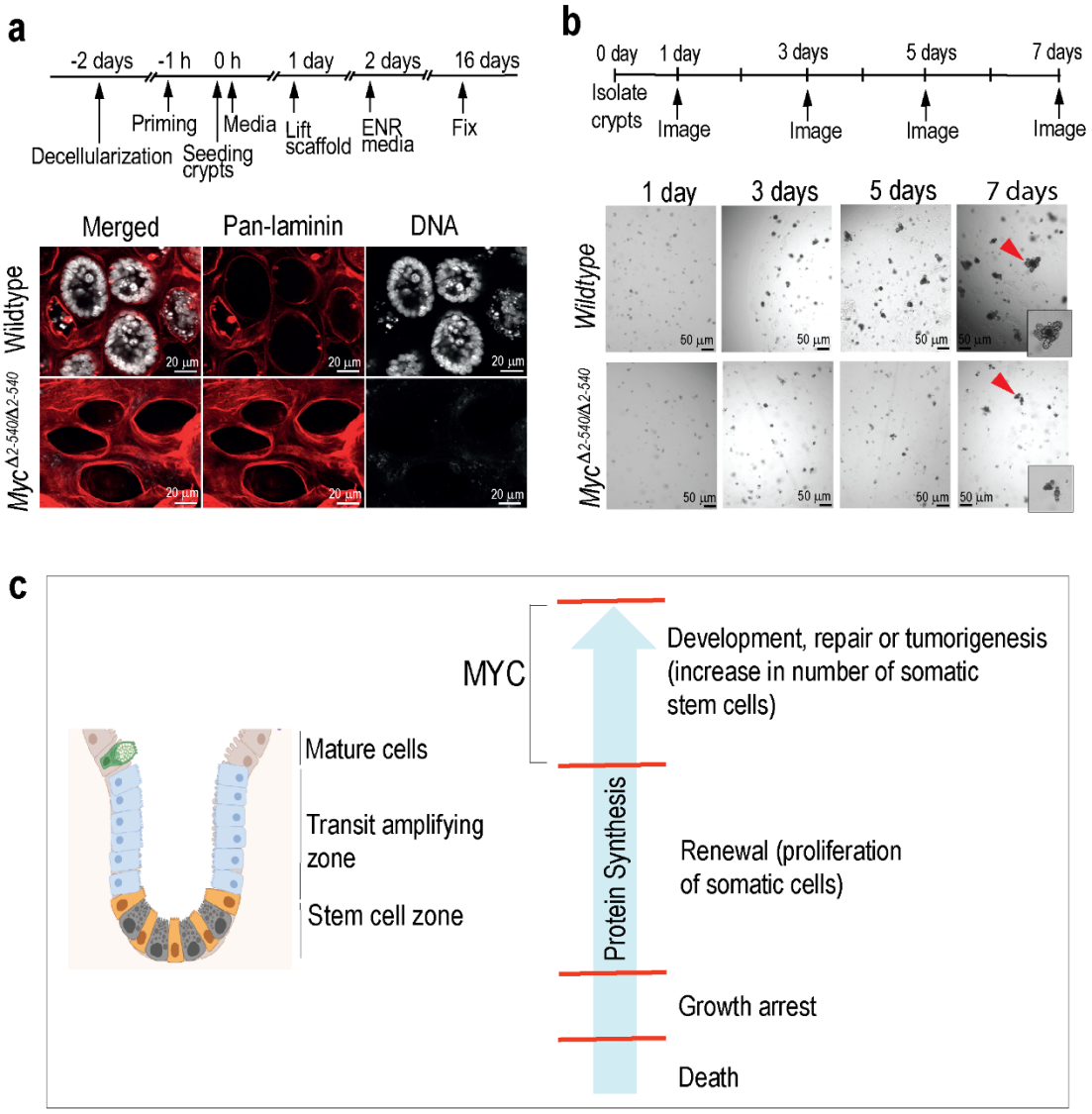

Fig. 4. *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ intestine fails to support regenerative growth *in vitro* 3-D cultures a, Immunofluorescence analysis showing failure of *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ crypts to re-epithelialize intestinal scaffolds. Representative images from two independent experiments are shown. **b,** The inability to re-epithelialize the scaffolds is correlated with the inability of crypts from adult *Myc*$^{\Delta 2\text{-}540/\Delta 2\text{-}540}$ intestine to support organoid growth in culture. Representative images from 4 independent experiment are shown. **c,** Model describing how different levels of protein synthesis can result in different proliferative cell fate choices ranging from normal homeostasis to regeneration and cancer.