

Intestinal label-retaining cells are secretory precursors expressing Lgr5

Simon J. A. Buczacki¹, Heather Ireland Zecchini¹, Anna M. Nicholson¹, Roslin Russell¹, Louis Vermeulen¹, Richard Kemp¹ & Douglas J. Winton¹

The rapid cell turnover of the intestinal epithelium is achieved from small numbers of stem cells located in the base of glandular crypts. These stem cells have been variously described as rapidly cycling or quiescent. A functional arrangement of stem cells that reconciles both of these behaviours has so far been difficult to obtain. Alternative explanations for quiescent cells have been that they act as a parallel or reserve population that replace rapidly cycling stem cells periodically or after injury; their exact nature remains unknown. Here we show mouse intestinal quiescent cells to be precursors that are committed to mature into differentiated secretory cells of the Paneth and enteroendocrine lineage. However, crucially we find that after intestinal injury they are capable of extensive proliferation and can give rise to clones comprising the main epithelial cell types. Thus, quiescent cells can be recalled to the stem-cell state. These findings establish quiescent cells as an effective clonogenic reserve and provide a motivation for investigating their role in pathologies such as colorectal cancers and intestinal inflammation.

The epithelial lining of the small intestine is continuously renewed from a small number of stem cells including Lgr5-expressing crypt base cells that have been shown to be a rapidly cycling stem-cell population in homeostasis¹. Alternative markers of intestinal stem cells have variously identified populations as rapidly cycling or quiescent with proposed roles for the latter as a parallel or reserve stem-cell population that act to replace rapidly cycling stem cells either periodically or after injury^{2–7}. Periodic replacement however, implies a hierarchical relationship that contrasts with the recent demonstration that the ongoing loss and replacement of rapidly cycling stem cells explains all the cellular output of the crypt^{8,9}. Hence, contemporary views consider the possibility of a reserve stem-cell population that is distinct from the rapidly cycling population responsible for homeostasis^{10,11}. Quiescent or slowly cycling stem-cell populations are defined by the property of label retention. Under conditions of expansion of the stem-cell compartment (during the later stages of gut growth or during regeneration after epithelial damage), proliferating cells in the intestine can be labelled with nucleotide analogues. Small numbers of label-retaining cells (LRCs) persist for around 4 weeks after homeostasis is (re)established¹². On subsequent regenerative challenge some LRCs can re-enter the cell cycle thereby demonstrating their proliferative potential¹³. Quiescent cells identified in this way are predominantly located in a supra-Paneth cell position within the crypt. Several candidate markers or regulators of quiescence with expression patterns that overlie this location have been identified. Bmi1-expressing cells in the proximal small intestine are capable of clonal expansion, have low Ki67 positivity and are responsible for repopulation of the Lgr5⁺ stem-cell population after conditional Lgr5 cell deletion^{4,11}. Wip1 phosphatase regulates stem-cell apoptosis in a p53-dependent manner and co-localizes with thymidine label-retaining cells in a supra-Paneth cell position¹⁴. The tumour suppressor Pten acts to limit the numbers of intestinal stem cells and its inactive phosphorylated form has been shown to co-localize with 5-bromodeoxyuridine (BrdU)-positive label-retaining cells^{15,16}. It has also been shown that Hopx marks cells in the +4 position, a small proportion of which are label-retaining and that

are also capable of converting to a Lgr5 population and maintaining clonogenic growth⁶. Most recently two publications dispute the role of the negative regulator of ErbB, Lrig1. One proposes that it is a unique marker of quiescent stem cells distinct from the bulk Lgr5 population, whereas the other identifies it as a general regulator of the stem-cell compartment^{3,7}. Furthermore, a recent report has demonstrated that all previously identified ‘quiescent stem-cell’ markers appear highly expressed by the Lgr5 population¹⁷. In summary, the exact nature of quiescent crypt cells remains unknown because candidate markers rather than the label-retaining population itself have been the focus of study.

Identification of label-retaining cells

To study quiescence directly we identified cells retaining nuclear-localized fluorescent H2B-YFP during a chase period following a pulse of induced expression. Transgenic *Cypla1-H2B-YFP* mice were evaluated before and after a pulse of β -naphthoflavone (BNF) treatment (Fig. 1a, b). The *Cypla1* promoter has an established pattern of expression on induction in all cells of the crypt-villus axis with the exception of the mature Paneth cells¹⁸. Twenty-four hours after induction (T0) cells throughout the crypt to villus axis expressed H2B-YFP with, as expected, the exception of Paneth cells (Fig. 1b–d and Supplementary Video 1). By 7 days crypt expression was restricted to multiple cells of the crypt base (Supplementary Fig. 1a). These cells decrease in frequency with time, becoming absent by 12 weeks (Supplementary Fig. 1b–d). Confocal microscopy analysis after 1 week suggested that a number of positive cells were Paneth cells due to characteristic enlarged and rounded nuclear morphology; this was confirmed using Paneth cell markers (Supplementary Fig. 1e, f). Non-Paneth LRCs were selected as the population of interest and, from 10 days post-induction, were defined as YFP-labelled LRCs (YFP-LRCs) in accord with published definitions of intestinal LRCs (Fig. 1e)^{5,6,12}. YFP-LRCs were present throughout the crypt base, slightly predominating in the +3 position, and were identifiable in decreasing numbers for up to 4 weeks (Fig. 1f, g). Subsequently, YFP⁺ Paneth cells were seen up to 8 weeks after this time (that is 12 weeks

¹Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK.

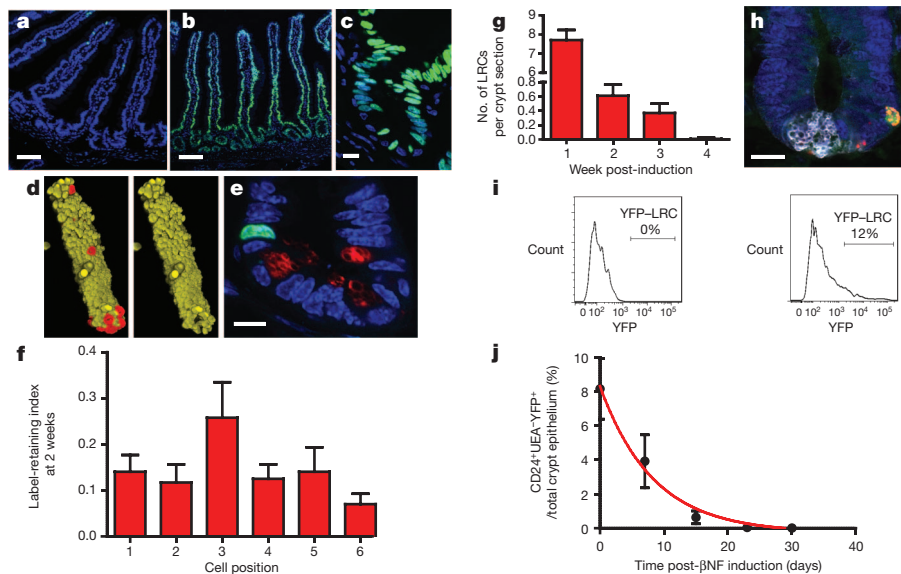


Figure 1 | Identification and isolation of YFP-LRCs. **a–f**, Time course images of intestinal epithelium from *Cyp1a1-H2B-YFP* mice after induction with βNF. **a**, Uninduced control. A very small number of background YFP⁺ cells are present on duodenal villus tips (**b**, **c**) T0. **d**, T0 three-dimensional reconstructed z-stack image of an isolated crypt showing that Paneth cells identified with UEA (red) do not initially express YFP (yellow). **e**, Three weeks post-induction image demonstrating a typically positioned YFP⁺ cell above UEA⁺ (red) Paneth cells. **f**, Non-Paneth YFP⁺ LRCs are distributed

post-induction) but not thereafter, which accords with the previously reported lifespan of terminally differentiated Paneth cells of 6–8 weeks¹⁸. By 3 weeks YFP-LRCs were reduced to 0.4 ± 0.13 per crypt section, approximating to 2 per whole crypt. No YFP-LRCs were seen in the colon.

The above estimate for YFP-LRC number is 50 times higher than previously described for LRCs identified by retention of nucleotide analogues (0.008 per crypt section)⁶. Confocal microscopy on dual pulse-chased EdU/H2B-YFP mice determined that EdU-LRCs were recognized at a frequency of 0.009 per crypt section and that EdU-LRCs (27 out of 30) are YFP⁺, demonstrating that the majority (90%) of EdU-LRCs can be identified by YFP label retention (Fig. 1h). The greater abundance of YFP-LRCs (2–3 per crypt) demonstrates that nucleotide labelling incompletely marks slowly cycling populations, presumably because the method only marks cells captured in S phase before acquiring quiescence¹⁹.

LRCs have a combined secretory and stem-cell signature

After tissue dissociation, YFP-LRCs were isolated by flow sorting for transcriptional profiling. A strategy combining UEA lectin (for Paneth cell selection) and CD24 (lower crypt specific) was used to compare three populations independent of stem markers associating with cell cycle status (Fig. 1i, j and Supplementary Figs 1f, g and 2a, b)⁷. Expression microarray analysis showed that YFP-LRCs (CD24⁺UEA⁻YFP⁺) are distinct from both Paneth (CD24⁺UEA⁺) and cycling lower crypt cells (LCCs) (CD24⁺UEA⁻YFP⁻) (Supplementary Fig. 2c). Notably, transcripts associated with the Paneth and enteroendocrine lineages (*Mmp7*, *Kit*, *Chga*, *Gip*, *Pax6*) as well as stem-cell markers (*Lgr5*, *Lrig1*, *CD133*, *CD44*, *Peg3*) were more abundant in YFP-LRCs compared to LCCs. To establish that the YFP-LRC signature did not arise as a composite of the other two differentiated cell populations, we undertook single-cell profiling of 48 cells from each of the three groups for 47 transcripts selected either as stem cell/differentiation markers or YFP-LRC-specific from the array comparison (see Methods). Principal component analysis demonstrated that YFP-LRCs were a discrete homogeneous population separate from both Paneth cells and LCCs but with no internal structure

throughout the crypt base and occur maximally at cell position +3 (mean ± s.e.m.). **g**, Non-Paneth YFP⁺ LRCs are present for up to 4 weeks after βNF induction (mean ± s.e.m.). **h**, Representative image of an EdU⁺YFP⁺ non-Paneth LRC. EdU, red; YFP, green; lysozyme, white. **i**, FACS YFP histograms of single, live, UEA⁻CD24⁺ cells in uninduced (left panel) and 10 days post-βNF-treated animals (right panel). **j**, The CD24⁺UEA⁻YFP⁺ population decreases in number with time (mean ± s.e.m.). Scale bars: 50 μm (**a**, **b**); 10 μm (**c**, **e**, **h**).

(Fig. 2a). *Nfatc3*, *CD83* and *Nfat5* were the top three principal components defining YFP-LRCs. It was confirmed that larger numbers of YFP-LRCs expressed higher levels of Paneth/enteroendocrine cell markers than LCCs including *Mmp7* and *Chgb* (Fig. 2b). Homogeneity of the YFP-LRC population was further validated using immunofluorescence for differentiation markers on flow-sorted populations (Supplementary Fig. 3).

Unexpectedly, given the quiescent status of YFP-LRCs, the microarray showed elevated expression of *Lgr5* in YFP-LRCs compared to the LCCs. To determine the degree of similarity with the *Lgr5* population, we used previously published microarray data²⁰ and found a highly significant degree of overlap in overexpressed genes between YFP-LRCs and *Lgr5*-GFP^{hi} cells ($P < 0.0001$) (Fig. 2c). Furthermore, quantitative polymerase chain reaction with reverse transcription (qRT-PCR) comparison of YFP-LRCs with *Lgr5*-GFP^{hi} cells isolated from *Lgr5*-EGFP-IRES-creERT2 mice demonstrated equivalent expression not only of *Lgr5* but also ‘quiescent’ stem-cell markers *Tert*, *Lrig1* and *Hopx*. In addition, there was upregulated expression of Paneth and enteroendocrine lineage genes in YFP-LRCs compared to *Lgr5*-GFP^{hi} cells (Fig. 2d). These comparisons suggest that most YFP-LRCs are a subpopulation of *Lgr5*-expressing cells. This was confirmed using spectral detection confocal microscopy to overlay YFP and GFP fluorescence on *Cyp1a1-H2B-YFP/Lgr5*-EGFP-IRES-creERT2 mice to confirm that 98% (40 out of 41) of YFP-LRCs also expressed *Lgr5* (Fig. 2e).

Elevated *Lgr5* defines LRC maturation into Paneth cells

YFP-LRCs have a combined secretory and stem-cell signature, indicating that they may have a secretory fate. Confocal microscopy was first used to track YFP inheritance into the Paneth cell population from 10 days (when all YFP-LRCs are formally defined) to 21 days post-induction (Fig. 3a). YFP-LRCs decrease in number over time whereas the number of YFP⁺ Paneth cells increases (Fig. 3b). This reciprocal relationship shows that one fate of YFP-LRCs is differentiation to Paneth cells. Moreover, the increasing number of YFP⁺ Paneth cells over the 11-day chase accounts for 3.6 new Paneth cells per crypt (see Methods). This chase period is relatively short compared

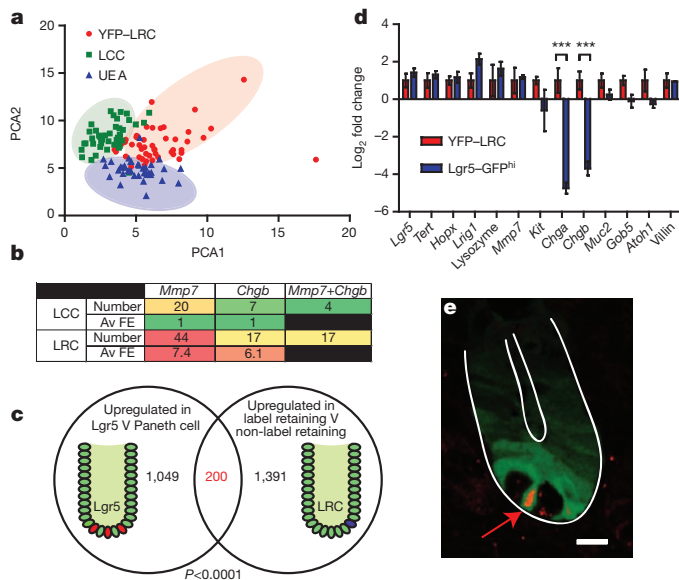


Figure 2 | YFP-LRCs are a discrete Lgr5-expressing subpopulation.

a, Principal component analysis (PCA) on single-cell Fluidigm qRT-PCR data from 48 single YFP-LRCs (red), LCCs (green) and CD24⁺UEA⁺ (blue) cells. **b**, Summary table of single-cell data showing the number of cells expressing secretory cell transcripts ($n = 48$) and the log fold change expression level of expressing cells (Av FE) relative to LCCs. **c**, Venn diagram demonstrating the overlap in overexpressed genes between YFP-LRCs and Lgr5⁺ cells. **d**, qRT-PCR data comparing the expression patterns of YFP-LRCs and Lgr5⁺ cells. Data are mean $\pm$ s.e.m. *** $P < 0.005$. Two-tailed Student's t -test. **e**, Representative confocal microscopy image of a Lgr5-GFP⁺/YFP-LRC cell. YFP, red; Lgr5, green. Scale bar, 10 μ m.

to the turnover time of Paneth cells (42 days), indicating that the direct conversion of LRCs to Paneth cells without any cell division could create around 14 ($3.6 \times 42/11$) Paneth cells per crypt over this time. With an estimated 15 Paneth cells per crypt it is likely that most, if not all, Paneth cells originate via LRCs.

We next investigated whether a subset of Lgr5⁺ cells could be identified with maturing Paneth cell features independent of the label-retaining marker and further differentiated in Paneth cell status to YFP-LRCs. A subset of Lgr5-GFP^{hi} cells that stained with UEA as a marker of Paneth cells was identified and isolated by flow sorting and confirmed to be enriched for Paneth cell transcripts *Mmp7* and *Defa5* (Fig. 3c–f).

To determine alternative fates for YFP-LRCs, the upper crypt region was analysed at 14 days post-induction for the presence of YFP⁺ cells transiting from the crypt base. Such cells are infrequent owing to the rapid rate of migration through the crypt (Fig. 3g)²¹. The vast majority of YFP⁺ cells (24 out of 25) present above the crypt base were chromogranin A⁺ enteroendocrine cells. This proportion effectively eliminates the numerically more abundant absorptive or Goblet cells as arising from YFP-LRCs. Overall, LRCs appear to be a Paneth cell precursor population but also capable of some enteroendocrine differentiation.

LRCs possess multi-lineage growth potential

To determine whether YFP-LRCs are irreversibly committed to a secretory fate we functionally assessed their growth potential in organoid culture²². YFP-LRCs and Lgr5-GFP^{hi} cells plated at the same densities were roughly equivalent in ability to form organoids (0.15% versus 0.25%). The organoids from YFP-LRCs demonstrated crypt budding at around 7 days and were found to resemble intact crypt-villus units as previously described (Supplementary Fig. 4). YFP-LRCs are therefore capable in a regenerative setting of acquiring stem-cell characteristics.

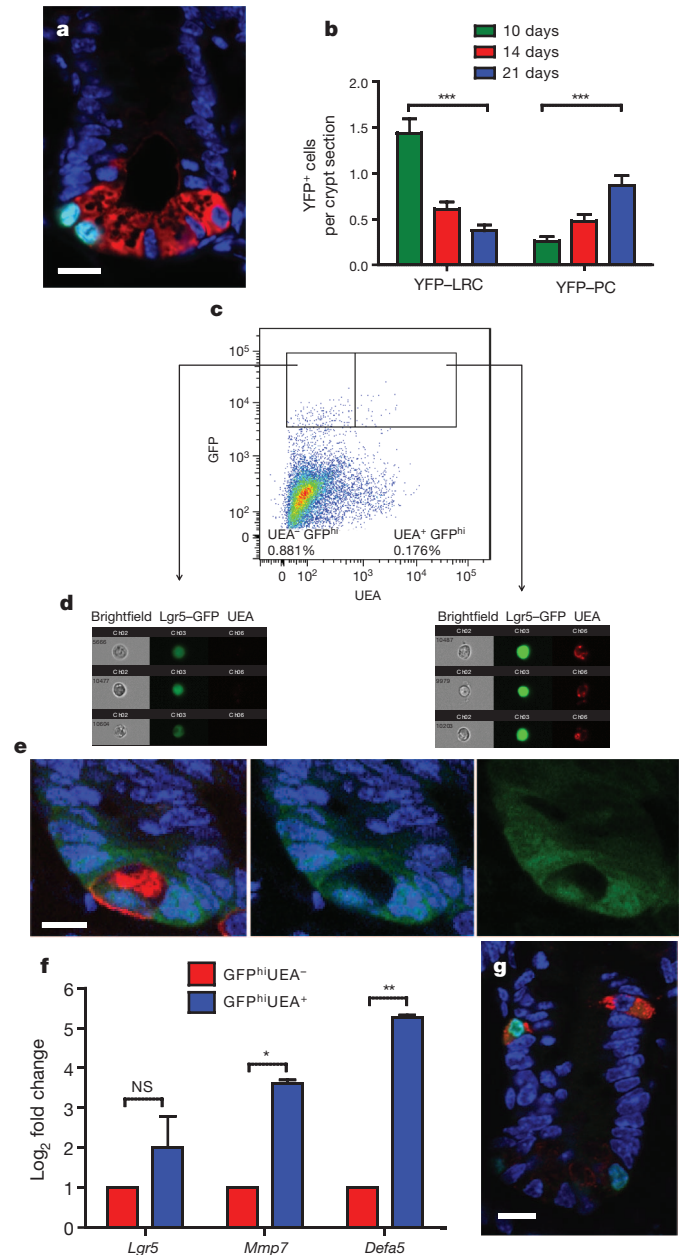


Figure 3 | Acquisition of Paneth and enteroendocrine cell characteristics from YFP-LRCs. **a**, Image demonstrating that after 4 weeks all YFP⁺ cells are Paneth cells as confirmed by lysozyme staining. **b**, Assessment of the number of YFP⁺ cells between 10 and 21 days demonstrates a reciprocal relationship in cell types: YFP⁺ Paneth cells (YFP-PC) increase in frequency as YFP-LRCs decrease (mean $\pm$ s.e.m.); $n = 100$. *** $P < 0.001$. Mann-Whitney U -test. **c**, FACS analysis showing that around 15% of Lgr5-GFP^{hi} cells are marked by UEA. **d**, Image-stream images showing examples of Lgr5-GFP^{hi}UEA[−] and Lgr5-GFP^{hi}UEA⁺ cells from **c**. Note the greater intensity of GFP in the latter. **e**, Confocal image showing a GFP⁺UEA⁺ cell. Left panel: UEA (red), Lgr5-GFP (green) and DAPI (blue). Middle panel: Lgr5-GFP and DAPI. Right panel: Lgr5-GFP alone. **f**, qRT-PCR data showing upregulation of Paneth cell transcripts *Mmp7* and *Defa5* in the GFP^{hi}UEA⁺ subpopulation. (mean $\pm$ s.e.m.); * $P < 0.05$; ** $P < 0.01$. Two-tailed Student's t -test. NS, not significant. **g**, Representative image of rare YFP⁺ chromogranin A⁺ enteroendocrine cells in the upper crypt at 2 weeks post-induction. Scale bars: 10 μ m (**a**, **e**, **g**).

LRCs contribute to the stem-cell pool only after injury

To functionally assess whether YFP-LRCs have stem-like characteristics *in vivo* we used a novel dimerizable Cre recombinase strategy (Fig. 4)²³. *dicreAB* mice are capable of clonally marking cells when two

inactive Cre fragments heterodimerize to form functional recombinase by the binding of Ariad rapamycin analogues to FKHB-ligand-binding domains present in both fragments. The carboxy-terminal peptide of Cre (CreB) is expressed constitutively from the *Rosa26* locus whereas the amino-terminal peptide (CreA) is expressed as an H2B-CreA fusion protein under the control of the β NF inducible promoter, *Cyp1a1* (Fig. 4a, b). Treatment of double transgenic *Rosa26-creB* and *Cyp1a1-H2B-creA* mice with β NF induces expression of

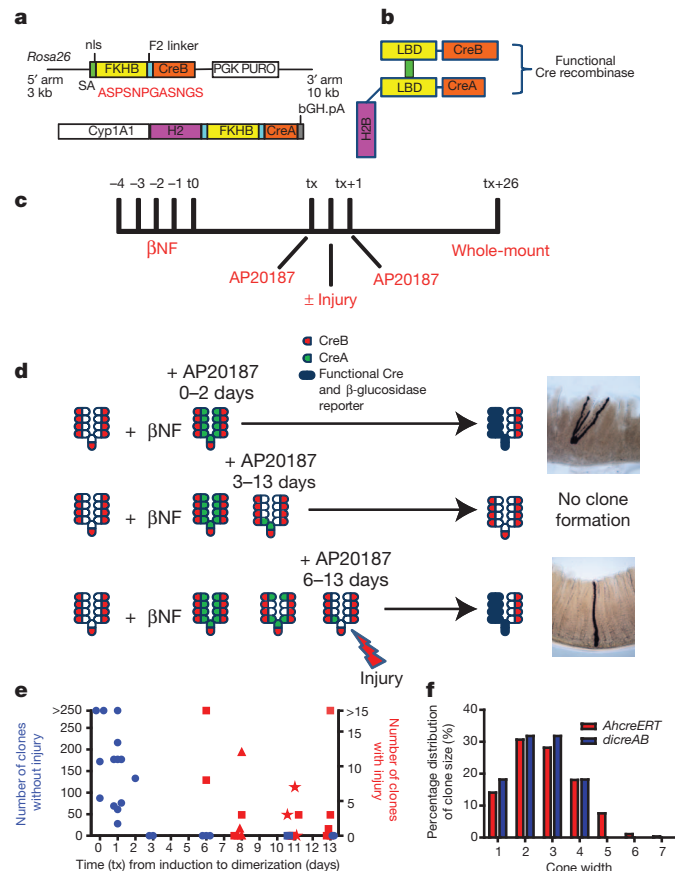


Figure 4 | LRCs only demonstrate clonogenicity after injury *in vivo*.

a, Diagrammatic representation of diCreAB construct showing position of FKHB dimerizing domain in protein fusions to both N terminus (*H2B-creA*) and C terminus (*creB*) regions of a split *cre* gene. **b**, Representation of the *in vivo* dimerization of CreA and CreB to form a functional Cre recombinase. Dimerizing agent (AP20187) (green) mediates dimerization of two FKHB domains thereby creating a functional Cre complex. **c**, Generic schematic of induction protocol to test *in vivo* clonogenicity of cells expressing *creA*. β NF induces *H2B-creA* against a background of constitutive *creB* expression. With time (tx) following β NF administration *H2B-creA* is first expressed and then becomes restricted. Expressing cells are tested for their ability to generate clones after treatment with dimerizing agent (AP20187) that allows a functional Cre complex to form and thereby mediate recombination of a *cre* reporter gene. **d**, Schematic showing conditions for clone formation. AP20187 is sufficient for clone formation 0–2 days post β NF (top row) but not between 3–13 days (middle row). Between 6–13 days combined injury and AP20187 are required for clone formation (bottom row). In the clone shown, the injury was hydroxyurea. **e**, Numbers of clones formed over the total murine intestinal tract in the absence of injury (blue) and in the presence of injury (red) and showing that treatment with AP20187 only generates clones in the absence of injury up to 2 days after β NF treatment. When AP20187 treatment is combined with injury, clones are generated up to 14 days post β NF induction of *H2B-creA* (squares, hydroxyurea; stars, irradiation; triangles, doxorubicin). **f**, Clone width (measured in cell diameters) comparisons from data acquired at 3 weeks after induction of clones with β NF in *AhcreERT* mice ($n = 277$)⁸ and 3 weeks after dimerization of Cre with AP20187 in *dicreAB* mouse ($n = 23$) induced with β NF and then dimerized and injured 13 days later.

H2B-creA throughout the epithelium on a background of constitutive *creB* expression. With time after β NF induction only LRCs retain H2B-CreA, in a manner analogous to H2B-YFP. Administration of intravenous dimerizing agent causes the two fragments of Cre to form a functional protein with subsequent downstream reporter expression.

By varying the interval between β NF induction and Cre dimerization we show that, in the absence of injury, crypt cells form clones only up to 3 days between β NF induction and dimerization (Fig. 4c, d). With longer induction and dimerization intervals (>3 days) clone formation is not observed. However, induction–dimerization intervals of up to 14 days induced clone formation when combined with intestinal injury (Fig. 4e). These clones were noted to contain all differentiated cell lineages at appropriate frequency (Supplementary Fig. 5). Clone size analysis compared to work we have previously published for pulse–chase clone induction after tamoxifen treatment demonstrates that LRC-derived clones occupy the predicted size distribution expected for clones of equivalent age (3 weeks) and that arises from clone expansion due to neutral drift and leads to clone fixation as clones occupy whole crypts (Fig. 4f)⁸.

These results demonstrate that rapidly cycling stem cells (not having diluted CreA within 3 days) are, as expected, competent to produce clonal progeny in the steady state. The lack of clone formation with longer induction–dimerization intervals could either be due to insufficient bioavailability of H2B-CreA for recombination or that H2B-CreA-retaining cells are not clonogenic. The injury models indicate the latter, demonstrating that LRCs do not maintain the stem-cell pool during homeostasis. Under conditions of injury/regeneration LRCs are capable of clonogenic growth by recall to the self-renewing pool of stem cells. To our knowledge this is the first demonstration of *in vivo* lineage tracing based on label retention in a mammalian system.

Mature Paneth cells are not proliferative

The differentiation of YFP-LRCs into Paneth cells has the consequence that both cell types co-exist. To confirm that Paneth cells are indeed terminally differentiated and not able to proliferate after damage, H2B-YFP mice were induced with β NF in a pulse–chase experiment with a chase period (5 weeks) that extended beyond the lifespan of YFP-LRCs. These animals were then irradiated and the response of the residual population of YFP⁺ Paneth cells determined. In control (non-irradiated) animals YFP⁺ cells were invariably isolated Paneth cells. After radiation treatment (at 1, 2 and 7 days), intestinal sections were examined for the doublets or strings of YFP⁺ cells that would arise if Paneth cells were undergoing post-irradiation division. The pattern of YFP⁺ positivity remained unchanged from the controls, suggesting that Paneth cells do not divide (Supplementary Fig. 6).

Discussion

There are a number of alternative functional roles for a quiescent population. First, they could be a stem population co-existing and in dynamic equilibrium with the rapidly cycling cells⁶. The dynamics of stem-cell renewal, however, have shown that the turnover of a single population of rapidly dividing cells is sufficient to drive changes in clone size distribution, leading to clones that occupy whole crypts^{8,9}. Second, they could be a dormant or reserve stem-cell population that is only active under conditions of epithelial stress or injury¹⁰. Third, they could be terminally differentiated Paneth cells². Last, and not previously described, they could be committed to differentiate and no longer contribute to stem-cell functions.

Our observations establish the functional role of LRCs. They are normally destined to become Paneth and enteroendocrine cells but retain the ability to reacquire stem-cell function and can be recruited to serve as a functional clonogenic stem population under conditions of regeneration. Historically it is accepted that following stem-cell ablation cells normally destined for differentiation can regenerate the stem-cell compartment¹². However, the current findings demonstrate

directly that a committed precursor can be recalled to the stem-cell pool after damage in a similar manner to that shown recently for Dll1⁺ secretory precursors²⁴.

Our observations resolve the apparent paradox that quiescence markers are expressed within a population of Lgr5⁺ cells that are viewed as rapidly cycling¹⁷. Around 20% of Lgr5⁺ cells are largely quiescent and continue to express Lgr5 before Paneth cell maturation.

Lgr5⁺ cells have been implicated as cancer stem cells in colonic cancers, and higher frequencies of Lgr5⁺ cells in intestinal cancers correlate with more extensive Paneth cell differentiation²⁵. Paneth cells may nurture cancer stem cells by niche generation as they do normal stem cells in culture. It will be important to determine the extent to which quiescent secretory intermediates retaining clonogenic potential are maintained in colon cancers and establish the factors that activate them.

METHODS SUMMARY

Mouse models and inductions. *Cyp1a1-H2B-YFP* mice have previously been described⁸. *dicreAB* mice were generated as follows: *creA* and *creB* are two segments of the *cre* coding sequence comprising codons 19–59 and 60–343, respectively, and were fused to other protein motifs with flexible F2 linkers as described previously²³ (see Methods for details).

Cyp1a1 was induced by $\times 3$ intraperitoneal β NF injections over 24 h (80 mg kg⁻¹). Dimerization of Cre fragments was performed by $\times 2$ intravenous AP20187 (Ariad) injections (10 mg kg⁻¹).

Single-cell qRT-PCR. FACS-sorted single cells underwent single-cell RNA amplification (see Methods for details). Generated cDNA libraries were then used for Fluidigm Biomark high-throughput qRT-PCR against custom-designed and validated primers (see Methods for details). Principal component analysis was carried out in R.

Quantitative PCR. qRT-PCR was performed using standard TaqMan or SybGreen assays. Custom primers were validated before use. For experiments where cell numbers were limiting, cDNA was amplified using a custom protocol (see Methods for details). Fluidigm qRT-PCR was performed as per the manufacturer's protocol.

Full Methods and any associated references are available in the online version of the paper.

Received 26 July 2012; accepted 30 January 2013.

Published online 27 February 2013.

1. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003–1007 (2007).
2. Roth, S. *et al.* Paneth cells in intestinal homeostasis and tissue injury. *PLoS ONE* **7**, e38965 (2012).
3. Powell, A. E. *et al.* The Pan-ErbB negative regulator Lrig1 is an intestinal stem cell marker that functions as a tumor suppressor. *Cell* **149**, 146–158 (2012).
4. Sangiorgi, E. & Capecchi, M. R. Bmi1 is expressed *in vivo* in intestinal stem cells. *Nature Genet.* **40**, 915–920 (2008).
5. Montgomery, R. K. *et al.* Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. *Proc. Natl Acad. Sci. USA* **108**, 179–184 (2011).
6. Takeda, N. *et al.* Interconversion between intestinal stem cell populations in distinct niches. *Science* **334**, 1420–1424 (2011).
7. Wong, V. W. *et al.* Lrig1 controls intestinal stem-cell homeostasis by negative regulation of ErbB signalling. *Nature Cell Biol.* **14**, 401–408 (2012).

8. 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).
9. Snippert, H. J. *et al.* Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010).
10. Li, L. & Clevers, H. Coexistence of quiescent and active adult stem cells in mammals. *Science* **327**, 542–545 (2010).
11. Tian, H. *et al.* A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. *Nature* **478**, 255–259 (2011).
12. Potten, C. S., Hume, W. J., Reid, P. & Cairns, J. The segregation of DNA in epithelial stem cells. *Cell* **15**, 899–906 (1978).
13. Potten, C. S., Owen, G. & Booth, D. Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J. Cell Sci.* **115**, 2381–2388 (2002).
14. Demidov, O. N. *et al.* Wip1 phosphatase regulates p53-dependent apoptosis of stem cells and tumorigenesis in the mouse intestine. *Cell Stem Cell* **1**, 180–190 (2007).
15. He, X. C. *et al.* BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt- β -catenin signaling. *Nature Genet.* **36**, 1117–1121 (2004).
16. He, X. C. *et al.* PTEN-deficient intestinal stem cells initiate intestinal polyposis. *Nature Genet.* **39**, 189–198 (2007).
17. Muñoz, J. *et al.* The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '4' cell markers. *EMBO J.* **31**, 3079–3091 (2012).
18. Ireland, H., Houghton, C., Howard, L. & Winton, D. J. Cellular inheritance of a Cre-activated reporter gene to determine Paneth cell longevity in the murine small intestine. *Dev. Dyn.* **233**, 1332–1336 (2005).
19. Foudi, A. *et al.* Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. *Nature Biotechnol.* **27**, 84–90 (2009).
20. Sato, T. *et al.* Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469**, 415–418 (2011).
21. Cheng, H. & Leblond, C. P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. *Am. J. Anat.* **141**, 537–561 (1974).
22. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
23. Jullien, N., Sampieri, F., Enjalbert, A. & Herman, J. P. Regulation of Cre recombinase by ligand-induced complementation of inactive fragments. *Nucleic Acids Res.* **31**, e131 (2003).
24. van Es, J. H. *et al.* Dll1⁺ secretory progenitor cells revert to stem cells upon crypt damage. *Nature Cell Biol.* **14**, 1099–1104 (2012).
25. Lewis, A. *et al.* Severe polyposis in *Apc*(1322T) mice is associated with submaximal Wnt signalling and increased expression of the stem cell marker Lgr5. *Gut* **59**, 1680–1686 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements This research was supported by Cancer Research UK (S.J.A.B., H.I.Z., A.M.N., R.R., R.K. and D.J.W.). L.V. was supported by a KWF fellowship. We thank D. Tan for providing the single-cell RNA amplification protocol. We also thank M. de la Roche, R. von Furstenberg and S. Henning for advice with the *in vitro* culture work. We acknowledge the following core facilities at CRUK/CRI: The Transgenic Laboratory, Biological Resource Unit, Flow Cytometry, Histopathology, Microscopy, Genomics and Bioinformatics and the CRUK Paterson Institute Microarray Facility. We thank R. J. Davies, A. Klein and A. Ibrahim for manuscript discussions.

Author Contributions S.J.A.B. designed and performed experiments and wrote the paper. H.I.Z. designed and developed the H2B-YFP model and performed experiments. R.R. performed the bioinformatic analysis. A.M.N. and L.V. performed experiments. R.K. designed experiments and performed bioinformatic analysis. D.J.W. designed experiments, developed the diCreAB model and wrote the paper.

Author Information Data were deposited to the GEO database under accession number GSE43772. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.J.W. (doug.winton@cruk.cam.ac.uk).

METHODS

Mice models. *Cyp1a1-H2B-YFP* mice have previously been described⁸. Briefly, *Cyp1a1-H2B-YFP* mice were created using conventional cloning techniques involving removing the human *H2B* coding sequence from the pBOS-H2BGFP vector (BD Pharmingen) and then ligating this into pIRES-eYFP (BD Biosciences). *Lgr5-EGFP-IRES-creERT2* mice were purchased from the Jackson Laboratory.

dicreAB mice were generated as follows: *creA* and *creB* are two segments of the *cre* coding sequence comprising codons 19–59 and 60–343, respectively, and were fused to other protein motifs with flexible F2 linkers as described previously²³. The FKBP dimerizing domain and AP20187 was a gift provided by Ariad Pharmaceuticals. *creA* was synthesized as a single fusion gene comprising in 5' to 3' order *H2B-F2 linker-FKBP-F2 linker-creA-pA* with flanking BglII sites to permit excision.

H2B-F2 linker-FKBP-F2 linker-creA-pA was subcloned using BglII into pAh1R1 such that the fusion gene was under the control of 12-kb of the rat *CYP1A1* promoter element²⁶. The cassette comprising *Cyp1a1-H2B-F2 linker-FKBP-F2 linker-creA-pA* (or *Cyp1a1-H2B-LBD-creA*) was excised from the plasmid backbone using NotI and this fragment purified for pronuclear injection. Oocyte injection was performed by the Transgenic Lab, within the CRI.

CreB was synthesized as *FKB-F2 linker-creB-pA* with flanking NheI sites for excision (sequences available on request). This cassette was cloned in the NheI site of *ROSMTs13* as described previously²⁷. The linearized plasmid was introduced into ES cells and selected with puromycin. The S6B6 ES cell line was derived from 129/sv/C57BL6/j hybrid embryos. Initial screening of ES-derived colonies was confirmed by long-range PCR and sequencing and mice generated as previously described²⁸.

Mice were bred and housed according to UK Home Office guidelines. For induction of *Cyp1a1-H2B-YFP* mice, 3 intraperitoneal injections of 80 mg kg⁻¹ of β NF over a 36-h period were used. T0 was defined as the same day as the last β NF injection. For inductions of *dicreAB* mice, 5 daily intraperitoneal injections of 80 mg kg⁻¹ β NF were used. Dimerization of Cre was performed using two intravenous injections of AP20187 10 mg kg⁻¹ (Ariad Pharmaceuticals) over 2 days. Injury models involved single doses: of hydroxyurea (intraperitoneally, 1 g kg⁻¹), doxorubicin (intraperitoneally, 10 mg kg⁻¹) that target DNA replicating cells (by inhibition of nucleotide incorporation and topoisomerase respectively) and γ -irradiation (6 Gy) that preferentially kills rapidly cycling cell as they attempt to replicate damaged DNA template.

Tissue preparation, immunofluorescence and whole-mounting. Preparation of tissues for immunofluorescence was performed by fixation in 4% PFA for 48 h followed by 20% sucrose for a further 48 h. Tissue was then embedded in OCT and sectioned at 6 μ m. Antigen retrieval was performed using 1% SDS for 5 min and then blocked with 5% NDS for 30 min. Sections were stained overnight at 4 °C for all primary antibodies and 2 h at room temperature for UEA. Secondary antibodies were applied for 1 h at room temperature. Nuclear staining was achieved using ProLong Gold antifade reagent with DAPI (Invitrogen). Sections were visualized on a C1Si Nikon confocal microscope.

For UEA and lysozyme co-expression experiments, primary anti-lysozyme was incubated overnight at 4 °C followed by incubation with UEA-AF647 and an AF555-anti-rabbit secondary for 2 h at room temperature. Differentiated Paneth cells were seen to overlay with both markers but in addition UEA marked a subset of goblet cells in both the upper crypt and villus.

Whole-mounting of intestines and detection of SYN β Glu activity was carried out as previously described⁸. Organoids were prepared for immunofluorescence as previously reported²⁹. Clone sizes were measured as previously described⁸. Comparisons were made between *dicreAB* clones at 3 weeks and those analysed in this publication at the same time.

Confocal counting experiments. For assessment of YFP-LRC position, YFP-LRC number and YFP-Paneth cell number, 100 crypts were counted for each time point in the proximal small intestine. Position +1 was defined as the apical two cells in the crypt and subsequent positions followed superiorly. Illuminating laser intensity was set uniformly between all comparisons to allow for reproducible and comparable detection of YFP positivity. Paneth cells were identified based on cytoplasmic and/or membranous UEA positivity. Countable crypts were initially detected by DAPI confirmation of a complete half crypt. The crypt of interest was then assessed for UEA or YFP positivity up to position +6. Label-retaining index was calculated as the mean number of YFP-LRCs at crypt positions +1 to +6 2 weeks after β NF induction.

EdU and YFP dual labelling experiments were performed by injecting mice with three injections of 1 mg per mouse (intraperitoneal) EdU over 24 h followed by three (intraperitoneal) β NF (80 mg kg⁻¹) injections over the next 24 h. Mice were killed 14 days later and EdU visualized using the Invitrogen Click-iT imaging detection kit as per the manufacturer's protocol.

The murine crypt has a population of 250 cells³⁰. Confocal assessment of an optimally sectioned 2 μ m optical section of a small intestinal crypt shows 39.3 ± 1.1 cellular nuclei. A single section through a crypt is therefore representative of 16% of the total number of crypt cells.

Images were acquired using a C1Si Nikon confocal microscope. Images were edited on EZ-C1 3.20 FreeViewer and Adobe Photoshop CS3. Volocity software was used to generate the H2B-YFP crypt three-dimensional reconstruction and video.

Flow cytometry. Intestinal epithelial single-cell preparation was performed as previously described³¹. Briefly, after dissection intestines were everted and fed onto either 3-mm (distal SI) or 4-mm (proximal SI) diameter glass rod spirals. They were then incubated in 50 ml of HBSS without Ca²⁺ and Mg²⁺ supplemented with 10 μ M EDTA and 10 mM NaOH at 37 °C for 1 h. Every 10 min the spirals were pulsed using a vibrating stirrer (Chemap AG, model CH-8604) and the HBSS collected and replaced with fresh HBSS. Fractions 2–6 were pooled and then re-suspended in $1 \times 0.05\%$ Trypsin-EDTA at 37 °C for 7 min with regular agitation. Epithelial samples were filtered through a 70 μ m mesh into cold 2% FBS/PBS and washed twice more.

After staining with antibodies and UEA, samples were sorted or analysed using either a BD FACS Aria SORP or BD LSR II, respectively. Single stained and unstained controls were always used, and propidium iodide (2.5 μ g ml⁻¹) was used to discriminate between live and dead cells.

For flow cytometric assessment of the percentage of YFP-LRCs, the denominator was defined as the number of single, live cells from analysis of 50,000 cells of pooled fractions 2–6 of a single cell epithelial preparation; that is, crypt enriched. The numerator was defined as the number of CD24⁺ UEA⁻ YFP⁺ cells; that is, non-Paneth-cell, crypt-located YFP-LRCs from the same population.

Single-cell fluorescent images were acquired using an Amnis ImageStream. FACS data were analysed using FlowJo v7.5.5 (TreeStar).

Affymetrix exon microarray and analysis. Six age- and sex-matched *Cyp1a1-H2B-YFP* mice 10 days after β NF induction were used comparing three cell populations from each mouse. After flow sorting 30,000 cells from the respective populations, RNA extraction was performed using the RNeasy Micro kit (Qiagen) with on column DNase digestion. RNA quality was assessed using a 2100 Bioanalyzer (Agilent Technologies).

RNA amplification and hybridization was performed at the Paterson Institute, Manchester, UK, Microarray Facility using Nugen Ovation Pico WTA System for amplification and then hybridization to a Mouse Exon 1.0 ST Array. Arrays were scanned using an Affymetrix GeneChip scanner 3000 running GCOS software.

The raw data were RMA (robust multi-array average)³² normalized using the Affymetrix Power Tools (APT) software package and the analysis was restricted to 'core' probe sets which are supported by Refseq³³ annotations. The data set was filtered on both the cross-hybridizing and undetected probe sets³⁴ and further processing was carried out in R³⁵ using a number of Bioconductor Packages³⁶. Differential expression analysis was carried out in *limma*³⁷ and the hypergeometric test was used to test for significant overlap with Lgr5⁺ cells. The Benjamini-Hochberg method for multiple correction was applied to *P* values from both methods. The data were deposited in the GEO database (accession number GSE43772).

Organoid culture. Culture of organoids was performed as previously described and including the use of Wnt3A (R&D) (100 ng ml⁻¹) for the first 3 days after plating²². Cells were plated in 24-well plates at a density of ~50,000 cells in 50 μ l matrigel. Organoids were counted as budding organoids at 10 days post seeding.

Single-cell expression profiling. For single-cell expression profiling, single cells were flow-sorted into 96-well PCR plates (Star Lab) with 4 μ l of first-strand synthesis buffer in each well containing 1 μ l 5 $\times$ SuperScript III buffer (Invitrogen), 0.5 μ l 5% NP-40 (Pierce), 0.25 μ l 1 mM dNTP mix (Thermo Scientific), 0.075 μ l 1 μ M MO4d(T) primer (DNA Technology A/S), 0.05 μ l 0.1 M DTT (Invitrogen), 0.25 μ l SuperRNaseIN (Ambion), 0.25 μ l RNase OUT (Invitrogen), 1 μ l Spike *Arabidopsis* DNase treated total RNA 10⁻¹ pg μ l⁻¹ (gift from A. Giakountis) and 0.625 μ l nuclease-free H₂O (Ambion); and then incubated at 65 °C for 5 min followed by 45 °C for 2 min. Then 0.5 μ l of SuperScript III (Invitrogen) was added to each aliquot and incubated at 45 °C for 15 min and then 65 °C for 10 min.

One microlitre of exonuclease mix was then added containing 0.2 μ l exonuclease I (Thermo Scientific), 0.6 μ l 75 mM MgCl₂ (Ambion) and 0.2 μ l nuclease-free H₂O (Ambion) and incubated at 37 °C for 30 min and 80 °C for 25 min.

Seven microlitres of a reaction mix was then added containing 2.6 μ l 5 $\times$ TdT buffer (Promega), 0.097 μ l 100 mM dATP (Thermo Scientific), 0.5 μ l RNase H (Invitrogen), 0.5 μ l TdT (Promega), 3.303 μ l nuclease-free H₂O (Ambion) and incubated at 37 °C for 15 min and 70 °C for 10 min.

Sixteen microlitres of amplification mix 1 was then added to 4 μ l of polyadenylated cDNA in triplicate, containing 2 μ l 10 $\times$ ExTaq buffer (Takara),

0.13 µl 100 mM dATP/dGTP/dCTP/dTTP (Thermo Scientific), 1.65 µl 100 µM MO₄(T) primer (DNA Technology A/S), 0.2 µl ExTaq polymerase (Takara) and 11.63 µl nuclease-free H₂O (Ambion) and cycled for 94 °C for 1 min, 50 °C for 2 min and 72 °C for 2 min followed by 35 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 2 min.

The triplicate products were then pooled and 2 µl of pooled product was added in duplicate to 18 µl of amplification mix 2 containing 2 µl 10× ExTaq buffer (Takara), 1.6 µl 2.5 mM dNTP mix (Takara), 0.4 µl 100 µM T7-MO4 primer (DNA Technology A/S), 0.2 µl ExTaq polymerase (Takara) and 13.8 µl of nuclease-free water (Ambion) and cycled for 35 cycles at 94 °C for 30 s, 60 °C for 30 s and 72 °C for 2 min. The duplicate samples were then pooled. Subsequent cDNA was then cleaned up using Zymo Clean and Concentrate kit and normalized based on Nanodrop quantification.

Normalized single-cell amplified cDNA was run on the Fluidigm Biomark EvaGreen qRT-PCR system against custom-designed primer pairs. Data were analysed using Fluidigm Real-Time PCR Analysis software (v3.0.2) and samples were normalized to spike *Arabidopsis* RNA primers LTP4 and TIM. Principal component analysis was carried using the prcomp function of the R statistical computing and graphics language. Cells were eliminated from PCA if the mouse housekeeping genes had failed to fire.

RNA amplification. For *Lgr5*/UEA qRT-PCR, cell numbers were limiting so a modified single-cell RNA amplification protocol was used³⁸. Briefly, 1 µl of extracted total RNA underwent first-strand synthesis in a buffer containing 1 µl 5× SuperScript III buffer (Invitrogen), 0.5 µl 5% NP-40 (Pierce), 0.25 µl 1 mM dNTP mix (Thermo Scientific), 0.075 µl 1 µM MO₄(T) primer (DNA Technology A/S), 0.05 µl 0.1 M DTT (Invitrogen), 0.25 µl SuperRNaseIN (Ambion), 0.25 µl RNase OUT (Invitrogen) and 0.625 µl nuclease-free H₂O (Ambion); and then incubated at 65 °C for 5 min followed by 45 °C for 2 min. 0.5 µl of SuperScript III (Invitrogen) was then added to each aliquot and incubated at 45 °C for 15 min and then 65 °C for 10 min.

One microlitre of exonuclease mix was then added containing 0.2 µl exonuclease I (Thermo Scientific), 0.6 µl 75 mM MgCl₂ (Ambion) and 0.2 µl nuclease-free H₂O (Ambion) and incubated at 37 °C for 30 min and 80 °C for 25 min.

Seven microlitres of a reaction mix was then added containing 2.6 µl 5× TdT buffer (Promega), 0.097 µl 100 mM dATP (Thermo Scientific), 0.5 µl RNase H (Invitrogen), 0.5 µl TdT (Promega), 3.303 µl nuclease-free H₂O (Ambion) and incubated at 37 °C for 15 min and 70 °C for 10 min.

Sixteen microlitres of amplification mix was then added to 4 µl of polyadenylated cDNA in triplicate, containing 2 µl 10× ExTaq buffer (Takara), 0.13 µl 100 mM dATP/dGTP/dCTP/dTTP (Thermo Scientific), 1.65 µl 100 µM MO₄(T) primer (DNA Technology A/S), 0.2 µl ExTaq polymerase (Takara) and 11.63 µl nuclease-free H₂O (Ambion) and cycled at 94 °C for 1 min, 50 °C for 2 min and 72 °C for 2 min followed by 35 cycles at 94 °C for 30 s, 60 °C for 30 s and 72 °C for 2 min.

Subsequent cDNA was then cleaned up using Zymo Clean and Concentrate kit and normalized based on Nanodrop quantification. qRT-PCR was then performed as follows using custom-designed 3' orientated primers.

qRT-PCR. Sybr green qRT-PCR was performed under standard conditions using a Rotorgene RG3000 (Corbett Research) and data analysed using Rotorgene 6 software. Custom primers were validated before use using standard Sybr green qRT-PCR and agarose gel electrophoresis of PCR products. Samples were normalized to housekeeping genes β-actin or β2 microglobulin.

TaqMan qRT-PCR was performed under standard conditions as above. Samples were normalized to housekeeping genes ribosomal protein L19 or glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*).

Antibodies and lectin. *Ulex europaeus* agglutinin I (UEA)-Alexa Fluor 647, 1:1,000 (custom conjugation; AbD Serotec). Anti-CD24-Pacific blue, 1:200 (Biolegend). Anti-substance P, 1:100 (Millipore). Anti-lysozyme, 1:1,000 (Dako). Anti-chromogranin A, 1:400 (Abcam). Anti-villin, 1:100 (Abcam).

Primers used for single cell-expression profiling. *Abca1* forward, 5'-TTGTTCCTCAAGAGCCATGTG-3'; *Abca1* reverse, 5'-GGAATGAGGGCC AATGATAA-3'; *Apc* forward, 5'-TGGGAGATGGTTGCCAGGGT-3'; *Apc* reverse, 5'-GTCGGAAGATGCATCAATGGCCT-3'; *Atll* forward, 5'-TGATCAAAAATGCACAAAATTG-3'; *Atll* reverse, 5'-AGAGCCAGTGA GCATTTGGT-3'; β-actin forward, 5'-AGGTGACAGCATTTGCTTCTG-3'; β-actin reverse, 5'-AGGGAGACCAAGCCTTCAT-3'; β2 microglobulin forward, 5'-ATTCACCCCCACTGAGACTG-3'; β2 microglobulin reverse, 5'-GCTATTCTTTCTGCGTGCAT-3'; *Bmi1* forward, 5'-TCATGGTGTGTACCT AAGACAAAAGAC-3'; *Bmi1* reverse, 5'-GGTAAAAAGCCTCATCCAGAG-3'; *Cadps* forward, 5'-TCACATTTTGTCTTTCTGTTTCA-3'; *Cadps* reverse, 5'-CCAGGGTCGCTCAGTTACAG-3'; *CD133* forward, 5'-ATCGGGGAAAC GAAGAAGTT-3'; *CD133* reverse, 5'-ACAGCCGGAAGTAAGAGCAC-3'; *CD44* forward, 5'-CCCCCTTTCTTTTCCAGTT-3'; *CD44* reverse, 5'-ACTTTC

TGCCCTCTCCACT-3'; *CD82* forward, 5'-CAGGGTAGGCAATTCTTCCA-3'; *CD82* reverse, 5'-AGCTGCCAAGAAACACCAGT-3'; *Cdk5rap2* forward, 5'-TGCCAAGATGGATGTTCAAA-3'; *Cdk5rap2* reverse, 5'-AAGAGCTTCA GCAACCTGGA-3'; *Chga* forward, 5'-GCAGAGGACCAGGAGCTAGA-3'; *Chga* reverse, 5'-CAGGGGCTGAGAACAAGAGA-3'; *Chgb* forward, 5'-ACAGGAAGAAGGCAGACGAA-3'; *Chgb* reverse, 5'-TCCTTCAGTGAAG GGTCTGT-3'; *Csrnp1* forward, 5'-GCCCCCTCTGTAAGATGGTGA-3'; *Csrnp1* reverse, 5'-GGCACACACACACACACAAA-3'; *DCAMKL-1* forward, 5'-AGGTGGGCTGGGGACTTGACA-3'; *DCAMKL-1* reverse, 5'-TGCAGCAAG TGACAAGGGCA-3'; *Ddahl1* forward, 5'-CCAAACACTAAGGCCGTCAT-3'; *Ddahl1* reverse, 5'-CACACATTGGCTGGAAGATG-3'; *Defa5* forward, 5'-TCTGTCTCAACAATTCTCCA-3'; *Defa5* reverse, 5'-GACACAGCCTGGT CCTCTTC-3'; *Dll1* forward, 5'-TGAGCCAGTCTTCTCTTGAA-3'; *Dll1* reverse, 5'-AGACCCGAAGTGCCTTTGTA-3'; *Dll4* forward, 5'-CCCCTTTTGCCAGC CAGGGG-3'; *Dll4* reverse, 5'-AGGTTGGTCTCGGGGCGAGCA-3'; *Dvl3* forward, 5'-ACACCTTAAGCCACCCCTTT-3'; *Dvl3* reverse, 5'-GCTCTCTG GGGTCTGATTG-3'; *EphB2* forward, 5'-TCCTCCCTGTCTGCTCTCAT-3'; *EphB2* reverse, 5'-CTGGGGTGTACAGTGAATG-3'; *EphB4* forward, 5'-TCC TGACTTGTCTGTCACCAC-3'; *EphB4* reverse, 5'-GAACCAGGTGCCCTTT AACA-3'; *Fgfr2* forward, 5'-TCGAAGGATGGCAAAGCCAGC-3'; *Fgfr2* reverse, 5'-AGTGAGTGGGCGTATCCAAAGCA-3'; *Frs2* forward, 5'-TCTACAGTGGGACTACCATTGCCA-3'; *Frs2* reverse, 5'-ACAGTGGTTACG TTTGCCAAGCA-3'; *Gip* forward, 5'-GGGAAAGGAGGACAAAGAGG-3'; *Gip* reverse, 5'-AGCCAAGCAAGCTAAGGTCA-3'; *Kif20a* forward, 5'-ACAG GAACAGCAGTGGGAAC-3'; *Kif20a* reverse, 5'-CACTAACTGGGTGCCA TGTG-3'; *Lgr5* forward, 5'-GGGCGTTAAGTCCACTGTGT-3'; *Lgr5* reverse, 5'-CGAACACCTGCGTGAATATG-3'; *Lrig1* forward, 5'-TTCCTTACCCGTG AGACTGG-3'; *Lrig1* reverse, 5'-CCATCACTGTGCCAACACTT-3'; *LTP4* forward, 5'-GTGCTGTGTCAGGAGTCAAAA-3'; *LTP4* reverse, 5'-TCTTCA GGCAAATGATGTCG-3'; *Mcm6* forward, 5'-CTGAAAGGCTCCAGTGAA GG-3'; *Mcm6* reverse, 5'-TCAAGCATTGCTCCAGACAC-3'; *Mctp2* forward, 5'-CATCTGGAGCGTAACCCCTA-3'; *Mctp2* reverse, 5'-CAGTCAAAAAG CAGGTGGA-3'; *Mmp7* forward, 5'-CCCGGTACTGTGATGTATCC-3'; *Mmp7* reverse, 5'-AATGGAGAGCCAGTGAGTG-3'; *Tert* forward, 5'-AGCC GCATTTGGCTCTGCT-3'; *Tert* reverse, 5'-TCGCTCTGGGAGCTTCCGG-3'; *Muc2* forward, 5'-GTCCGAAGTGTATACCCTGGA-3'; *Muc2* reverse, 5'-CCAGGAGTGAGAGGTCAG-3'; *Nfat5* forward, 5'-ACTGACCTGCCTT CTGTCAT-3'; *Nfat5* reverse, 5'-TTGACTGTCAGCTGTTTACAGAAA-3'; *Nfatc1* forward, 5'-TCATTTGCTCTGCACCAGTC-3'; *Nfatc1* reverse, 5'-GAC GAGAGTCAGGGAAGCAG-3'; *Nfatc3* forward, 5'-TGGGCTCAAGATGA AACCTT-3'; *Nfatc3* reverse, 5'-TAACTGAGGAGGAGCCTGGA-3'; *Ngn3* forward, 5'-ATCTGCCGCGCTCCGACCAT-3'; *Ngn3* reverse, 5'-GCCTGG TCTCCCTTGGGGGA-3'; *Notch1* forward, 5'-TCAGTGTGACCAGACC TTG-3'; *Notch1* reverse, 5'-CAAAAGGCCAGAAAGAGCTG-3'; *Pax6* forward, 5'-GAACAACACAGGCTGTTGGA-3'; *Pax6* reverse, 5'-TGTGTGTGTGCC CAGGTTCA-3'; *Pbx1* forward, 5'-AGAAAGCCATGTGTTGGAC-3'; *Pbx1* reverse, 5'-TGAAAAGGGTGCCTTTCATC-3'; *Rfx6* forward, 5'-TTTTCTT CCTCTTTGGTTTC-3'; *Rfx6* reverse, 5'-GCACAGGGGTAGAAGGTCAA-3'; *Ssh2* forward, 5'-CCACAGCCCAAGTATTC-3'; *Ssh2* reverse, 5'-GGAAGGCCTCCTTCAGAACT-3'; *Tac1* forward, 5'-GATGAAGGAGCT GTCCAAGC-3'; *Tac1* reverse, 5'-TCACGAAACAGGAAACATGC-3'; *TIM* forward, 5'-ACGGAAATCGCTAGTCTCCA-3'; *TIM* reverse, 5'-TGTCAAAC CAGAGCTCACA-3'; *Tph1* forward, 5'-AATTTGCCAAGACCGTGAAG-3'; *Tph1* reverse, 5'-GCCCTGGCTCTAGACTGATG-3'; villin forward, 5'-ATT AGCGTCTGGGGGTTTCT-3'; villin reverse, 5'-CGTAGCAAACCCATGTT CCT-3'; *Zbtb48* forward, 5'-CACATGGAGATCCACGACAG-3'; *Zbtb48* reverse, 5'-CAGGAGCTCCACAATGACCT-3'. MO4(T) primer, 5'-AAGCA GTGGTATCAACGACAGATGGCCATTACGGCCGTACTTTTTTTTTTTTTTTT- TTTTTTTTTTTTTTTTTT-3'; T7-MO4 primer, 5'-GGCCAGTGAATTGTAA TACGACTACTATAGGGAGGCGGAAGCAGTGGTATCAACGCAGAGTG GCCATTACGGCCGTAC-3'.

Taqman probes. *Atoh1*, Mm00476035_s1; *Chgb*, Mm00483287_m1; *Chga*, Mm00514341_m1; *Clca3*, Mm01320697_m1; *Kit*, Mm03053277_s1; *Gapdh*, Mm03302249_g1; *Hoxp*, Mm00558630_m1; *Lgr5*, Mm00438890_m1; *Lrig1*, Mm00456116_m1; *Lyz1*, Mm00657323_m1; *Mmp7*, Mm00487724_m1; *Muc2*, Mm00458299_m1; *Rpl19*, Mm01606039_g1; *Tert*, Mm00436931_m1; villin, Mm00457074_m1.

26. Campbell, S. J. *et al.* Regulation of the CYP1A1 promoter in transgenic mice: an exquisitely sensitive on-off system for cell specific gene regulation. *J. Cell Sci.* **109**, 2619–2625 (1996).

27. Vooijs, M., Jonkers, J. & Berns, A. A highly efficient ligand-regulated Cre recombinase mouse line shows that LoxP recombination is position dependent. *EMBO Rep.* **2**, 292–297 (2001).

28. McCutcheon, S. C. *et al.* Characterization of a heat resistant β -glucosidase as a new reporter in cells and mice. *BMC Biol.* **8**, 89 (2010).
29. Gracz, A. D., Ramalingam, S. & Magness, S. T. Sox9 expression marks a subset of CD24-expressing small intestine epithelial stem cells that form organoids in vitro. *Am. J. Physiol. Gastrointest. Liver Physiol.* **298**, G590–G600 (2010).
30. Potten, C. S. & Loeffler, M. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* **110**, 1001–1020 (1990).
31. Bjerknes, M. & Cheng, H. Methods for the isolation of intact epithelium from the mouse intestine. *Anat. Rec.* **199**, 565–574 (1981).
32. Irizarry, R. A. *et al.* Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31**, e15 (2003).
33. Pruitt, K. D., Tatusova, T., Klimke, W. & Maglott, D. R. NCBI Reference Sequences: current status, policy and new initiatives. *Nucleic Acids Res.* **37**, D32–D36 (2009).
34. Lockstone, H. E. Exon array data analysis using Affymetrix power tools and R statistical software. *Brief. Bioinform.* **12**, 634–644 (2011).
35. R Development Core Team. *R: A Language and Environment for Statistical Computing* (2009).
36. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
37. Smyth, G. K. in *Bioinformatics and Computational Biology Solutions using R and Bioconductor Solutions using R and Bioconductor* (eds Carey, V., Gentleman, R., Dudoit, S., Huber, W. & Irizarry, R.) 397–420 (Springer, 2005).
38. Jensen, K. B. & Watt, F. M. Single-cell expression profiling of human epidermal stem and transit-amplifying cells: Lrig1 is a regulator of stem cell quiescence. *Proc. Natl Acad. Sci. USA* **103**, 11958–11963 (2006).