Supplementary Methods

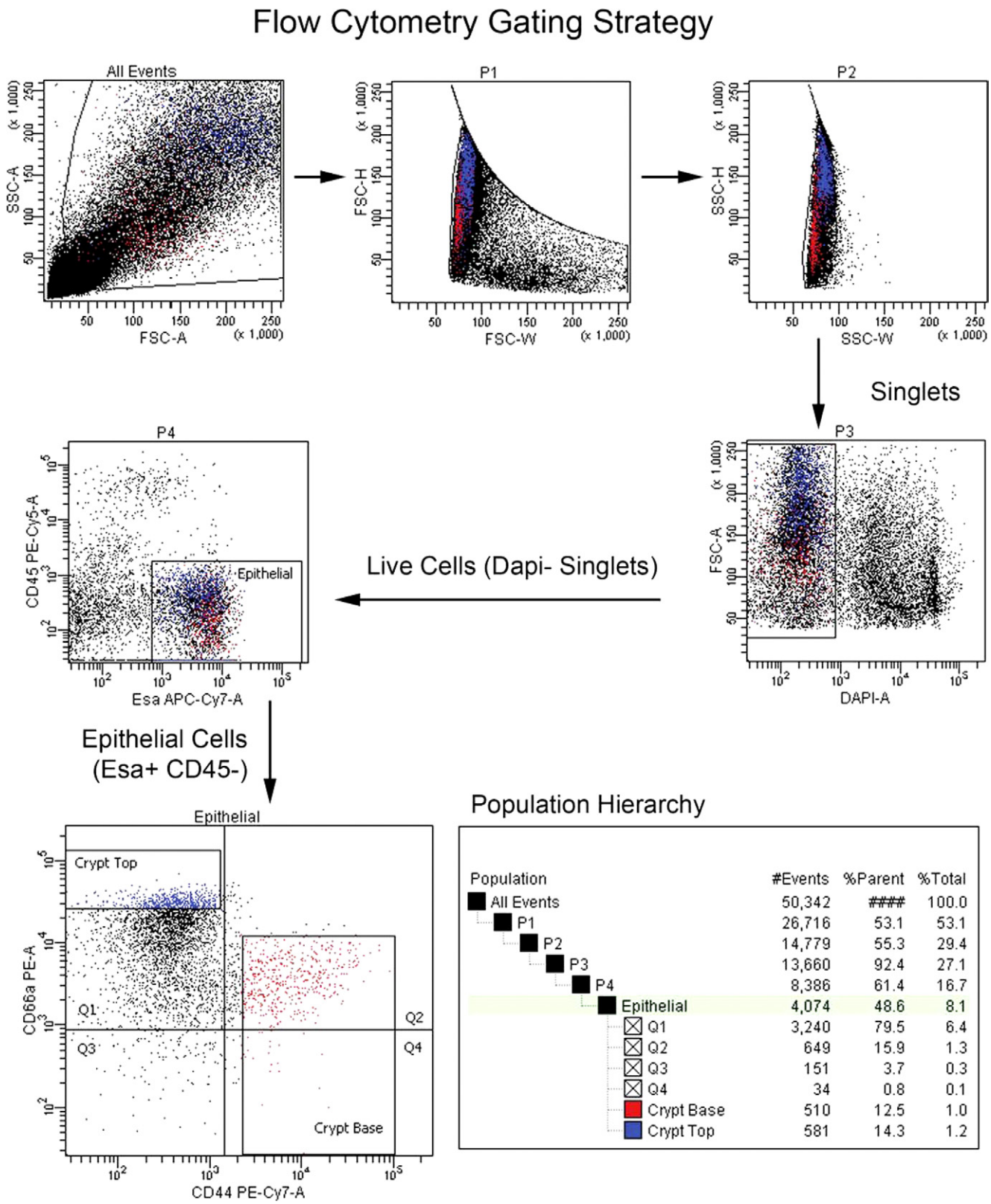
Flow Cytometry

For each flow cytometry experiment, multiple stainings with different antibody combinations were done to ensure that antibody combinations did not cause gating artifacts. All flow Cytometry was performed with a 130 uM nozzle on a BD FacsAria II using FacsDiva software. Optimal voltages for each fluorophore were determined prior to each experiment, and the CST was checked daily. Compensation was performed before each experiment using dissociated colon cells stained with the individual conjugated antibodies used in the experiment. The compensation matrix was calculated with the FacsDiva software, and was then checked manually and adjusted as necessary using single color controls. Multicolor stains were also checked with FMO (fluorescence minus one) controls. Debris and doublets were excluded by sequential gating on FSC-area vs SSC-area, followed by FSC-width vs FSC-height, followed by SSC-height vs SSC-width (Fig S1). Viable cells were identified by exclusion of DAPI (Molecular Probes) and were typically 60–70% of live singlets.

Single Cell Gene Expression

mRNA for a predefined set of genes was reverse transcribed and pre-amplified for 20 cycles using multiplex PCR and thermocycling protocols as described[17]. The resulting preamplified cDNA from each cell was then diluted 1:3 in TE buffer. A 2.25 uL aliquot was mixed with 2.5 uL of 2x Taqman Universal Master Mix (Applied Biosystems) and 0.25 uL of Fluidigm sample loading agent, and was loaded into the sample inlets of a Dynamic Array M96 Fluidigm microfluidic chip using a Hamilton StarLET pipetting robot. Individual Taqman assays (Table S1) were diluted 1:1 with TE buffer, and a 2.5 μl aliquot of each diluted assay was then mixed with 2.5 μl of Fluidigm assay loading agent and individually loaded into the M96 microfluidic chip assay inlets. Loaded microfluidic chips were then transferred to a BioMark real-time PCR reader (Fluidigm), and loaded chips underwent thermocycling and fluorescent quantification according to manufacturer instructions. Each experiment consisted of at least 4 M96 chips, i.e., 336 sorted cells. 12 samples on each 96 well plate (the bottom row) were used as a positive control for Taqman assays by creating a serial dilution of control mRNA, which was a 1:1 mixture of total mouse intestine mRNA and mouse colon total RNA (Biochain).

For analysis, we removed low quality assays-those in which the qPCR amplification curves showed non-exponential increases. We also removed low quality cells (<10% total) that did not express housekeeping genes (Actb or GAPDH) or ESA/EpCAM. For each gene we normalized by mean-centering and dividing by three times the standard deviation of the qPCR threshold cycles for all expressing cells. Hierarchical clustering (Matlab command clustergram) was performed on both cells and genes, with a Euclidean or correlation distance metric and complete linkage. Positive or negative associations among pairs of genes were tested by Spearman correlation, and p-values were calculated using n = 10,000 permutations.



Supplementary Figure 1. *FACS gating strategy for isolation of colonic crypt subregions.*
In the first 3 plots (top row), we exclude debris and doublets by sequentially gating on FSC-A vs. SSC-A, then FSC-W vs. FSC-H, and then SSC-H vs. SSC-W. Live single cells (box in Dapi vs FSC-A plot) are then plotted for Esa vs. CD45, which shows three major populations: hematopoietic Esa-CD45+ cells, stromal Esa-CD45− cells, and epithelial Esa+CD45− cells (box). Epithelial cells represent approximately half (45–60%) of all live single cells. These are next plotted on a CD44 vs. CD66a plot to separate crypt top (blue) from crypt bottom (red) cells. A population hierarchy is shown (lower right). Gating was done with FacsDiva software.

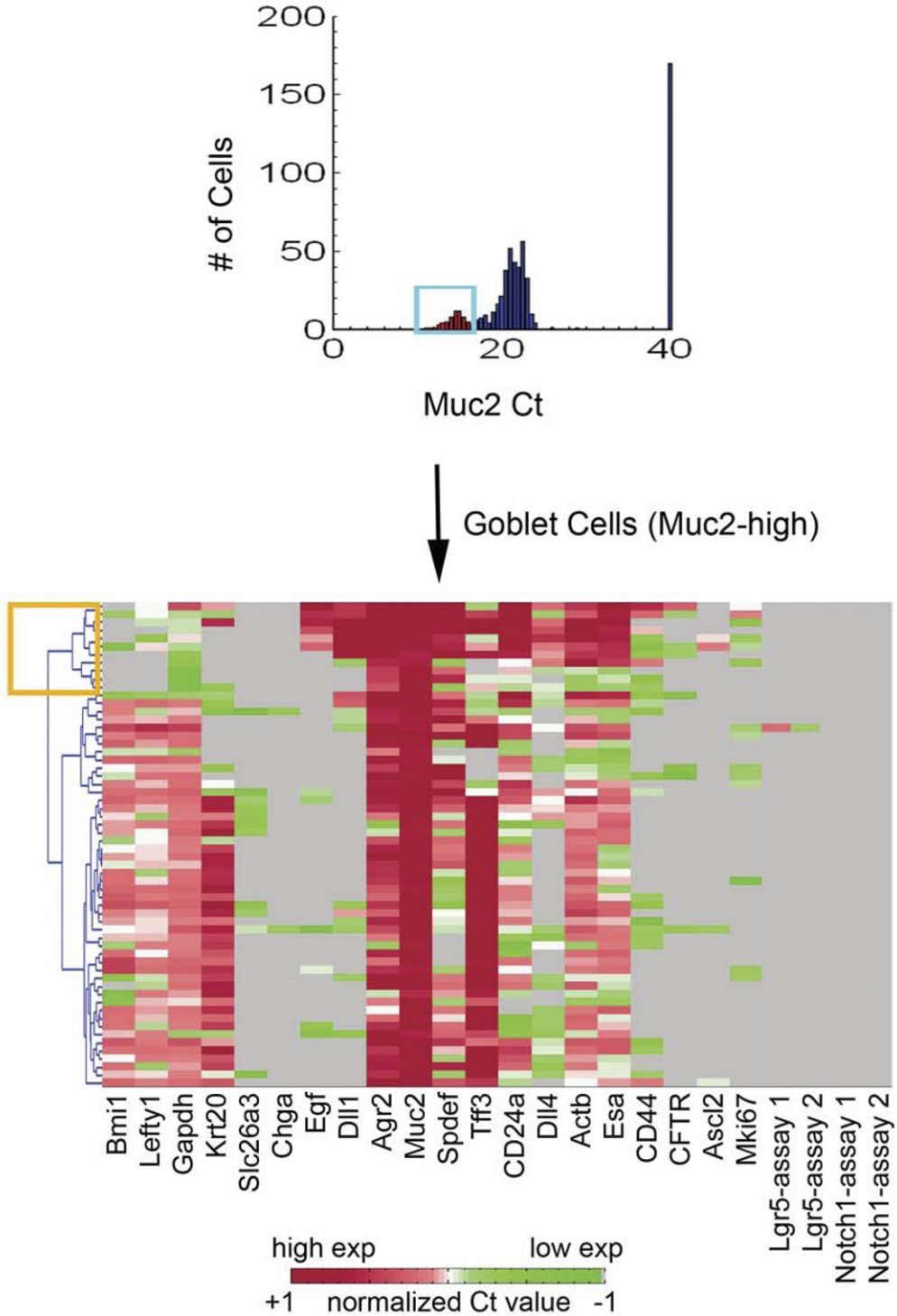
Supplementary Table 1. Taqman Assays used for Single Cell and Conventional qPCR Gene Expression Analysis

Assay Number	Gene
Mm00607939_s1	Actb
Mm00507853_m1	Agr2
Mm01268891_g1	Ascl2
Mm00443610_m1	Axin2
Mm03053308_g1	Bmi1
Mm00782538_sH	Cd24a
Mm01277160_m1	Cd44
Mm00445197_m1	Cftr
Mm00432841_m1	Dll1
Mm01279269_m1	Dll1
Mm00444619_m1	Dll4
Mm01338018_m1	Dll4
Mm01316968_m1	Egf
Mm99999915_g1	Gapdh
Mm00468601_m1	Hes1
Mm00445212_m1	Kit
Mm00442972_m1	Kitl
Mm00508106_m1	Krt20
Mm03053915_s1	Lefty1
Mm00438890_m1	Lgr5
Mm01251805_m1	Lgr5
Mm01278616_m1	Mki67
Mm00458299_m1	Muc2
Mm00435245_m1	Notch1
Mm00471115_m1	Reg4
Mm00445313_m1	Slc26a3
Mm00600221_m1	Spdef
Mm00803437_m1	Spink4
Mm01227384_m1	Tacstd1 (EpCam, ESA)
Mm00436930_g1	Tert
Mm00436931_m1	Tert
Mm00495590_m1	Tff3
Mm00495732_g1	Top2a

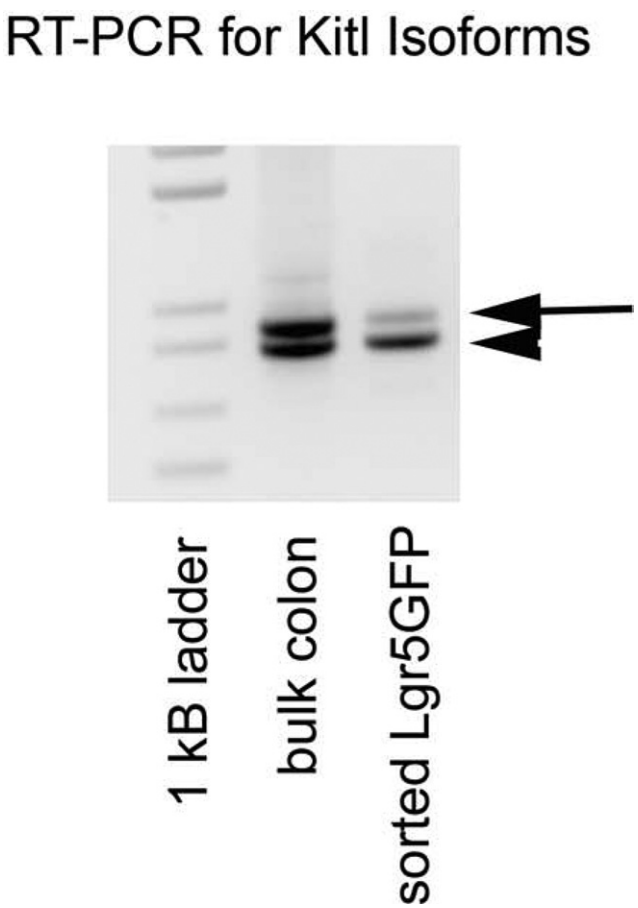
Supplementary Table 2. Antibodies used in immunostaining

Antibody	Dilution	Vendor	Catalog/Clone Num
Goat anti-cKit	1:200	R&D	AF1356
Rabbit anti-Muc2	1:50	Santa Cruz Biotechnology	H300
Mouse anti-human Ki67	1:5	BD	clone B56
Goat anti-human Reg4	1:200	R&D	AF1379
Rabbit anti-human lysozyme	1:1000	Dako	EC 3.2.1.17
ESA-Alexa488	1:200	Biolegend	clone G8.8
CD45-biotin	1:200	Biolegend	clone 30F-11
CD44-biotin	1:100	Biolegend	clone IM7
CD66a-PE	1:200	Biolegend	clone MAb-CC1
Donkey anti-goat Alexa594	1:400	Jackson Labs	705-515-147
Donkey anti-rabbit Alexa488	1:400	Jackson Labs	711-545-152
Donkey anti-rabbit Alexa594	1:400	Jackson Labs	711-585-152
Streptavidin-Alexa594	1:300	Invitrogen	S-11227

Supplementary Figure 2. *MKi67 staining of adult mouse colon.*
Immunostaining for MKi67 (green) and Dapi (blue) shows that proliferating cells are confined to the lower half of the crypts, consistent with published work and with our single cell profiling in Figure 2. *Scale bar:* 50 uM.



Supplementary Figure 3. *Single cell transcriptional profiling of Goblet Cells*.
In this experiment, crypt base epithelial cells were analyzed by single cell gene expression analysis as described. A histogram of Muc2 expression (top) shows 3 populations: Muc2 non-expressing cells (Ct = 40), Muc2 low cells (dark blue peak), and Muc2 high cells, i.e., goblet cells (red peak, enclosed in light blue box). The Ct cutoff for Muc2 high cells was 16.5. Nearly all cells express high levels of Agr2, which is required for Muc2 production. Hierarchical clustering shows a subpopulation of EGF+Dll1+ goblet cells (yellow box), as seen in Figure 2. They also express high levels of Dll4, Esa, CD24, and Spdef. cKit was not included in this experiment.



Supplementary Figure 4. *Expression of secreted and transmembrane cKit isoforms in Lgr5+ colon cells.*
RT-PCR on total mouse colon (lane 2) and FACS-sorted Lgr5-GFP+ cells (lane 3) for membrane-bound (arrow, ~910 bp) and secreted (arrowhead, ~830bp) cKit isoforms shows that both are detected. Lane 1 is 1 kB DNA ladder.