SUPPLEMENTARY MATERIALS

Notch signaling mediates differentiation in Barrett's Esophagus and progression to adenocarcinoma

*Bettina Kunze[1], *Frederik Wein[1], Hsin-Yu Fang[1], Akanksha Anand[1], Theresa Baumeister[1], Julia Knörndel[1], Sophie Gerland[1], Jonas Ingermann[1], Natasha Stephens Münch[1], Maria Wiethaler[1], Vincenz Sahm[1], Ana Hidalgo Sastre[1], Sebastian Lange[1], Charles J Lightdale[2], Aqiba Bokhari[3], Gary W Falk[4], Richard A Friedman[5,6], Gregory G Ginsberg[4], Prasad G Iyer[7], Zhezhen Jin[8], Hiroshi Nakagawa[2,6], Carrie J Shawber[9], TheAnh Nguyen[10], William J. Raab[11], Piero Dalerba[2,6,11,12], Anil K Rustgi[2,6], Antonia R Sepulveda[11], Kenneth K Wang[7], Roland M Schmid[1] Timothy C. Wang[2,6], #Julian A. Abrams[2,6], #Michael Quante[1]

[1]II. Medizinische Klinik, Technische Universitat München, Munich, Germany
[2]Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA
[3]Yosemite Pathology Medical Group, Modesto, CA, USA
[4]Department of Medicine, Division of Gastroenterology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA
[5]Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY, USA
[6]Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY, USA
[7]Department of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN, USA
[8]Department of Biostatistics, Columbia University Mailman School of Public Health, New York, NY USA
[9]Department of Obstetrics and Gynecology, Columbia University Irving Medical Center, New York, NY, USA
[10]Oregon Health and Science University, Portland, OR, USA
[11]Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, NY, USA
[12]Columbia Stem Cell Initiative, Columbia University Irving Medical Center, New York, NY, USA

* These authors contributed equally
# Co-corresponding authors

SUPPLEMENTARY METHODS

Institutional Review Board Approval

The prosptective, multi-center cross-sectional study was approved by the Institutional Review Boards at Columbia University Medical Center (IRB protocol AAAI6700), the University of Pennsylvania (IRB protocol 814641), and Mayo Clinic (Rochester) (IRB protocol 11-006510).

Histologic classification of BE subjects

As the presence of dysplasia can vary within an individual over consecutive endoscopies, and there is significant inter-observer variability in the reading of dysplasia, we developed an algorithm to rigorously classify subjects based on their "worst" confirmed histology ever based on the following: Biopsies taken during the study endoscopy were interpreted by each study site's expert pathologist. These biopsies were also interpreted by a central pathologist (ARS), blinded to the local pathology read. If there was disagreement between the two reads, then a consensus diagnosis was reached after review by a third pathologist. If the subject had a history of prior dysplasia or EAC that had been confirmed by an expert GI pathologist, then these data were recorded. Each subject was then categorized based on comparison of the highest grade of neoplasia (dysplasia or cancer) from the study endoscopy and the worst from prior endoscopies.

Goblet cell density assessment

All biopsies obtained from the Barrett's segment for each subject were stained with hematoxylin and eosin and assessed for goblet cell density. The slides were then digitally scanned using a Leica scanner, and 200 μm x 200 μm grids were placed over the scanned slide image. (*Supplementary Figure 1A and B*) All areas of biopsies containing columnar-lined epithelium were evaluated at 10x magnification, excluding crushed tissue, oxyntic mucosa, and pancreatic metaplasia. As dysplasia is known to harbour reduced numbers of goblet cells, we only evaluated grids that contained at least 50% columnar lined epithelium and that contained exclusively non-dysplastic epithelium.

Each evaluable grid was assessed for the presence of goblet cells, using a semi-quantitative scoring method. The following scores were assigned to each evaluable grid: 0 (no goblet cells in the grid), 1 (≤ 3 goblet cells in the grid), or 2 (> 3 goblet cells in the grid). The primary outcome for goblet cell density was calculated for each subject as the number of grids with any goblet cells (score of 1 or 2) divided by the total number of evaluable grids. A secondary outcome for goblet cell density was calculated as the number of grids with high numbers of goblet cells (score of 2) divided by the total number of evaluable grids.

Gene expression analyses of human samples

Frozen esophageal biopsies were thawed, carefully removed from the Qiagen AllProtect® preservative and individually transferred into clean 1.5 mL polypropylene microcentrifuge snap-cap tubes (Eppendorf), where they were resuspended into 100 µL of TRIzol® Reagent (Ambion) and immediately snap-frozen again in liquid nitrogen. Each biopsy was then mechanically pulverized on dry ice, together with the solid 100 µL pellet of TRIzol® Reagent into which it had been resuspended, using a RNase-free disposable micro-pestle (Kimble Chase Life Science).[15] The pulverized material was subsequently dissolved by mixture with additional 400 µL of TRIzol® Reagent, which brought the final volume of the solution to 500 µL. The RNA was then isolated using a classic *"one-step"* acid guanidinium thiocyanate-phenol-chloroform extraction protocol followed by isopropanol precipitation,[16] and finally purified using RNeasy® spin-columns with silica-based membranes (Qiagen). For each sample, RNA concentration and integrity were measured using a BioAnalyzer (Agilent). 92% of the samples yielded RNA with a RNA Integrity Number (RIN) > 5 and 81% with a RIN > 7, two commonly used metrics to define, respectively, good and high quality RNA preparations, suitable for qRT-PCR applications.[17, 18]

Immunohistochemistry analyses for NOTCH1, NOTCH3, and Ki67 were performed on Barrett's esophagus biopsies from a subset (n=36) of the subjects. The epithelial and stromal

compartments were analyzed and scored separately in a semi-quantitative fashion (range 0-4). Further details are provided in the *Supplemental Methods* section.

Gene expression analyses of human samples

Samples were reverse transcribed according to the manufacturer's instructions for the High Capacity Reverse Transcription kit (Applied Biosystems, Foster City, CA). Briefly, 50 ng of total RNA was reverse transcribed in a 20 µl reaction mixture containing 0.8 µl of 100nM dNTP, 2.0 µl RT buffer, 1.0 µl of reverse transcriptase (50U/µl), 2 µl of RT primer. The reaction mixture was mixed and incubated as follows: 25ºC for 10 min, 37ºC for 2 h, and then 85ºC for 5 sec, followed by a 4°C hold.

Pre-amplification of cDNA was initiated by creating a pool of 96 TaqMan mRNA assays at a final concentration of 0.2X for each assay. The pre-PCR amplification reaction was performed in a 5 µl reaction mixture containing 2.5 µl TaqMan PreAmp Master Mix (2X), 1.25 µl of 24-pooled TaqMan assay mix (0.2X) and 1.25 µl of cDNA. The pre-amplification PCR was performed according to the following cycling conditions: one cycle at 95ºC for 10 min, 14 cycles of 95ºC for 15 seconds and 60ºC for 4 minutes, followed by a 4°C hold. After pre-amplification PCR, the product was diluted 1:5 with dH_2O and stored at -20° C until needed for final amplification.

Quantitative PCR of the mRNA targets was carried out using the 96.96 dynamic array IFC (Fluidigm, South San Francisco, CA) following the manufacturer's protocol. Briefly, a 5 µl sample mixture was prepared for each sample and contained 2x TaqMan Universal Master Mix (with UNG), 20X GE Sample Loading Reagent and each diluted pre-amplified cDNA. Five microliters of Assay mix was prepared with one 20X TaqMan mRNA assay (final concentration 10x) and 2X Assay Loading Reagent. The dynamic array was primed with control line fluid in the IFC controller and assay and sample mixes were loaded into the appropriate inlets. The chip was then returned to the IFC controller for loading and mixing, and then placed in the

BioMark Instrument for the final amplification step. The standard protocol for a 96x96 IFC consists of a thermal mix of 50ºC for 2 min, 70ºC for 30 min and 25ºC for 10 min; then a series of amplification steps: 50ºC for 2 min and 95ºC for 10 min, followed by 40 cycles at 95ºC for 15 sec and 60ºC for 1 min. The data was analyzed with Real-Time PCR Analysis Software in the BioMark instrument (Fluidigm, South San Francisco, CA).

The gene expression of 96 transcripts representing 80 genes in 192 samples was measured by PCR on two 96 samples plates with a Fluidigm Biomark M96 high-throughput PCR instrument.[1-3] Numbers of samples of each kind were closely balanced between plates. Samples with a significant portion of genes undetected were discarded, as were samples with RIN scores <5.[4] Samples were also discarded which were outliers from the whole dataset, based upon principal component analysis performed using the R affycoretools package[5], and hierarchical clustering with the R heatmap function using Euclidean distances and complete linkage clustering.[6] The expression of *EPCAM* vs the RIN score was plotted for the remaining samples as an additional check, and there were no significant outliers. Similarly, the expression of the house housekeeping genes *ACTB*, *GAPDH*, and *POLRA* were plotted against the RIN score and against each other, and there were no significant outliers. Gene expression was normalized against the average of the *ACTB*, *GAPDH*, and *POLRA* housekeeping genes. Differential expression was analyzed using Limma[7] blocked by plate and corrected for multiple testing using the Benjamini-Hochberg false discovery rate.[8] The primary comparison to identify differentially expressed genes associated with neoplastic progression was for samples from patients with HGD or EAC vs. those with non-dysplastic BE. The linear dependence and Spearman and Pearson Correlation coefficients of select genes relevant to Notch signaling from BE samples were estimated. Univariate dependencies were estimated and plotted using standard R functions.[9] Multivariate dependencies were estimated using standard R functions and plotted with the car R package.[10] A list of pairs of genes studied in this manner is given in *Supplementary Table 2*. Correlograms were generated to visualize correlations graphically. A correlogram is a heatmap displaying correlations between

variables.[11] Correlograms differ from the usual heatmaps in microarray analysis, in that whereas the usual heatmaps depict quantity of expression, correlograms depict extent of correlation. A correlogram displaying gene expression was generated by calculating the Pearson correlation coefficient in the R stats package[12], and then calculating Euclidean distances[13], clustering the genes by complete linkage clustering[13] and displaying the heatmap with the R gplots package.[14]

Immunohistochemistry of Human Samples

Immunohistochemistry analyses were performed for Notch1, Notch3, and Ki67 on Barrett's esophagus biopsies from a subset (n=36) of subjects. Standard immunohistochemical procedures with 1X100ml Target retrieval solution (DAKO, S1699) were performed using the following antibodies: Notch1 antibody (cell signaling #3608), 1:50 incubated 1.5 h at RT (antibody diluent using DAKO company Catalog Numer:S0809), Notch3 antibody (ab23426) 1:1000 incubated 1.5 h at RT. (antibody diluent using cell signaling Catalog Numer:8112) and CONFIRM anti-Ki-67 (30-9) Rabbit Monoclonal Primary Antibody. Quantification was accessed as percentage of positive cells or areas within BE regions as previously described.[15]

Lgr5 In Situ Hybridization (ISH)

Detection of Lgr5 within murine tissue was performed using the Advanced Cell Diagnostics RNAscope 2.5HD Assay- BROWN kit with the Mm-Lgr5 probe (ACD, USA) on FFPE slides following the manufacturer's protocol.

Flow cytometry

Single-cell suspensions of murine esophageal, cardia tissue, forestomach and colon regions were generated by chopping tissue with scissors in EDTA solution. Then the tissue- and EDTA solution was transferred into digestion medium followed by incubation at 150 rpm at 37°C for 30 minutes. Digestion medium for esophagueal tissue consisted of 5ml Krebs Ringer buffer +4% (w/v) BSA (0.2g)+2mg/ml collagenase P (Roche). Digestion buffer for cardia and

forestomach consisted of 5ml DMEM + 2 mg/ml collagenase P +2mg/ml Pronase (Roche). Blood and spleen samples were used as controls. The following antibodies (eBioscience) were used for T-cell staining: purchased from eBioscience eFluor450-labelled anti-CD45 (48045182), APC-eFluor780-labelled anti-CD11b (47011282), Alexa-eFlour700-labelled anti-Ly-6G (56593182), FITC -labelled anti- CD11c (11011481), PE-labelled anti-Ly-6C (12593282), anti-F4/80 (14480181) for myeloid staining and e450-labelled anti-CD4 (48004180), APC-labelled anti-CD8a (17008182), FITC-labelled anti-CD3 (11003381), APC-eFluor780-labelled anti-NK1.1 (47594182), PE-labelled anti- γδTCR (12571182).

To measure MUC/Muc2 levels in 3D cultured human or mouse derived organoids, cells were singularized by enzymatic digestion with trypsin (0.25% Trypsin-EDTA solution, Merck) and fixed with PFA (eBioscience 88-8824-00) prior treatment with APC-labelled anti-MUC2 antibody from Novus Biologicals. Frequencies of immune cells, isolated from relevant organs, were measured to evaluate potential shifts of the immune response due to genetic modifications.

Real-time PCR analysis

Subsequent to RNA isolation the QuantiTect Whole Transcriptome Kit (100) (207045, Qiagen) was used to enhance cDNA quantity. Quantitative PCR was performed using the LightCycler® 480 Instrument (Roche) and the QuantiFast SYBR Green PCR Kit (4000) (204057, Qiagen) according to the manufacturer's instructions. RNA levels were normalized to *GAPDH* levels. Primer sequences were designed with the NCBI PrimerBLAST tool and are listed in *Supplementary Table 3*.

DNA isolation and low coverage genome sequencing

FFPE tissue was microdissected and approximately 250 ng DNA was isolated using the Maxwell 16 LEV Blood Kit with 1-Thioglycerol and Incubation Buffer for deparaffinization. Library preparation was performed with an average of 50 ng DNA per sample using the

NEBNext Ultra II FS DNA Library Prep Kit for Illumina. Samples were sequenced on an Illumina NextSeq system, resulting in ~20 Mio. single-end, 75 bp long, reads per sample.

Resulting sequencing data was processed using a standardised set of pipelines.[16] Briefly, reads were trimmed using Trimmomatic and mapped to the mouse reference genome GRCm38.p6 using bwa mem. The GATK toolkit was used for base recalibration. Copy number alterations were called by HMMCopy, using data from the tail of backcrossed C57BL/6J mice as control.

Immunohistochemistry

Standard immunohistochemical procedures with citrate buffer antigen retrieval (1.00244.1000, Merck) were performed using the following antibodies: primary, rabbit anti-mouse Ki67 antibody (Abcam, ab15580, 1:500, 1h at room temperature) and secondary, goat anti-rabbit antibody (Vector Labs, BA-1000, 1:200, 30 min at RT); primary rabbit anti-human pIKKα/β antibody (Cell Signaling, 2697P, 1:75, overnight at 4°C); primary rabbit anti-human Notch2-IC antibody[17] (DSHB Hybridoma Bank, C651.6DbHN-c, 1:500; overnight at 4°C) and secondary, rabbit anti-rabbit antibody (Vector Labs, BA-4000, 1:500, 30 min at RT); primary anti-human p63 antibody (Abcam, ab735, 1:70 in M.O.M diluent; 30 min at RT) and secondary M.O.M. antibody (Vector Labs, BMK-2202,10 min at RT). Quantification was accessed as the percentage of positive cells in BE regions, which were defined as the defined BE region between squamous epithelium and oxyntic mucosa of the stomach or as the number of positive cells in 10 high-power fields of vision.

Transcriptional profile analysis

Total RNA from cardia and forestomach tissues of three mice for each phenotype were extracted using the RNA/Protein Kit (50) (80404, Qiagen) according to the manufacturers` instructions. RNA concentration and quality was assessed in a NanoDrop 2000 spectrophotometer (ThemoFisher). An Ambion WT Expression Kit (Thermo Fisher) was used to generate amplified sense-strand cDNA. Mouse gene 2.1st affymetrix arrays were used. Raw

data have been deposited in National Center for Biotechnology Information's Gene Expression Omnibus (GEO) are accessible through GEO accession (GSE103616). Gene set enrichment analysis (GSEA) was performed on the entire gene list ranked according to fold changes observed among WT mice and the phenotypes *pL2.Lgr5*, *pL2.Lgr5.N2fl/fl* and *pL2.Lgr5.N2IC*. Functional analysis was performed on the collapsed gene symbol list using GSEA with the MSigDB_v3.0 (Molecular Signatures Database) C2-C7 gene sets.[18] Additional gene sets generated with genetically engineered mouse prostate cancer epithelial cells with increased Notch signalling (GSE76822_Notch_signaling_up) were generated accessing the public GEO database using the top 50 significant genes with a Notch signalling dependent gene expression (GSE76822).[19]

Tissue preparation and disease evaluation

Mice were sacrificed after 7, 10, 13, and 16 months and at terminal stage. Subsequently, the organs were removed and subjected to downstream application. For macroscopic scoring, the stomach was opened along the large curvature and flattened for documentation. Each stomach and esophagus was evaluated for tumor coverage, individual tumor size, total tumor size, and summed for an overall macroscopic score as previously described.[25] Histopathology was evaluated using a previously described scoring system.[25] Macroscopic scoring of the squamocolumnar junction (SCJ) and the esophagus was performed following an established methods for dysplasia assessment in the mice that our group has published previously[20-24]. For further RNA or protein analysis the SCJ was macroscopically identified as the first 2 mm of columnar tissue and cut with a magnifying glass dissection-scope to eliminate squamous tissue contamination as good as possible. Mouse tissues were fixed in formalin and paraffin-embedded then cut and stained with H&E (Haematoxylins and eosin). *Histopathological scores were performed by an experienced mouse pathologist b*y previously established criteria for the influx of immune cells per high-power field, metaplasia and dysplasia[25]: *Inflammation was scored by the percentage of different immune cells (mostly neutrophilic myeloid cells) in a defined tissue area of the SCJ in a high-power field evaluation. Metaplasia was evaluated by*

the abundance of mucus producing or cells per gland and the abundance of glands with mucus producing cells in the BE area at the SCJ. Dysplasia was evaluated by the amount of cellular atypia and the presence of low or high grade dysplasia in single or multiple glands as assessed by experienced mouse pathologists. Mucus production was assessed by Periodic Acid-Schiff-(PAS) staining and quantified as percentage of PAS positive area in BE regions. Crypt fission was quantified by counting fused crypts in the BE region per 10 high power fields. Crypt fission was quantified by counting fused crypts in the BE region by experienced pathologists (KS and AS), using previously described methods [26].

3D organoid culture

The cardia and forestomach tissue of mice was extracted for organoid culture as described previously and as shown in *Supplementary Figure 2B*.[14, 26] To date there is no molecular technique to determine the specificity of BE organoids compared to gastric organoids other than growth characteristics and location of tissue sampling. In our experiments, we isolated organoids only from mouse tissue specimens taken from the SCJ. We followed our published protocol using media that supports the growth of BE organoids, whereas gastric gland organoids typically require additional growths factors.[26] In control experiments with SCJ tissue from WT mice, we were unable to generate any organoids. At least three independent primary organoid lines that were freshly isolated from our mice or thawed and expanded from our human Barrett's biobank were used for cell proliferation or differentiation experiments only in early passages 3-5 during optimal expansion rates. Cells were exposed to small molecule inhibitors DAPT (D5942, Sigma; 50μM) and JSH-23 (J4455, Sigma; 10μM) for 72h to block Notch and NF-κB signaling. Inhibitor concentrations were adjusted due to 3D matrix conditions and non-specific toxicity. Organoid growth and cell differentiation was evaluated according to microscopic analyses, cell activity assay using a MTT cell activity assay (M2128-Sigma-Aldrich) or flow cytometry. For immunohistochemistry, H&E and PAS staining in paraffin imbedded organoids were stained and 8-12 organoids on 3-5 slides were included in each experiment. Organoid survival was microscopically assessed according to the number of

organoids two days after isolation relative to those with a viable morphology at day seven after isolation.

SUPPLEMENTARY REFERENCES

1. Thorsen T, Maerkl SJ, Quake SR. Microfluidic large-scale integration. Science 2002;298:580-4.
2. Melin J, Quake SR. Microfluidic large-scale integration: the evolution of design rules for biological automation. Annu Rev Biophys Biomol Struct 2007;36:213-31.
3. Dalerba P, Kalisky T, Sahoo D, et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. Nat Biotechnol 2011;29:1120-7.
4. Schroeder A, Mueller O, Stocker S, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol 2006;7:3.
5. MacDonald JW. affycoretools: Functions useful for those doing repetitive analyses with Affymetrix GeneChips. 1.46.5 ed, 2008.
6. R Core Team. R: A Language and Environment for Statistical Computing}. 3.5,1 ed. Vienna, Austria: R Foundation for Statistical Computing, 2018.
7. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
8. Benjamini Y, Hochberg Y. Controlling the false discovery rate; A practical and powerful approach to multiple testing. J. Roy. Stat. Soc. Ser. B 1995;57:289-300.
9. Crawley M. The R Book. Hoboken: John Wiley & Sons, 2007.
10. Fox J, Weisberg HS. An R Companion to Applied Regression. Thousand Oaks CA: Sage, 2011.
11. Friendly M. Corrgrams: Exploratory displays for correlation matrices. The American Statistician 2002;56:316–324.
12. R Core Team. R: A Language and Environment for Statistical Computing}. 3.5.3 ed. Vienna, Austria: R Foundation for Statistical Computing, 2019.
13. Everitt BS, Landau S, Leese M. Cluster Analysis: Wiley, 2011.
14. Warnes GR, Bolker B, Bonebakker L, et al. gplots: Various R Programming Tools for Plotting Data. 3.0.1.1 ed, 2019.
15. Fox JG, Beck P, Dangler CA, et al. Concurrent enteric helminth infection modulates inflammation and gastric immune responses and reduces helicobacter-induced gastric atrophy. Nat Med 2000;6:536-42.
16. Lange S, Engleitner T, Mueller S, et al. Analysis pipelines for cancer genome sequencing in mice. Nat Protoc 2020.
17. Zagouras P, Stifani S, Blaumueller CM, et al. Alterations in Notch signaling in neoplastic lesions of the human cervix. Proc Natl Acad Sci U S A 1995;92:6414-8.
18. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545-50.
19. Revandkar A, Perciato ML, Toso A, et al. Inhibition of Notch pathway arrests PTEN-deficient advanced prostate cancer by triggering p27-driven cellular senescence. Nat Commun 2016;7:13719.
20. Munch NS, Fang HY, Ingermann J, et al. High-fat Diet Accelerates Carcinogenesis in a Mouse Model of Barrett's Esophagus via IL8 and Alterations to the Gut Microbiome. Gastroenterology 2019.
21. Fang HY, Munch NS, Schottelius M, et al. CXCR4 is a potential target for diagnostic PET/CT imaging in Barrett's dysplasia and esophageal adenocarcinoma. Clin Cancer Res 2017.
22. Schellnegger R, Quante A, Rospleszcz S, et al. Goblet Cell Ratio in Combination with Differentiation and Stem Cell Markers in Barrett Esophagus Allow Distinction of Patients with and without Esophageal Adenocarcinoma. Cancer Prev Res (Phila) 2017;10:55-66.
23. Lee Y, Urbanska AM, Hayakawa Y, et al. Gastrin stimulates a cholecystokinin-2-receptor-expressing cardia progenitor cell and promotes progression of Barrett's-like esophagus. Oncotarget 2016.

24. Quante M, Bhagat G, Abrams JA, et al. Bile acid and inflammation activate gastric cardia stem cells in a mouse model of Barrett-like metaplasia. Cancer Cell 2012;21:36-51.
25. Quante M, Bhagat G, Abrams JA, et al. Bile Acid and Inflammation Activate Gastric Cardia Stem Cells in a Mouse Model of Barrett-Like Metaplasia. Cancer Cell 2012;21:36-51.
26. Jin G, Ramanathan V, Quante M, et al. Inactivating cholecystokinin-2 receptor inhibits progastrin-dependent colonic crypt fission, proliferation, and colorectal cancer in mice. J Clin Invest 2009;119:2691-701.
27. Munch NS, Fang HY, Ingermann J, et al. High-Fat Diet Accelerates Carcinogenesis in a Mouse Model of Barrett's Esophagus via Interleukin 8 and Alterations to the Gut Microbiome. Gastroenterology 2019;157:492-506 e2.

SUPPLEMENTARY TABLES

Supplementary Table 1. Patient characteristics.

	All (n=164)	BE (n=116)	Controls (n=48)	p-value
Age	64.4 (11.4)	66.0 (10.4)	60.5 (13.0)	0.005
Sex, male	125 (76.2%)	98 (84.5%)	27 (56.3%)	<0.001
Race, white	162 (98.9%)	115 (99.1%)	47 (97.9%)	0.52
Ethnicity, non-Hispanic	162 (98.9%)	115 (99.1%)	47 (97.9%)	0.52
BMI*	29.0 (5.1)	29.4 (5.1)	27.8 (5.0)	0.10
WHR†	0.99 (0.06)	1.01 (0.06)	0.96 (0.06)	<0.001
PPI use‡	144 (92.9%)	111 (97.4%)	33 (80.5%)	<0.006
Aspirin	81 (49.4%)	64 (55.2%)	17 (35.4%)	0.08
Statins	49 (30.4%)	39 (34.2%)	10 (21.3%)	0.11
Smoking, ever	95 (57.9%)	73 (62.9%)	22 (45.8%)	0.06
Family hx BE/EAC	19 (11.6%)	16 (13.8%)	3 (6.3%)	0.36
Endoscopy and pathology:				
Hiatal hernia	117 (71.3%)	113 (97.4%)	4 (8.3%)	<0.001
HH size, median (IQR)	3 (2-4)	3 (2-4)	0 (0-1)	0.01
BE length, cm				
C, median (IQR)	--	1 (0-3)	--	--
M, median (IQR)	--	4 (2-7)	--	--
BE histology				
No dysplasia	--	44 (37.9%)	--	--
Indefinite	--	5 (4.3%)	--	--
LGD	--	21 (18.1%)	--	--
HGD	--	31 (26.7%)	--	--
EAC	--	15 (12.9%)	--	--

BMI, body mass index; WHR, waist-to-hip ratio; PPI, proton pump inhibitor; BE, Barrett's esophagus; EAC, esophageal adenocarcinoma
*BMI data missing on 21 subjects
†WHR data missing on 53 subjects
‡PPI data missing on 9 subjects

Supplementary Table 2. List of genes analyzed and associated TaqMan assay IDs.

Gene	*TaqMan assay ID*	*Gene*	*TaqMan assay ID*
GAPDH	Hs99999905_m1	*BMP4*	*Hs00370078_m1*
ACTB	Hs00357333_g1	CCK2R	Hs00176123_m1
POLR2A	Hs00172187_m1	*CCK2R*	*Hs01086284_g1*
NOTCH1	Hs00413187_m1	CCND1	Hs00765553_m1
NOTCH1	*Hs01062014_m1*	SOX2	Hs01053049_s1
NOTCH2	Hs01050702_m1	MYC	Hs00153408_m1
NOTCH2	*Hs00225747_m1*	NFKB1	Hs00765730_m1
NOTCH3	Hs01128537_m1	IL1B	Hs00174097_m1
NOTCH3	*Hs01128547_m1*	IL6	Hs00985639_m1
NOTCH4	Hs00965889_m1	CXCL8	Hs00174103_m1
NOTCH4	*Hs00965892_m1*	CXCR4	Hs00607978_s1
DLL1	Hs01011325_g1	TGFB1	Hs00998133_m1
DLL1	*Hs00194509_m1*	VEGFA	Hs00903129_m1
DLL3	Hs01085096_m1	EGFR	Hs00193306_m1
DLL3	*Hs01085097_m1*	PTGS2	Hs00153133_m1
DLL4	Hs01117332_g1	TLR4	Hs00152939_m1
DLL4	*Hs00184092_m1*	TLR3	Hs01551078_m1
JAG1	Hs01070032_m1	TLR7	Hs01933259_s1
JAG1	*Hs00164982_m1*	TLR9	Hs00370913_s1
JAG2	Hs01000098_g1	NOD1	Hs00196075_m1
HES1	Hs01118948_g1	NOD2	Hs00223394_m1
HES1	*Hs00172878_m1*	WNT7A	Hs00171699_m1
HEY1	Hs01114113_m1	WNT7B	Hs00536497_m1
HEY1	*Hs00232618_m1*	WNT9A	Hs00243321_m1
MAML1	Hs01070499_m1	LEF1	Hs01547250_m1
MAML1	*Hs00207373_m1*	CDH1	Hs01023894_m1
NRARP	Hs01104102_s1	CTNNB1	Hs00355049_m1
NRARP	*Hs04183811_s1*	TP53	Hs01034249_m1
BIRC5	Hs00153353_m1	CDKN2A	Hs00923894_m1
MKI67	Hs00267195_m1	SMAD4	Hs00929647_m1
TOP2A	Hs01032137_m1	ARID1A	Hs00195664_m1
PARP1	Hs00242302_m1	PIK3CA	Hs00907957_m1
CASP3	Hs00234387_m1	ERBB2	Hs01001580_m1
ANXA5	Hs00996187_m1	GATA6	Hs00232018_m1
EPCAM	Hs00158980_m1	KRAS	Hs00364284_g1
KRT20	Hs00300643_m1	MET	Hs01565576_m1
CDX2	Hs01078080_m1	HNF4A	Hs00230853_m1
TFF1	Hs00907239_m1	SP1	Hs00916521_m1
TFF2	Hs00193719_m1	NFYA	Hs00953589_m1
TFF3	Hs00173625_m1	E2F1	Hs00153451_m1
MUC2	Hs03005094_m1	ESR1	Hs00174860_m1
MUC5AC	Hs00873651_mH	HIF1A	Hs00153153_m1
MUC6	Hs01674026_g1	NR1H4	*Hs01026590_m1*
LGR5	Hs00969421_m1	FABP6	*Hs01031183_m1*
LGR5	*Hs00969423_m1*	CCL20	*Hs00355476_m1*
DCLK1	Hs00178027_m1	NR0B2	*Hs00222677_m1*
DCLK1	*Hs00973855_m1*	FGF19	*Hs00192780_m1*
BMP4	Hs03676628_s1	MIR221	*Hs04231481_s1*

*Assays in italics represent secondary assays analyzed in the event the primary chosen assay did not produce high quality data.

Supplementary Table 3: Primers for Real-time Quantitative PCR

Target	Sequence Forward	Sequence Reverse	Amplicon Size
Notch1	ACAGTGCAACCCCCTGTATG	TCTAGGCCATCCCACTCACA	102bp
Notch2	CCCAGAACCAATCAGGTTAGC	GCCGAGACTCTAGCAATCACAA	109bp
GapDH	GACATCAAGAAGGTGGTGAAGCAG	ATACCAGGAAATGAGCTTGACAAA	174bp
CyclophilinA	ATGGTCAACCCCACCGTGT	TTCTGCTGTCTTTGGAACTTTGT	102bp
Hes1	CAACACGACACCGGACAAAC	GGAATGCCGGGAGCTATCTT	157bp
Atoh1	CTTCGTTGAACTGGGTTGCC	TAGACGGGAAGGTCTCTCGC	202bp
Muc2	GTCCCGACTTCAACCCAAGTGA	TGGTGCAGCCATTGTAGGAAAT	150bp
TFF2	CTGGTAGAGGGCGAGAAACC	TGCTCCGATTCTTGGTTTGGA	182bp
RelA-A	ACGAGGCTCCTTTTCTCAAGCT	GTCGGCGTACGGAGGAGTC	73bp
RelA-B	TCTGCCGAGTAAACCGGAAC	GCACCTTGTCGCACAGCA	72bp
Jag1	ATGCAGAACGTGAATGGAGAG	GCGGGACTGATACTCCTTGAG	132bp
Jag2	CAATGACACCACTCCAGATGAG	GGCCAAAGAAGTCGTTGAG	203bp
Ki67	ATCATTGACCGCTCCTTTAGGT	GCTCGCCTTGATGGTTCCT	104bp
Lgr5	GACGCTGGGTTATTTCAAGTTCA	CAGCCAGCTACCAAATAGGTGCT	150bp
EGFR	CATGCGAAGACGTCACATTGTT	GGGTGTGAGAGGTTCCACGA	82bp

Supplementary Table 4. Association between gene expression of markers of Notch signaling with markers of goblet cell differentiation.

	TFF3				***MUC2***	
	rho	p			rho	p
NOTCH3	-0.60	$<2.2 \times 10^{-16}$		*NOTCH3*	-0.56	1.5×10^{-9}
JAG2	-0.60	$<2.2 \times 10^{-16}$		*JAG2*	-0.55	5.4×10^{-9}
NOTCH2	-0.56	2.5×10^{-9}		*JAG1*	-0.53	2.2×10^{-8}
JAG1	-0.55	6.9×10^{-9}		*NOTCH4*	-0.51	1.1×10^{-7}
NOTCH4	-0.48	5.6×10^{-7}		*NOTCH2*	-0.47	8.7×10^{-7}
NOTCH1	-0.47	1.4×10^{-6}		*NOTCH1*	-0.42	2.0×10^{-5}
HES1	-0.35	4.7×10^{-4}		*HES1*	-0.30	0.002
HEY1	-0.22	0.03		*HEY1*	-0.27	0.007
DLL1	-0.14	0.16		*MAML1*	-0.17	0.10
MAML1	-0.13	0.19		*DLL1*	-0.14	0.16
DLL3	0.18	0.08		*DLL3*	0.09	0.38
DLL4	0.21	0.03		*DLL4*	0.13	0.18
NRARP	0.37	1.4×10^{-4}		*NRARP*	0.34	6.8×10^{-4}

Supplementary Table 5. Notch gene expression in patients with Barrett's esophagus, comparing those with HGD or EAC vs. no dysplasia.

Gene	***log-2-fold change***	***FDR***
JAG2	0.91	0.04
NOTCH3	0.61	0.08
JAG1	0.23	0.16
NOTCH1	0.27	0.19
NOTCH4	0.25	0.33
NOTCH2	0.16	0.43
HEY1	0.28	0.46
DLL3	-1.95	0.58
DLL1	0.09	0.72
NRARP	0.07	0.72
HES1	-0.04	0.83
DLL4	-0.08	0.85
MAML1	0.00	1.00

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1.

Biopsy slides stained with hematoxylin and eosin were digitally scanned, and then overlaid with 200 µm x 200 µm grids, shown at low power (A, top image) and high power (B, bottom image). (C) Each evaluable grid was then scored for goblet cell density based on prespecified criteria. (D) There were significant correlations between expression of *NOTCH3* with *TGFB1*, *VEGF*, *MYC*, *CASP3*, and *CTTNB1* (E); and *JAG2* with *TGFB1*, *EGFR*, *MYC*, *CASP3*, and *CTTNB1*.

Supplementary Figure 2.

(A) IHC staining of Notch-IC. Shown are representative examples of indicated mouse strains and varying age. (B) Macroscopic image of murine esophagus, SCJ, and stomach. A similar image was shown in our previous publication[27] (C) Representative images of IHC for the indicated marker proteins using 3D cultured organoids from *pL2.Lgr5.N2IC* mice, exposed to the gamma-secretase inhibitor DAPT.

Supplementary Figure 3

Global gene expression analyses reveal Notch-dependent gene signatures. GSEA that are based on the global transcriptome of indicated mouse strains with mitigated Notch signaling.

Supplementary Figure 4

To demonstrate DNA alteration in *pL2.Lgr5.N2IC* mice compared to *L2-IL1B* mice dysplastic tissue areals from six *pL2.Lgr5.N2IC* and six *L2-IL1B* mice were microdissected and DNA was isolated for Library preparation (NEBNext Ultra II FS DNA Library Prep Kit for Illumina). Samples were sequenced on an Illumina NextSeq system and processed (Lange et al., Nature Protocols 2019) demonstrating increased copy number alterations (arrows) compared to controls.

Supplementary Figure 5

(A) Representative IHC images of small intestine tissue of *pL2.Lgr5*, *pL2.Lgr5N2fl/fl,* and *pL2.Lgr5.N2IC* mice using Ki67 antibody and performing Alcian Blue staining at indicated time points. The diagrams show the respective statistical evaluations of (B) goblet cell and (C) Ki67+ cell frequencies. Data are presented as means ± standard deviation. Statistical analysis was performed using one-way ANOVA and Tukeys multiple comparison test. *p<.05.

Supplementary Figure 6

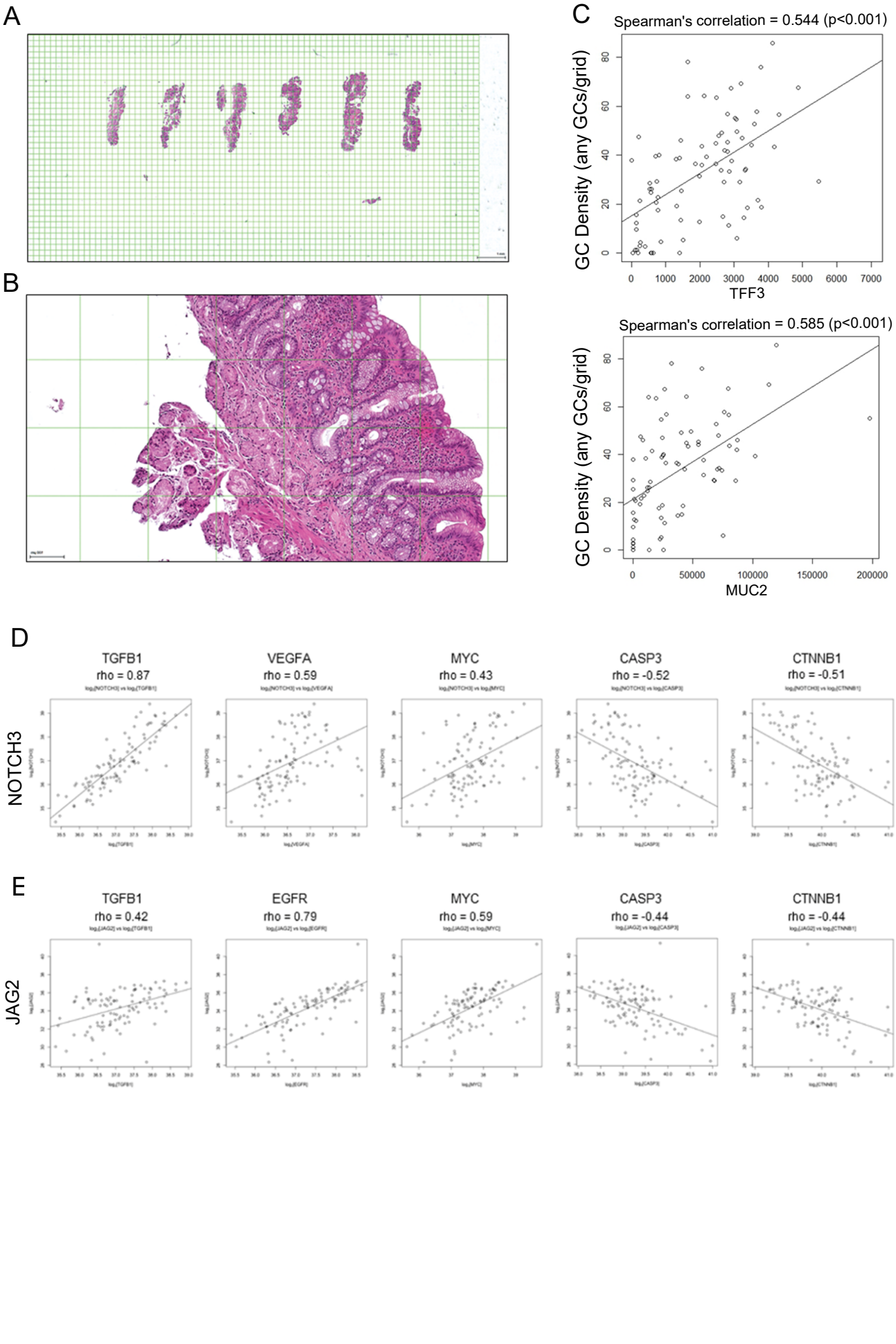
Immune cell frequencies show no distinct differences in *pL2* mice strains with varying Notch signaling. Gating strategy of myeloid lineage cells (A) and relevant T cell populations (B) from the cardia and the forestomach region. Dot plots display are representative analyses of *pL2-IL1b* mice that were pre-selected for CD45 and viability. The statistical summary of each population, using at least three mice, is indicated according to the following surface patterns: Macrophages (CD11b+F4/80+Ly6Glow), neutrophils (CD11b+F4/80-Ly6Ghigh), immature myeloid cells (IMC, CD11b+Ly6ChighLy6Ghig), T helper cells (CD3+CD4+), Cytotoxic T cells (CD3+CD8+), NK cells (CD3+CD4-CD8-NK1.1+), and gamma delta T cells (CD3+CD4-CD8-gdT+). (C) FACS results were statistically summarized. Data is presented as means ± standard deviation. Statistical analysis was performed using one-way ANOVA and Tukeys multiple comparison test. *p<.05.

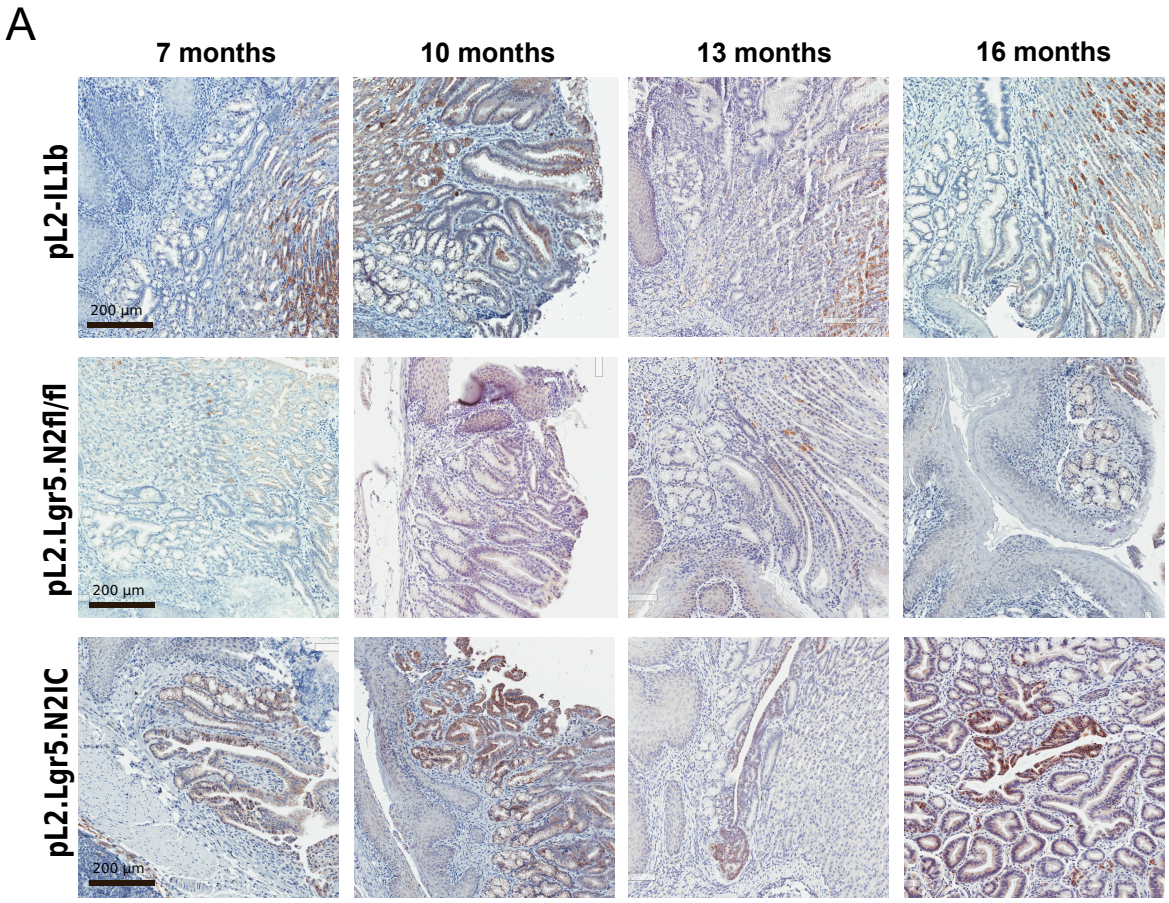
Supplementary Figure 7

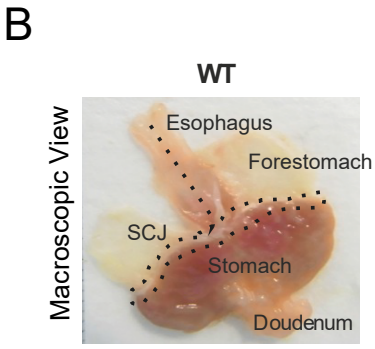
Disease progression directly correlates with expression intensity of marker genes. Statistical summary of Pearson correlations that were performed using gene chip data in regard to Notch and NFkB signaling, goblet cell differentiation and stem cell maintenance. The table summarizes the displayed diagrams.

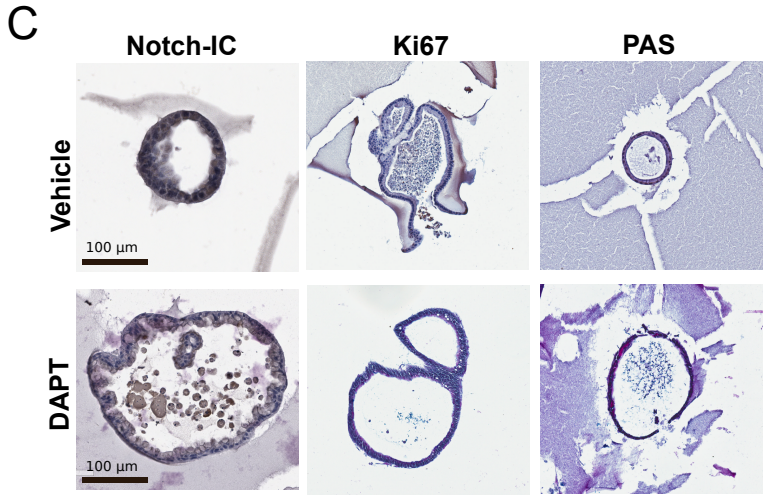
Supplementary Figure 8

Statistical summary of Pearson correlations that were performed using pathological scores and average gene chip data of each mouse strain in regard to Notch and NFkB signaling, goblet cell differentiation and stem cell maintenance. The table summarizes the displayed diagrams.

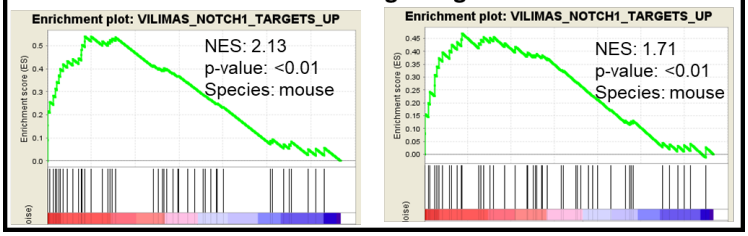

A
B
C
Spearman's correlation = 0.544 (p<0.001)
GC Density (any GCs/grid)
TFF3
Spearman's correlation = 0.585 (p<0.001)
GC Density (any GCs/grid)
MUC2
D
NOTCH3
TGFB1
rho = 0.87
VEGFA
rho = 0.59
MYC
rho = 0.43
CASP3
rho = -0.52
CTNNB1
rho = -0.51
E
JAG2
TGFB1
rho = 0.42
EGFR
rho = 0.79
MYC
rho = 0.59
CASP3
rho = -0.44
CTNNB1
rho = -0.44


A
7 months
10 months
13 months
16 months
pL2-IL1b
pL2.Lgr5.N2fl/fl
pL2.Lgr5.N2IC
200 μm
200 μm
200 μm

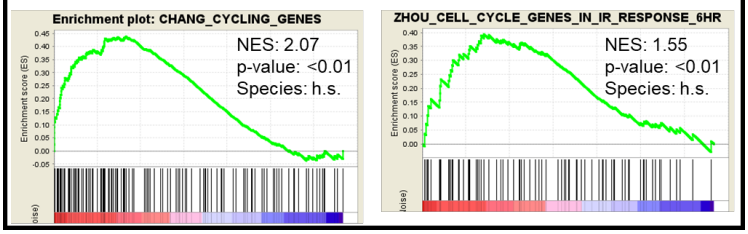

B
WT
Macroscopic View
Esophagus
Forestomach
SCJ
Stomach
Doudenum


C
Notch-IC
Ki67
PAS
Vehicle
DAPT
100 μm
100 μm

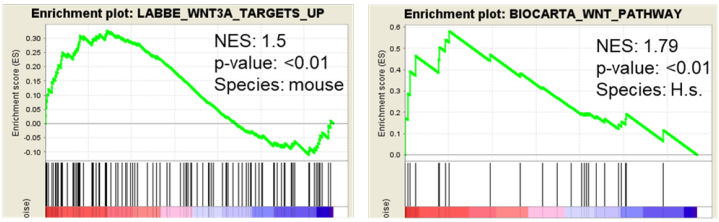
Notch signaling



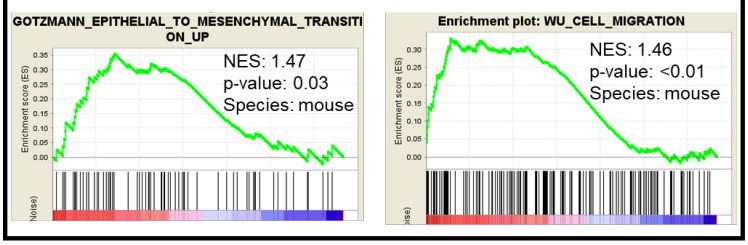
Proliferation



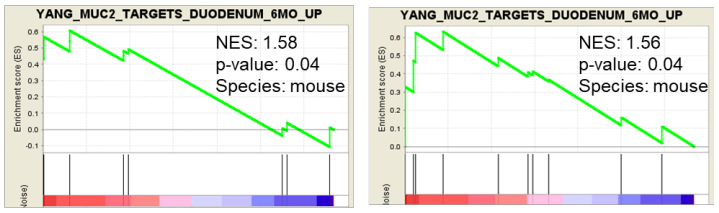
Wnt signaling



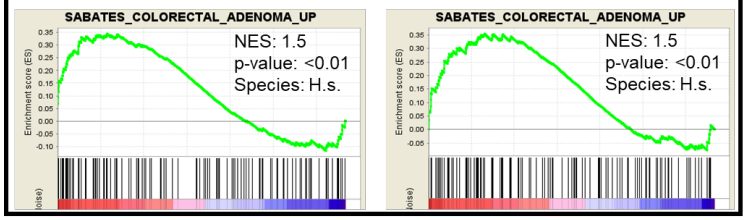
Migration



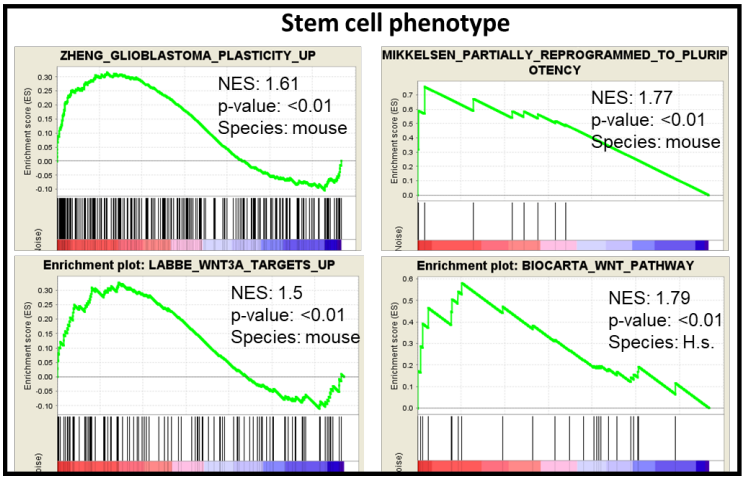
Goblet cell maturation inhibition



Cancer

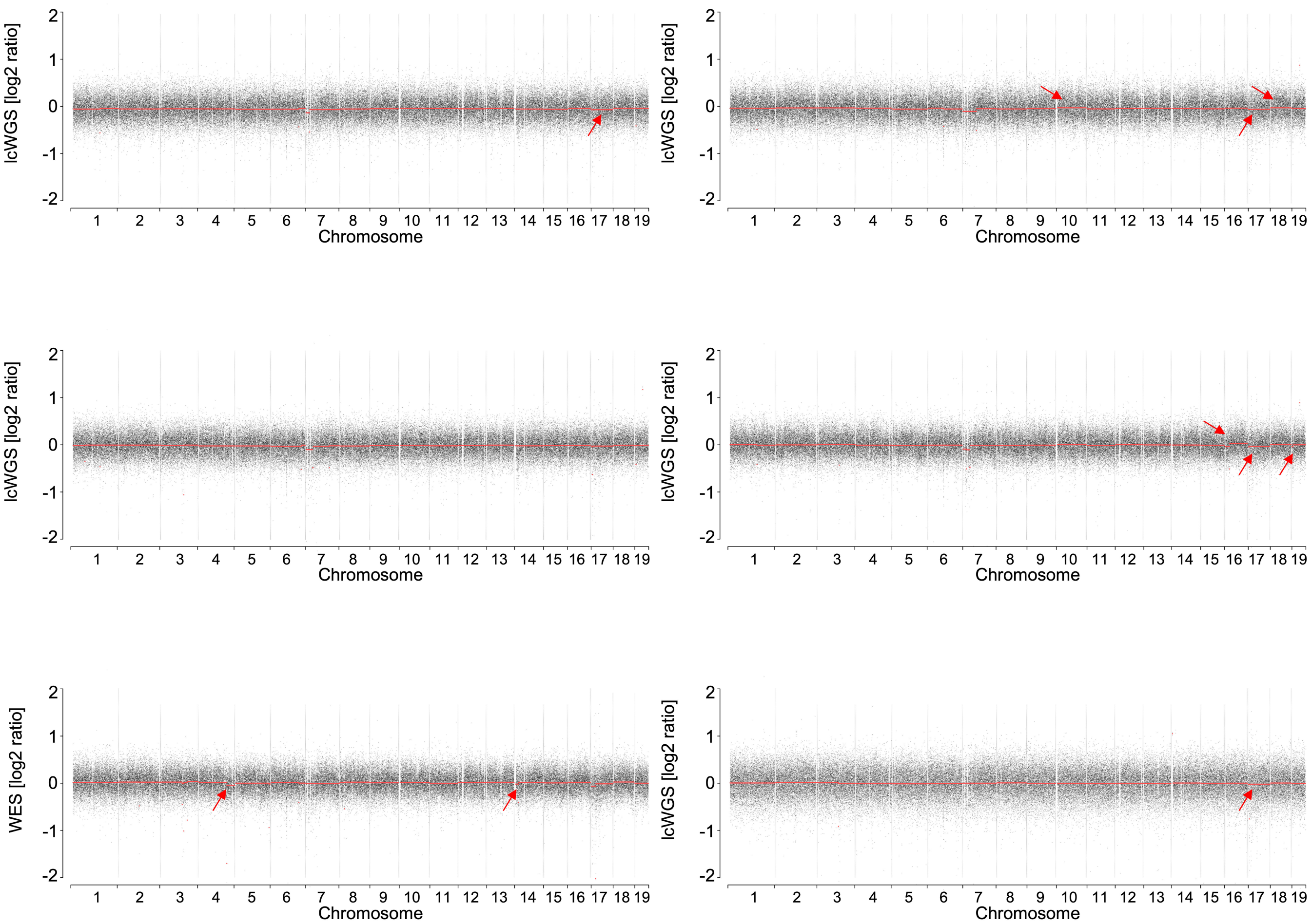


Stem cell phenotype

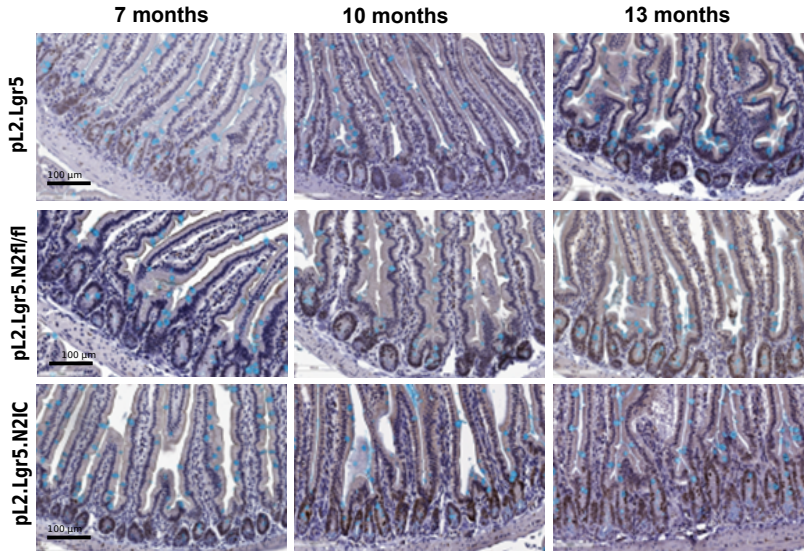


pL2.Lgr5.N2IC vs pL2-IL1b pL2.Lgr5.N2IC vs pL2.Lgr5.N2fl/fl

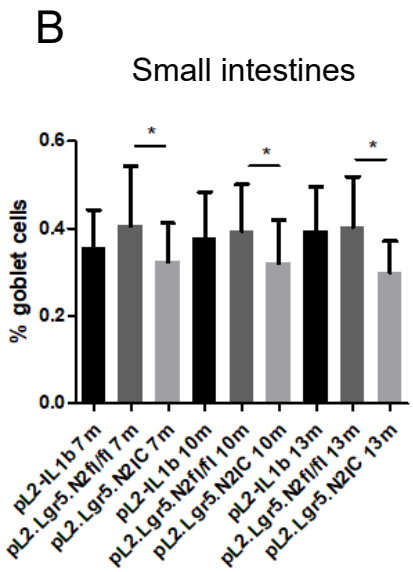
pL2.Lgr5.N2IC vs pL2-IL1b pL2.Lgr5.N2IC vs pL2.Lgr5.N2fl/fl


lcWGS [log2 ratio]
WES [log2 ratio]
Chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
2 1 0 -1 -2

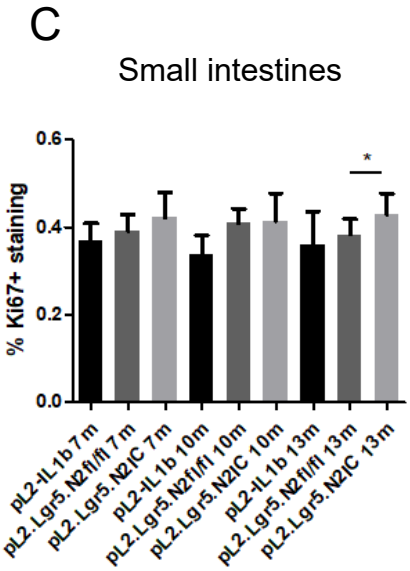
A

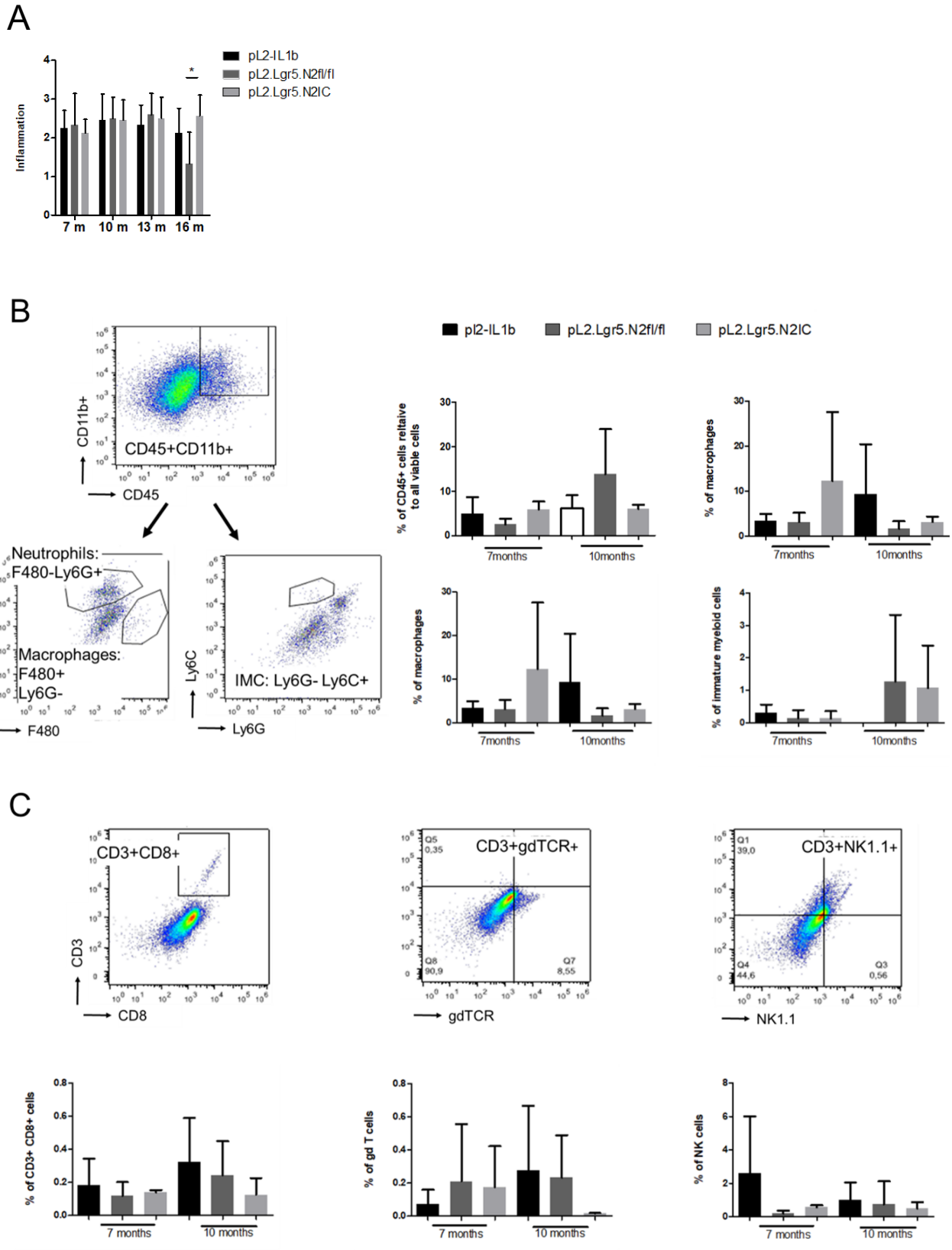

7 months
10 months
13 months
pL2.Lgr5
pL2.Lgr5.N2fl/fl
pL2.Lgr5.N2IC
100 µm

B


Small intestines
% goblet cells
0.6
0.4
0.2
0.0
*
*
*
pL2-IL1b 7m
pL2.Lgr5. N2fl/fl 7m
pL2.Lgr5.N2IC 7m
pL2-IL1b 10m
pL2.Lgr5.N2fl/fl 10m
pL2.Lgr5.N2IC 10m
pL2-IL1b 13m
pL2.Lgr5.N2fl/fl 13m
pL2.Lgr5.N2IC 13m

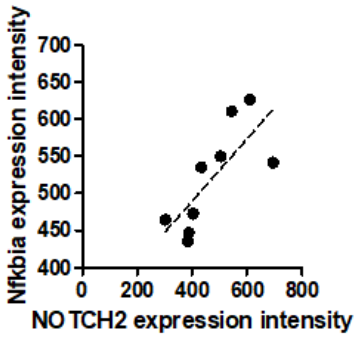
C

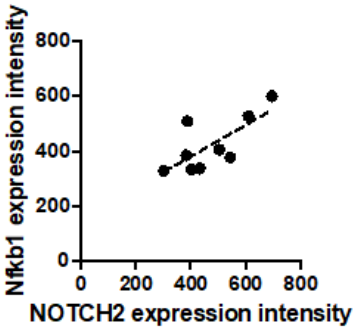

Small intestines
% Ki67+ staining
0.6
0.4
0.2
0.0
*
pL2-IL1b 7m
pL2.Lgr5. N2fl/fl 7m
pL2.Lgr5.N2IC 7m
pL2-IL1b 10m
pL2.Lgr5.N2fl/fl 10m
pL2.Lgr5.N2IC 10m
pL2-IL1b 13m
pL2.Lgr5.N2fl/fl 13m
pL2.Lgr5.N2IC 13m

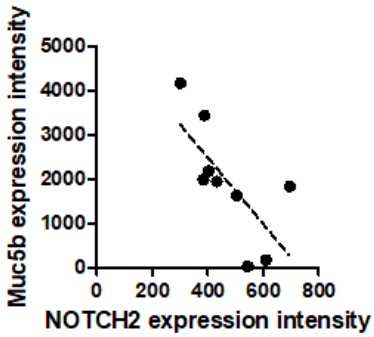

A
pL2-IL1b
pL2.Lgr5.N2fl/fl
pL2.Lgr5.N2IC
Inflammation
7 m
10 m
13 m
16 m
*
B
pl2-IL1b
pL2.Lgr5.N2fl/fl
pL2.Lgr5.N2IC
CD11b+
CD45
CD45+CD11b+
Neutrophils:
F480-Ly6G+
Macrophages:
F480+
Ly6G-
Ly6G
F480
Ly6C
Ly6G
IMC: Ly6G- Ly6C+
% of CD45+ cells reltaive to all viable cells
% of macrophages
% of macrophages
% of immature myeloid cells
7months
10months
C
CD3+CD8+
CD3
CD8
CD3+gdTCR+
gdTCR
CD3+NK1.1+
NK1.1
% of CD3+ CD8+ cells
% of gd T cells
% of NK cells
7 months
10 months

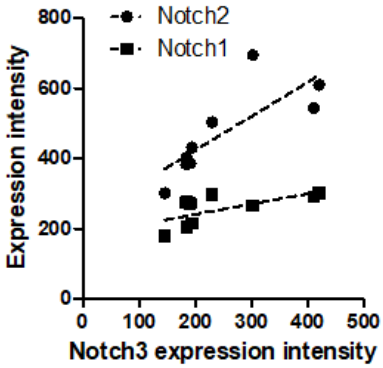
	Notch2 expression vs			Notch3 expression vs					Hey1 expression vs		
	Nfkbia	Nfkb1	Muc5b	Notch2	Notch1	Wnt5a	Muc5b	Muc5ac	Rela	Cxcr4	Dtx1
Implication	**b**	**b**	**c**	**a**	**a**	**d**	**c**	**c**	**b**		**a**
Number of XY Pairs	9	9	9	9	9	9	9	9	9	9	9
Pearson r	0.7544	0.7266	-0.7067	0.7928	0.6733	0.7066	-0.8680	-0.6977	0.6799	0.7776	0.8648
P value (two-tailed)	0.0188	0.0266	0.0333	0.0108	0.0468	0.0333	0.0024	0.0366	0.0439	0.0136	0.0026
R squared	0.5691	0.5279	0.4994	0.6286	0.4534	0.4993	0.7534	0.4868	0.4622	0.6047	0.7479

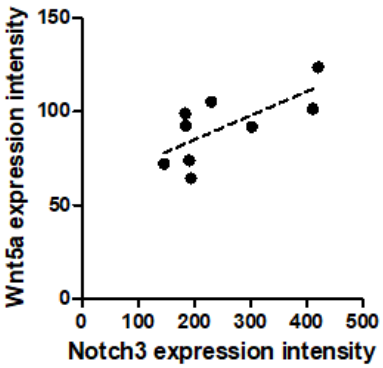
a, involved in Notch signaling; b, involved in NFkB signaling; c, goblet cell marker; d, stem cell associated

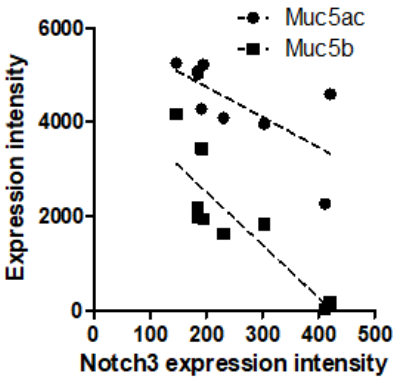


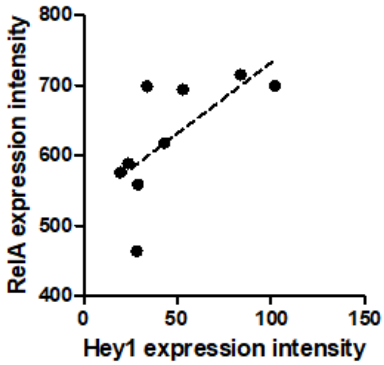


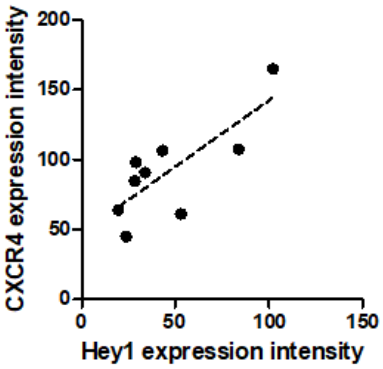


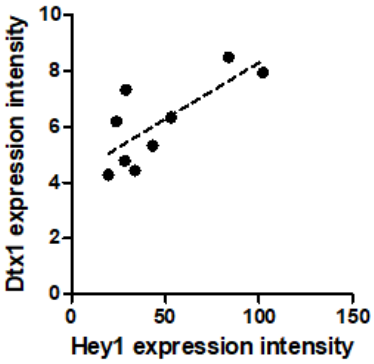












	Macroscopic score vs gene expression			Dysplasia score vs gene expression			Goblet cell ratio vs gene expression		
	Notch2	Dtx2	Muc5ac	Nfkbia	Irf1	Cxcr2	Nfkbia	Bcl3	Nanog
Implication	**a**	**a**	**c**	**b**	**b**	**b, d**	**b**	**b**	**d**
Number of XY Pairs	3	3	3	3	3	3	3	3	3
Pearson r	0.9970	0.9999	-0.9972	0.9984	0.9988	0.9993	-0.9987	-1.000	-0.9999
P value (two-tailed)	0.0492	0.0107	0.0479	0.0365	0.0318	0.0237	0.0330	0.0045	0.0087
R squared	0.9940	0.9997	0.9944	0.9967	0.9975	0.9986	0.9973	0.9999	0.9998

a, involved in Notch signaling; b, involved in NFkB signaling; c, goblet cell marker; d, stem cell associated