EPITHELIAL STEM AND PROGENITOR CELLS: RESPONSES TO HOMEOSTATIC AND ONCOGENIC CUES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Preetish Kadur Lakshminarasimha Murthy

August 2019

EPITHELIAL STEM AND PROGENITOR CELLS: RESPONSES TO HOMEOSTATIC AND ONCOGENIC CUES

Preetish Kadur Lakshminarasimha Murthy, Ph. D.

Cornell University 2019

In this study we have tried to understand the epithelial tissue, which froms the barrier between the external environment and the body, in homeostatic and oncogenic states. Notch signalling maintains epithelial stem cell regeneration at the mouse intestinal crypt base and balances the absorptive and secretory lineages in the upper crypt and villus. In the first part of this dissertation, we report the role of Fringe family of glycosyltransferases in modulating Notch activity in the two compartments. At the crypt base, RFNG is enriched in the Paneth cells and increases cell surface expression of DLL1 and DLL4. This promotes Notch activity in the neighbouring *Lgr5*+ stem cells assisting their self-renewal. Expressed by various secretory cells in the upper crypt and villus, LFNG promotes DLL surface expression and suppresses the secretory lineage. Hence, in the intestinal epithelium, Fringes are present in the ligand-presenting 'sender' secretory cells and promote Notch activity in the neighbouring 'receiver' cells. Fringes thereby provide for targeted modulation of Notch activity and thus the cell fate in the stem cell zone, or the upper crypt and villus.

Mutations in the KRAS gene are frequently found in human lung tumours and are known to drive their proliferation. In the second half of this study, we show that mutant *Kras* remodels the epigenetic landscape of Club and AEC2 cells of the lung.

We find that FOSL1 based AP-1 transcription factor directly binds to and guides the nucleosome remodelling SWI/SNF complex to increase chromatin accessibility at genomic loci controlling the expression of genes necessary for neoplastic transformation. Pharmacological inhibition of AP-1 activity abrogated proliferation specifically of tumour cells but not normal cells. We observe that tumours retain a signature of their cell of origin and demonstrate that AP-1 mediated epigenetic reprogramming is not restricted to pulmonary epithelium but can also be seen in intestinal stem cells, suggesting that it is a general mechanism downstream of mutant *Kras*.

BIOGRAPHICAL SKETCH

Preetish was born in Birur, Karnataka, India in 1991. He finished his schooling in Bangalore and went on to obtain a degree in mechanical engineering from Indian Institute of Technology Madras in 2013. He worked on understanding fluid flows in microchannels using Extended Transport Equations during his undergraduate days. With an interest in academic research he joined the graduate program at Cornell University where he worked, in Dr. Xiling Shen's lab, on understanding the epithelial tissues under homeostatic and pathologic conditions.

Dedicated to my mother.

ACKNOWLEDGMENTS

I would like to acknowledge the contribution of many a friend, family member and colleague without whom I would not have been able to complete this work. Firstly, I would like to thank my thesis committee chair Prof. Brian Kirby for being supportive throughout my stay at Cornell and making it possible for me to work *in absentia* at Duke University for over three years to complete my thesis. I thank Dr. Xiling Shen, who provided me the opportunity and trusted me to work on intestinal and pulmonary epithelial tissues for the past six years. I would like to thank Dr. Ankur Singh for serving on my thesis committee and providing invaluable assistance.

I extend my sincere gratitude to my lab-mate Rui Xi for her continued scientific and moral support over the years. I thank Tara Srinivasan, Diana Arguijo, Matthew Bochter, Susan Cole and Mirjana Maletic-Savatic for their help with the experiments. Scientific advice from a multitude of Duke faculty including Drs. Brigid Hogan, PR Tata, Gregory Crawford, Jeffrey Everitt and Tim Reddy has been instrumental in the training I received during my stint as a doctoral candidate.

I thank my parents, brother and wife for their constant support and trust which made my stay in the United States a lot easier. Lastly, I thank Marcia Sawyer for her incredible administrative support.

TABLE OF CONTENTS

CHAPTER 1

BACKGROUND

1.1 Epithelium

Epithelial cells form a tight barrier on the outer surface of our organs and protect them from the external environment. They are susceptible to injury from the harsh environments they are often exposed to. For example, the epidermis of the skin can be exposed to UV radiation, the pulmonary epithelium can be harmed by toxic gases and pollutants, and the gastrointestinal epithelium can be targeted by pathogenic microbes (Peterson and Artis, 2014; Whitsett and Alenghat, 2015). Hence, they are either under a constant state of proliferation to replace the dying cells, termed homeostasis, as in the small intestinal epithelium or epidermis of skin, or possess a remarkable capability to regenerate after injury as seen in the pulmonary or hepatic epithelial tissue (Desbarats and Newell, 2000; Kotton and Morrisey, 2014; Morasso and Tomic-Canic, 2005; van der Flier and Clevers, 2009).

Cellular proliferation under both homeostatic and regenerative conditions is fueled by tissue resident stem cells (Rawlins and Hogan, 2006). These cells are often located in a mesenchymal niche that provides the necessary growth and supportive factors to help them maintain their proliferative potential and ‘stemness’, defined as the ability to self-renew and differentiate into mature cell-types (Rawlins and Hogan, 2006). Stem cells divide and often give rise to a highly proliferative progenitor population which further divides and differentiates into fully functioning mature cells (Rawlins and Hogan, 2006). It is interesting to observe the dynamic equilibrium at which these tissues

exist and is essential to understand the signals that help them be. Here, we have tried to understand how the epithelium of the intestine is maintained in a state of dynamic equilibrium.

Tissues have robust mechanisms to correct deviations from the steady homeostatic state (Rawlins and Hogan, 2006). However, somatic mutations to genome of stem or progenitor cells upon exposure to radiation, chemicals or chronic inflammation can lead to an irreversible change: neoplastic transformation (Greenman et al., 2007). In this work, we have tried to understand the initial stages of transformation and have concentrated on characterising the changes to the epigenome that makes this perturbation irreversible.

1.2 Notch signalling pathway

Notch is an evolutionarily conserved juxtacrine signalling pathway comprising mainly of Notch receptors and ligands. Mammals have four types of Notch receptors (NOTCH1, NOTCH2, NOTCH3 and NOTCH4) (Bray, 2006). There are three main delta like ligands (DLL1, DLL3 and DLL4) and two serrate like (JAG1 and JAG2) classified based on their homology to Delta and Serrate ligands of *Drosophila* (D'souza et al., 2008). Both the receptors and the ligands are transmembrane proteins with their extracellular domain (ECD) containing 29-36 epidermal growth factor (EGF) like repeats (Bray, 2006; Stahl et al., 2008). In the canonical Notch pathway, the ligand binds to the receptor on a neighboring cell and exposes the transmembrane domain of the receptor for S2 cleavage by ADAM/TACE metalloproteases (Bray, 2006; Kopan,

2002). The receptor undergoes another cleavage (S3) by γ-secretase and the intra-cellular domain (NICD) translocates to the nucleus (Bray, 2006; Kopan, 2002). Here, it binds with CSL complex and Mastermind to regulate transcription of a host of genes including *Hes* and *Hey* families (Iso et al., 2003; Kopan, 2002).

Post-translational modification (PTM) of the Notch receptor, especially by glycosylatransferases in the Golgi complex is an important determinant of the signalling pathway (Xu and Egan, 2018). Glycosylation of EGF like repeats on the ECD of Notch receptor by *O*-fucosyltransferase1 (POFUT1) is necessary for its binding to DLL and JAG ligands (Stahl et al., 2008). The fucosylated product can further be elongated with the addition of β-N-acetylglucosamine by fringe family proteins (Harvey et al., 2016). Other than the addition of *O*-fucose, *O*-glucose is also added to the ECD of Notch by POGLUT1 and might influence Notch activity (Rana et al., 2011). Other than that, N-acetylglucosamine can also be linked to the Serine or Threonine residues on the Notch ECD by glycosyltransferase EOGT and the modification affects notch signalling (Sawaguchi et al., 2017).

Among the PTMs of Notch receptor, those by the Fringe proteins have received researchers' attention owing to their interesting effects (Xu and Egan, 2018). Three Fringe proteins, Radical (RFNG), Lunatic (LFNG) and Manic (MFNG), are found in mammals (Haines and Irvine, 2003). Modification of NOTCH1 by Lunatic and Manic fringes supports its activation by DLL1 and decreases it activation by JAG1 (Haines and Irvine, 2003; Hicks et al., 2000; Murthy et al., 2018; Panin et al., 1997). But, RFNG

mediated glycosylation of NOTCH1 increases its activation by both DLL1 and JAG1 (LeBon et al., 2014; Murthy et al., 2018).

Notch pathway has a role in many a developmental, homeostatic, regenerative and pathologic context. It is essential for proper formation segmental boundaries between somites during development (Weinmaster and Kintner, 2003). It is also important in development of most organs in the body (Weinmaster and Kintner, 2003). It has a role in maintaining the epithelium of stomach, intestine, colon and epidermis under homeostatic conditions (Kim and Shivdasani, 2011a; Miyamoto and Rosenberg, 2011; Pellegrinet et al., 2011; Watt et al., 2008). Notch signalling is important for the regeneration of exocrine pancreas after acute pancreatitis (Siveke et al., 2008). Genomic translocation leading to the expression of constitutively active form of NOTCH1 leads to the formation of T cell Acute Lymphoblastic Leukemia (T ALL) (Allenspach et al., 2002). Notch pathway also plays a role in skin, colon, lung, cervical and prostate cancers (Allenspach et al., 2002).

1.3 Notch signalling and the intestinal epithelium

Mammalian intestinal surface has invaginations called crypts and finger-like projections called villi (Figure 1.1, adapted from (Murthy et al., 2018)). The epithelial lining is simple columnar and is one of the fastest regenerating tissues in the body. At the base of the crypts *Lgr5* expressing stem cells are found (Barker et al., 2007). They undergo constant symmetric cell division to replenish the lost cells (Snippert et al., 2010). A mesenchymal niche provides WNT signal to maintain their stemness (Kabiri

et al., 2014). As they divide, the daughter cells move up the crypt and leave the stem cell niche (Noah et al., 2011). They continue to proliferate as progenitors (also called transit amplifying cells) in the upper crypt and differentiate into mature cell types (Noah et al., 2011). Two major types of differentiated cells are found lining the intestine, the secretory cells and absorptive cells. Enterocytes form the absorptive lineage and as the name suggests, they are involved in nutrient absorption (Ziv and Bendayan, 2000). Four main types of secretory cells are seen:

1. Goblet cells

   These are mucus secreting cells found scattered throughout the upper crypt and villus and are the most abundant among the secretory lineages. They are found in large numbers around Peyer's patches (Pelaseyed et al., 2014).

2. Enteroendocrine cells (EEC)

   These are specialized cells capable of secreting neuropeptides and are found to be in direct contact with the neurons (Kaelberer et al., 2018). They scattered throughout the upper crypt and villus are involved in sensing the gut luminal contents and signalling to the brain (Latorre et al., 2016).

3. Tuft cells

   These are rare cells constituting barely 0.4% of the total epithelium and are found scattered in the upper crypt and villus. They appear to be similar to cells in taste buds (Gerbe et al., 2012).

4. Paneth cells

   They are involved in the secretion of anti-microbial peptides like α-defensins (Ayabe et al., 2000). In contrast to all the other secretory cell types, Paneth cells

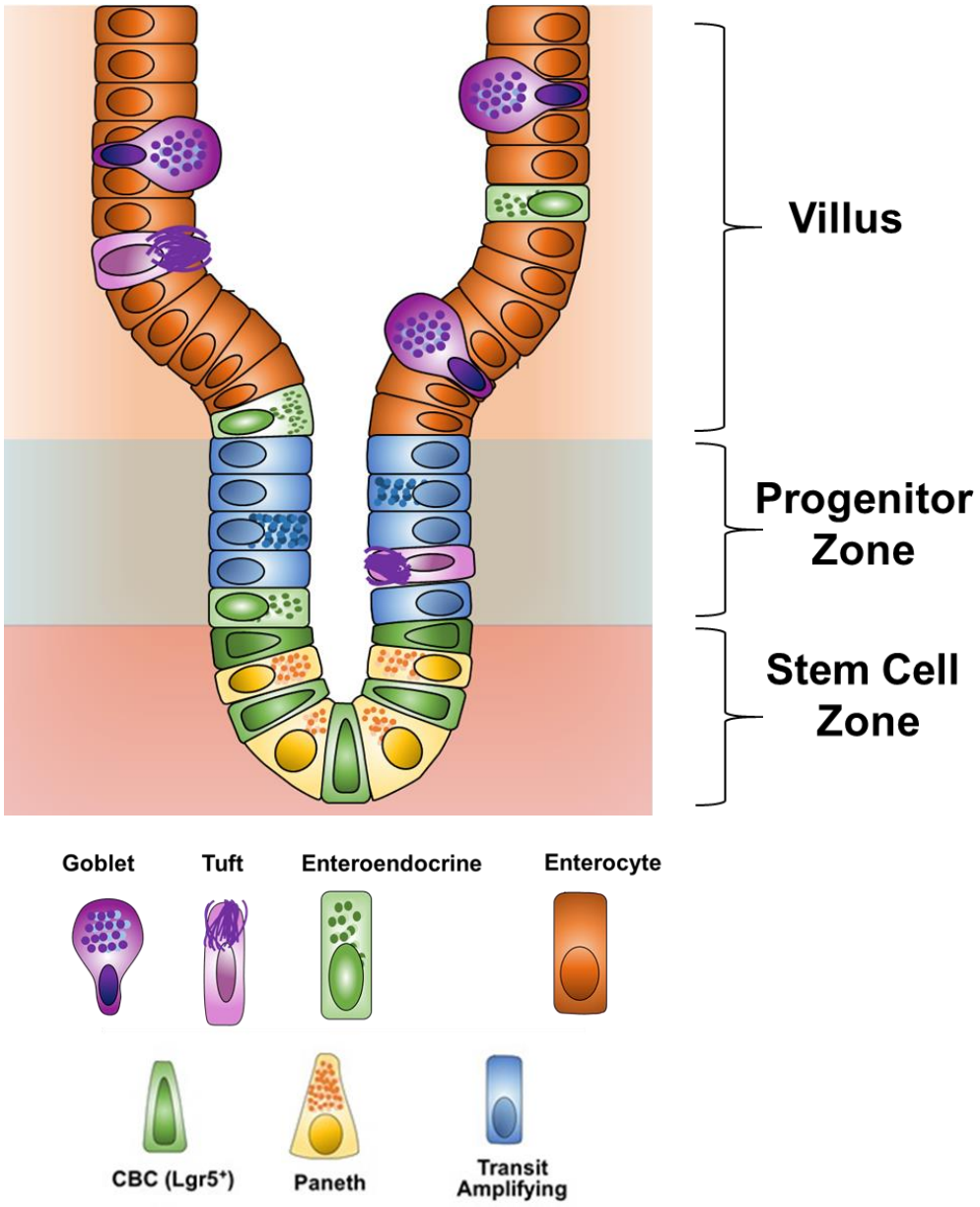


Figure 1.1 Schematic representation of the intestinal epithelium showing various zones of the crypt and villus, and the cell types present in them.

migrate to the bottom of the crypts and provide a supportive niche for the *Lgr5*+ stem cells (Noah et al., 2011).

Notch pathway is involved in both the maintenance of *Lgr5*+ stem cells and proper differentiation of progenitors into mature cell types. At the crypt bottom, Paneth cells

express Notch ligands DLL1 and DLL4 and activate the Notch receptors, primarily NOTCH1 and NOTCH2 (Pellegrinet et al., 2011). In the upper crypt, secretory

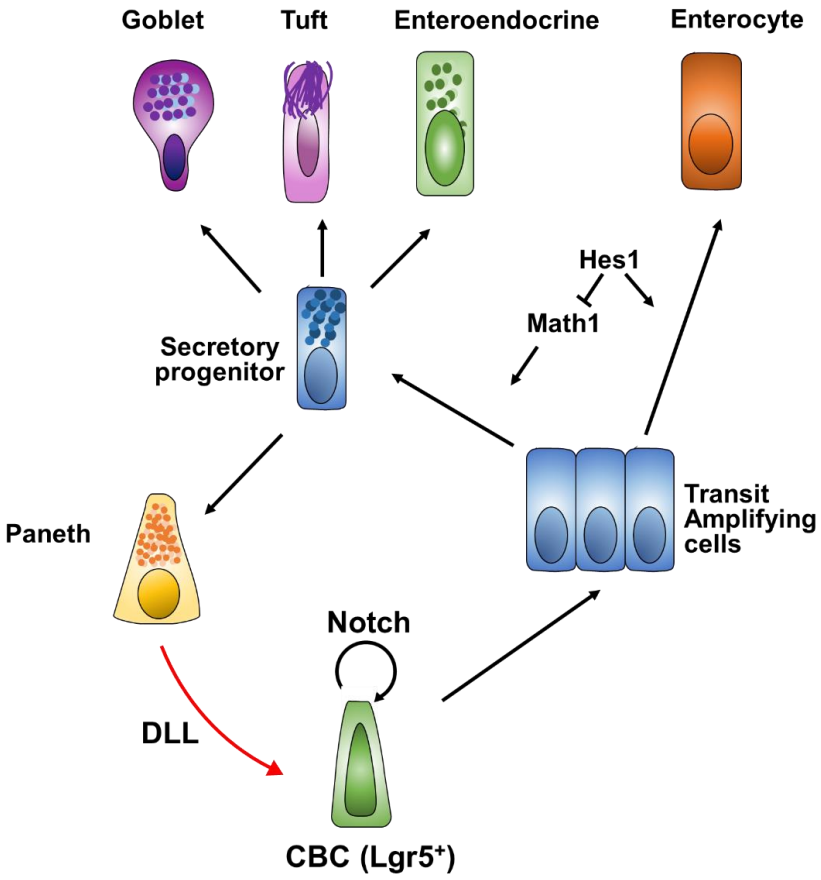


Figure 1.2 Schematic showing the role of Notch signalling in maintaining the intestinal epithelium.

progenitors and mature secretory cells express DLL1 and DLL4 to promote Notch activity in the neighbouring enterocyte progenitors and aid in their differentiation (Pellegrinet et al., 2011) (Figure 1.2, adapted from (Vooijs et al., 2011)).

Notch activity in the intestinal cells leads to expression of *Hes1* which represses the expression of Atonal homolog 1 (*Atoh1* or *Math1*) gene (Noah et al., 2011). *Atoh1* is essential for the maturation of secretory cells (Shroyer et al., 2007). In *Atoh1* null

mice, the *Lgr5*+ stem cell proliferation and the enterocyte differentiation does not appear to be hampered suggesting that the role of Notch signalling in the intestine is to repress *Atoh1* expression (Durand et al., 2012; Noah et al., 2011).

1.4 *KRAS* oncogene

In the second part of this dissertation, the discussion focusses on the transition of the epithelium from its homeostatic state to a neoplastic one. One of the hallmarks of cancer as described in (Hanahan and Weinberg, 2011) is non-reliance on external growth signals. Mitogen activated protein kinase (MAPK) pathway is one of the major intracellular signalling cascades that promote cellular proliferation (Katz et al., 2007). Extracellular signalling through growth factors like Epidermal Growth Factor (EGF) or Fibroblast Growth Factor (FGF) etc. is detected by plasma membrane bound receptors (EGFR or FGFR etc.) and gets transduced downstream through a cascade of kinases to MAPK (Katz et al., 2007). Activating mutations in this cascade can result in self-sufficiency in growth signals (Hanahan and Weinberg, 2011). Evidently, mutations in *EGFR*, *RAS* and *RAF* are frequently found in human cancers (Pylayeva-Gupta et al., 2011). There are three nearly identical RAS proteins in mammals: KRAS, HRAS and NRAS (Colicelli, 2004). RAS proteins exist in two states: active GTP bound one and an inactive GDP bound one. Guanine nucleotide exchange factors (GEFs) push them to the GTP bound state while the GTPase-activating proteins (GAP) bring them back to the inactive state (Pylayeva-Gupta et al., 2011). Oncogenic mutants are non-responsive to GAPs and continue to remain in the active state resulting in unrestrained growth

signalling (Pylayeva-Gupta et al., 2011). Among epithelial cancers, referred to as carcinomas, mutant KRAS is frequently found and hence it has been the focus of our study.

Mutant KRAS is found in about 97% of pancreatic ductal adenocarcinomas (PDAC), 45% of colorectal adenocarcinomas (COAD) and 30% of lung adenocarcinomas (LUAD) (Cox et al., 2014). Common mutations of KRAS include amino acid substitutions G12C, G12D, G12V, G13C, G13D, A18D, Q61H, K117N (Stolze et al., 2015). Expression of *Kras*^{G12D} in adult mouse pancreatic cells leads to the formation of pancreatic intra-epithelial neoplasia or PanIn lesions and not invasive adenocarcinomas (Habbe et al., 2008). Additional mutations such as loss of tumour suppressor *P53* are necessary for the progression of those lesions (Habbe et al., 2008). Expression of *Kras*^{G12D} in intestinal epithelial cells also does not lead to the formation of tumours (Snippert et al., 2014). In fact, KRAS mutations have been found in histologically normal human colon (Aivado et al., 2000; Zhu et al., 1997). However, expression of *Kras*^{G12D} in pulmonary epithelial cells leads to the formation of adenocarcinoma (Sutherland et al., 2014). Hence, we chose to study the effects of *Kras* oncogene in the context of lung adenocarcinoma.

1.5 Pulmonary epithelium

Lungs are the sites of gas exchange in mammals. The lining of the respiratory tract consisting of the nasal cavity, trachea, bronchi and bronchioles consists of ciliated pseudostratified epithelium and suddenly changes to simple squamous epithelium in the

alveoli of the lungs (Tata and Rajagopal, 2017). Lung adenocarcinoma is thought to arise from the cells in the alveoli or the terminal bronchioles (Sutherland et al., 2014). The major cell types found in the terminal bronchioles of mice are (Figure 1.3, adapted from (Rock and Hogan, 2011)):

1. Club cells

   Previously known as Clara cells, Club cells are secretory cells identified by the expression of *Scgb1a1* gene (Tata and Rajagopal, 2017). They can act as stem cells in the airway and can generate ciliated cells (Rawlins et al., 2009).

2. Ciliated cells

   They are identified by the expression of transcription factor *Foxj1*. And as the name implies, they contain cilia on their apical surface (Tata and Rajagopal, 2017).

3. Bronchioalveolar stem cells (BASC)

   These are special cells found at the bronchio-alveolar duct junction and have been shown to contribute to the formation of both bronchiolar and alveolar cell types (Kim et al., 2005). They can be identified by co-expression of Club cell marker *Scgb1a1* and alveolar type II (AEC2) cell marker *Sftpc* (Liu et al., 2019). Interestingly, they are not found in human lungs (Tata and Rajagopal, 2017).

The alveoli of lungs contain two major cell types:

1. Alveolar type I (AEC1) cells

   They are thin and elongated cells comprising of most of the alveolar space(Logan and Desai, 2015). They are in close proximity to endothelial cells and are involved

in gas exchange (Logan and Desai, 2015). AEC1 cells can be identified by their expression of *Hopx* gene (Logan and Desai, 2015).

2. Alveolar type II (AEC2) cells

They are cuboidal cells found interspersed between AEC1 cells (Logan and Desai, 2015). They have been shown to contribute to the regeneration of the alveoli and are considered to be the stem cells (Barkauskas et al., 2013). They are involved in the secretion of surfactants which help in maintaining the surface tension of the lung while breathing and can be identified by their expression of *Sftpc* gene (Tata

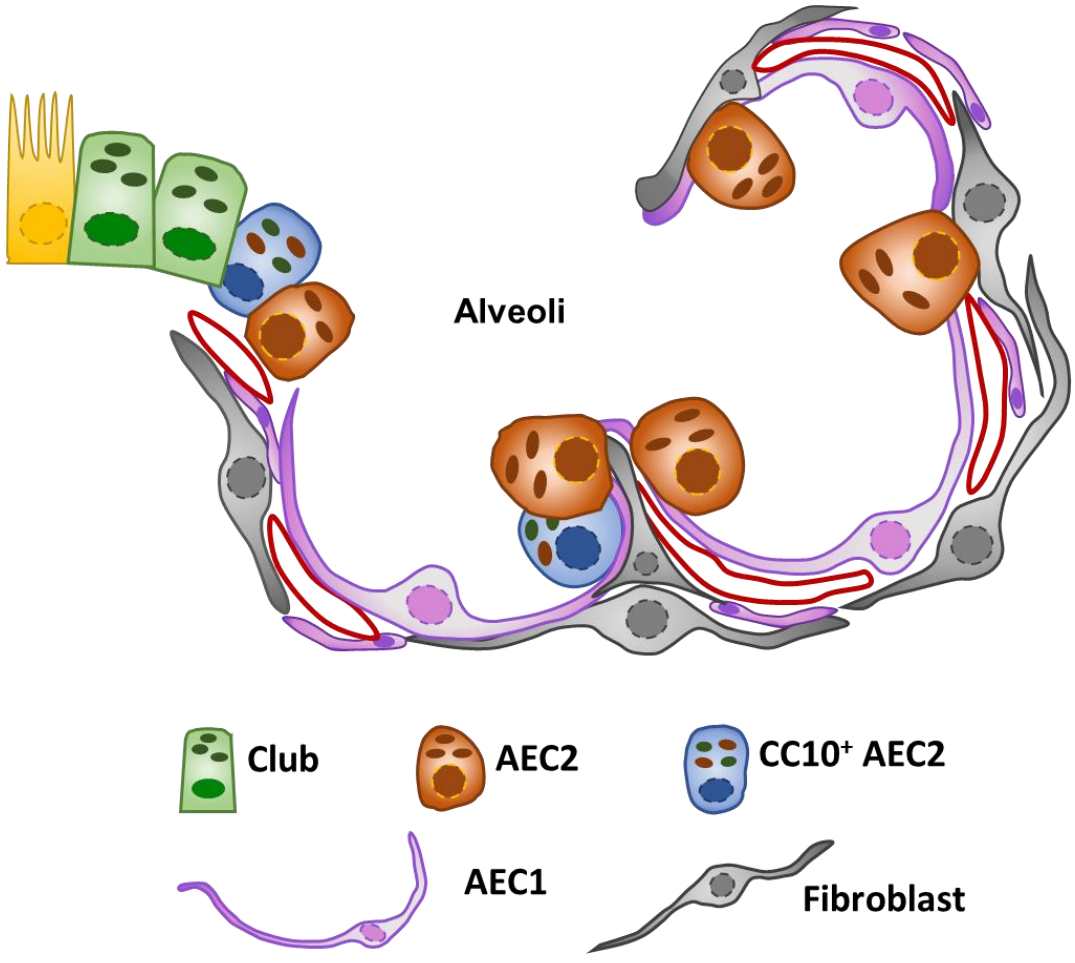


Figure 1.3 Schematic showing the organization of pulmonary alveoli.

and Rajagopal, 2017). Recent work has shown that a *Wnt* expressing fibroblast niche is important to maintain their stemness (Nabhan et al., 2018).

Cellular plasticity observed in the pulmonary epithelium makes it an interesting system to study. It has been shown that mature club cells of the airway can de-differentiate into basal cells (Tata et al., 2013). And in the alveoli, AEC1 can generate alveolar type II cells (Jain et al., 2015).

REFERENCES

Aivado, M., Gynes, M., Gorelov, V., Schmidt, W., Röher, H., and Goretzki, P. (2000). " Field cancerization"--an additional phenomenon in development of colon tumors? K-ras codon 12 mutations in normal colonic mucosa of patients with colorectal neoplasms. Der Chirurg; Zeitschrift fur alle Gebiete der operativen Medizen *71*, 1230-1234; discussion 1234-1235.

Allenspach, E.J., Maillard, I., Aster, J.C., and Pear, W.S. (2002). Notch Signaling in Cancer. Cancer Biology & Therapy *1*, 466-476.

Ayabe, T., Satchell, D.P., Wilson, C.L., Parks, W.C., Selsted, M.E., and Ouellette, A.J. (2000). Secretion of microbicidal α-defensins by intestinal Paneth cells in response to bacteria. Nature immunology *1*, 113.

Barkauskas, C.E., Cronce, M.J., Rackley, C.R., Bowie, E.J., Keene, D.R., Stripp, B.R., Randell, S.H., Noble, P.W., and Hogan, B. (2013). Type 2 alveolar cells are stem cells in adult lung. Journal of Clinical Investigation *123*, 3025-3036.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J.*, et al.* (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature *449*, 1003-1007.

Bray, S.J. (2006). Notch signalling: a simple pathway becomes complex. Nature reviews Molecular cell biology *7*, 678-689.

Colicelli, J. (2004). Human RAS superfamily proteins and related GTPases. Sci STKE *2004*, re13-re13.

Cox, A.D., Fesik, S.W., Kimmelman, A.C., Luo, J., and Der, C.J. (2014). Drugging the undruggable RAS: mission possible? Nature reviews Drug discovery *13*, 828.

D'souza, B., Miyamoto, A., and Weinmaster, G. (2008). The many facets of Notch ligands. Oncogene *27*, 5148.

Desbarats, J., and Newell, M.K. (2000). Fas engagement accelerates liver regeneration after partial hepatectomy. Nature medicine *6*, 920.

Durand, A., Donahue, B., Peignon, G., Letourneur, F., Cagnard, N., Slomianny, C., Perret, C., Shroyer, N.F., and Romagnolo, B. (2012). Functional intestinal stem cells after Paneth cell ablation induced by the loss of transcription factor Math1 (Atoh1). Proceedings of the National Academy of Sciences *109*, 8965-8970.

Gerbe, F., Legraverend, C., and Jay, P. (2012). The intestinal epithelium tuft cells: specification and function. Cellular and Molecular Life Sciences *69*, 2907-2917.

Greenman, C., Stephens, P., Smith, R., Dalgliesh, G.L., Hunter, C., Bignell, G., Davies, H., Teague, J., Butler, A., and Stevens, C. (2007). Patterns of somatic mutation in human cancer genomes. Nature *446*, 153.

Habbe, N., Shi, G., Meguid, R.A., Fendrich, V., Esni, F., Chen, H., Feldmann, G., Stoffers, D.A., Konieczny, S.F., Leach, S.D.*, et al.* (2008). Spontaneous induction of murine pancreatic intraepithelial neoplasia (mPanIN) by acinar cell targeting of oncogenic Kras in adult mice. Proceedings of the National Academy of Sciences *105*, 18913-18918.

Haines, N., and Irvine, K.D. (2003). Glycosylation regulates Notch signalling. Nature reviews Molecular cell biology *4*, 786-797.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. cell *144*, 646-674.

Harvey, B.M., Rana, N.A., Moss, H., Leonardi, J., Jafar-Nejad, H., and Haltiwanger, R.S. (2016). Mapping sites of O-glycosylation and fringe elongation on Drosophila Notch. Journal of Biological Chemistry *291*, 16348-16360.

Hicks, C., Johnston, S.H., diSibio, G., Collazo, A., Vogt, T.F., and Weinmaster, G. (2000). Fringe differentially modulates Jagged1 and Delta1 signalling through Notch1 and Notch2. Nature Cell Biology *2*, 515-520.

Iso, T., Kedes, L., and Hamamori, Y. (2003). HES and HERP families: Multiple effectors of the notch signaling pathway. Journal of Cellular Physiology *194*, 237-255.

Jain, R., Barkauskas, C.E., Takeda, N., Bowie, E.J., Aghajanian, H., Wang, Q., Padmanabhan, A., Manderfield, L.J., Gupta, M., and Li, D. (2015). Plasticity of Hopx+

type I alveolar cells to regenerate type II cells in the lung. Nature communications *6*, 6727.

Kabiri, Z., Greicius, G., Madan, B., Biechele, S., Zhong, Z., Zaribafzadeh, H., Aliyev, J., Wu, Y., Bunte, R., and Williams, B.O. (2014). Stroma provides an intestinal stem cell niche in the absence of epithelial Wnts. Development *141*, 2206-2215.

Kaelberer, M.M., Buchanan, K.L., Klein, M.E., Barth, B.B., Montoya, M.M., Shen, X., and Bohórquez, D.V. (2018). A gut-brain neural circuit for nutrient sensory transduction. Science *361*, eaat5236.

Katz, M., Amit, I., and Yarden, Y. (2007). Regulation of MAPKs by growth factors and receptor tyrosine kinases. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research *1773*, 1161-1176.

Kim, C.F., Jackson, E.L., Woolfenden, A.E., Lawrence, S., Babar, I., Vogel, S., Crowley, D., Bronson, R.T., and Jacks, T. (2005). Identification of Bronchioalveolar Stem Cells in Normal Lung and Lung Cancer. Cell *121*, 823-835.

Kim, T.-H., and Shivdasani, R.A. (2011). Notch signaling in stomach epithelial stem cell homeostasis. Journal of Experimental Medicine *208*, 677-688.

Kopan, R. (2002). Notch: a membrane-bound transcription factor. Journal of cell science *115*, 1095-1097.

Kotton, D.N., and Morrisey, E.E. (2014). Lung regeneration: mechanisms, applications and emerging stem cell populations. Nature medicine *20*, 822.

Latorre, R., Sternini, C., De Giorgio, R., and Greenwood-Van Meerveld, B. (2016). Enteroendocrine cells: a review of their role in brain–gut communication. Neurogastroenterology & Motility *28*, 620-630.

LeBon, L., Lee, T.V., Sprinzak, D., Jafar-Nejad, H., and Elowitz, M.B. (2014). Fringe proteins modulate Notch-ligand cis and trans interactions to specify signaling states. eLife *3*.

Liu, Q., Liu, K., Cui, G., Huang, X., Yao, S., Guo, W., Qin, Z., Li, Y., Yang, R., and Pu, W. (2019). Lung regeneration by multipotent stem cells residing at the bronchioalveolar-duct junction. Nature genetics, 1.

Logan, C.Y., and Desai, T.J. (2015). Keeping it together: Pulmonary alveoli are maintained by a hierarchy of cellular programs. BioEssays *37*, 1028-1037.

Miyamoto, S., and Rosenberg, D.W. (2011). Role of Notch signaling in colon homeostasis and carcinogenesis. Cancer science *102*, 1938-1942.

Morasso, M.I., and Tomic-Canic, M. (2005). Epidermal stem cells: the cradle of epidermal determination, differentiation and wound healing. Biology of the Cell *97*, 173-183.

Murthy, P.K.L., Srinivasan, T., Bochter, M.S., Xi, R., Varanko, A.K., Tung, K.-L., Semerci, F., Xu, K., Maletic-Savatic, M., and Cole, S.E. (2018). Radical and lunatic fringes modulate notch ligands to support mammalian intestinal homeostasis. eLife *7*, e35710.

Nabhan, A.N., Brownfield, D.G., Harbury, P.B., Krasnow, M.A., and Desai, T.J. (2018). Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. Science *359*, 1118-1123.

Noah, T.K., Donahue, B., and Shroyer, N.F. (2011). Intestinal development and differentiation. Experimental cell research *317*, 2702-2710.

Panin, V.M., Papayannopoulos, V., Wilson, R., and Irvine, K.D. (1997). Fringe modulates Notch–ligand interactions. Nature *387*, 908-912.

Pelaseyed, T., Bergström, J.H., Gustafsson, J.K., Ermund, A., Birchenough, G.M., Schütte, A., van der Post, S., Svensson, F., Rodríguez-Piñeiro, A.M., and Nyström, E.E. (2014). The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. Immunological reviews *260*, 8-20.

Pellegrinet, L., Rodilla, V., Liu, Z., Chen, S., Koch, U., Espinosa, L., Kaestner, K.H., Kopan, R., Lewis, J., and Radtke, F. (2011). Dll1- and dll4-mediated notch signaling are required for homeostasis of intestinal stem cells. Gastroenterology *140*, 1230.

Peterson, L.W., and Artis, D. (2014). Intestinal epithelial cells: regulators of barrier function and immune homeostasis. Nature Reviews Immunology *14*, 141.

Pylayeva-Gupta, Y., Grabocka, E., and Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web. Nature Reviews Cancer *11*, 761.

Rana, N.A., Nita-Lazar, A., Takeuchi, H., Kakuda, S., Luther, K.B., and Haltiwanger, R.S. (2011). O-Glucose Trisaccharide Is Present at High but Variable Stoichiometry at Multiple Sites on Mouse Notch1. Journal of Biological Chemistry *286*, 31623-31637.

Rawlins, E.L., and Hogan, B.L. (2006). Epithelial stem cells of the lung: privileged few or opportunities for many? Development *133*, 2455-2465.

Rawlins, E.L., Okubo, T., Xue, Y., Brass, D.M., Auten, R.L., Hasegawa, H., Wang, F., and Hogan, B.L. (2009). The role of Scgb1a1+ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. Cell stem cell *4*, 525-534.

Rock, J.R., and Hogan, B.L. (2011). Epithelial progenitor cells in lung development, maintenance, repair, and disease. Annual review of cell and developmental biology *27*, 493-512.

Sawaguchi, S., Varshney, S., Ogawa, M., Sakaidani, Y., Yagi, H., Takeshita, K., Murohara, T., Kato, K., Sundaram, S., and Stanley, P. (2017). O-GlcNAc on NOTCH1 EGF repeats regulates ligand-induced Notch signaling and vascular development in mammals. Elife *6*, e24419.

Shroyer, N.F., Helmrath, M.A., Wang, V., Antalffy, B., Henning, S.J., and Zoghbi, H.Y. (2007). Intestine-Specific Ablation of Mouse atonal homolog 1 (Math1) Reveals a Role in Cellular Homeostasis. Gastroenterology *132*, 2478-2488.

Siveke, J.T., Lubeseder–Martellato, C., Lee, M., Mazur, P.K., Nakhai, H., Radtke, F., and Schmid, R.M. (2008). Notch signaling is required for exocrine regeneration after acute pancreatitis. Gastroenterology *134*, 544-555. e543.

Snippert, H.J., Schepers, A.G., van Es, J.H., Simons, B.D., and Clevers, H. (2014). Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion. EMBO reports *15*, 62-69.

Snippert, H.J., van der Flier, L.G., Sato, T., van Es, J.H., van den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., van Rheenen, J., Simons, B.D.*, et al.* (2010). Intestinal Crypt Homeostasis Results from Neutral Competition between Symmetrically Dividing Lgr5 Stem Cells. Cell *143*, 134-144.

Stahl, M., Uemura, K., Ge, C., Shi, S., Tashima, Y., and Stanley, P. (2008). Roles of Pofut1 and O-fucose in mammalian Notch signaling. Journal of Biological Chemistry *283*, 13638-13651.

Stolze, B., Reinhart, S., Bulllinger, L., Fröhling, S., and Scholl, C. (2015). Comparative analysis of KRAS codon 12, 13, 18, 61, and 117 mutations using human MCF10A isogenic cell lines. Scientific Reports *5*, 8535.

Sutherland, K.D., Song, J.-Y., Kwon, M.C., Proost, N., Zevenhoven, J., and Berns, A. (2014). Multiple cells-of-origin of mutant K-Ras–induced mouse lung adenocarcinoma. Proceedings of the National Academy of Sciences *111*, 4952-4957.

Tata, P.R., Mou, H., Pardo-Saganta, A., Zhao, R., Prabhu, M., Law, B.M., Vinarsky, V., Cho, J.L., Breton, S., and Sahay, A. (2013). Dedifferentiation of committed epithelial cells into stem cells in vivo. Nature *503*, 218.

Tata, P.R., and Rajagopal, J. (2017). Plasticity in the lung: making and breaking cell identity. Development *144*, 755-766.

van der Flier, L.G., and Clevers, H. (2009). Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium. Annual Review of Physiology *71*, 241-260.

Vooijs, M., Liu, Z., and Kopan, R. (2011). Notch: architect, landscaper, and guardian of the intestine. Gastroenterology *141*, 448-459.

Watt, F.M., Estrach, S., and Ambler, C.A. (2008). Epidermal Notch signalling: differentiation, cancer and adhesion. Current opinion in cell biology *20*, 171-179.

Weinmaster, G., and Kintner, C. (2003). Modulation of notch signaling during somitogenesis. Annual review of cell and developmental biology *19*, 367-395.

Whitsett, J.A., and Alenghat, T. (2015). Respiratory epithelial cells orchestrate pulmonary innate immunity. Nature immunology *16*, 27.

Xu, K., and Egan, S.E. (2018). Out on the Fringe: Modulation of Notch Signaling by Glycosylation. In Targeting Notch in Cancer (Springer), pp. 87-126.

Zhu, D., Keohavong, P., Finkelstein, S.D., Swalsky, P., Bakker, A., Weissfeld, J., Srivastava, S., and Whiteside, T.L. (1997). K-ras gene mutations in normal colorectal

tissues from K-ras mutation-positive colorectal cancer patients. Cancer research *57*, 2485-2492.

Ziv, E., and Bendayan, M. (2000). Intestinal absorption of peptides through the enterocytes. Microscopy research and technique *49*, 346-352.

CHAPTER 2

RADICAL AND LUNATIC FRINGES MODULATE NOTCH LIGANDS TO SUPPORT MAMMALIAN INTESTINAL HOMEOSTASIS

Preetish Kadur Lakshminarasimha Murthy, Tara Srinivasan, Matthew S Bochter, Rui Xi, Anastasia Kristine Varanko, Kuei-Ling Tung, Fatih Semerci, Keli Xu, Mirjana Maletic-Savatic, Susan E Cole, Xiling Shen

Adapted from P. K. L. Murthy *et al.*, *eLife* **7**, e35710 (2018).

Author Contributions:

PKLM: Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing—original draft, Project administration, Writing—review and editing.

TS (contributed equally with PKLM): Data curation, Formal analysis, Investigation, Visualization, Writing—original draft, Writing—review and editing.

MSB: Resources. Maintained the mouse colonies and, harvested and processed intestinal samples.

RX: Formal analysis, Investigation, Visualization.

AKV and KLT: Investigation. Performed experiments to generate supporting data not shown in the manuscript.

FS, KX and MMS: Resources, Writing—review and editing.

SEC: Conceptualization, Resources, Writing—review and editing.

XS: Conceptualization, Supervision, Funding acquisition, Writing—review and editing.

INTRODUCTION

Lgr5+ Crypt Base Columnar cells (CBCs) located at the bottom of the crypts constantly self-renew to maintain the small intestinal epithelium, which is one of the fastest regenerative tissues in the body (Barker et al., 2007; van der Flier and Clevers, 2009). CBCs divide and move up the crypt into the progenitor or transit-amplifying zone where the cells rapidly proliferate and terminally differentiate into two major types: absorptive (enterocytes) and secretory (mainly Paneth and goblet cells). Enterocytes and goblet cells populate the villi while the Paneth cells move to the bottom of the crypt and provide a niche for the CBCs (van der Flier and Clevers, 2009).

Notch signalling pathway primarily consists of Notch receptors (NOTCH1-4) and ligands (DLL1-4 and JAG1-2) (Bray, 2006). Upon activation of a Notch receptor by a ligand, it undergoes successive cleavages by ADAM/TACE and γ-secretase (Bray, 2006). The cleaved intracellular domain (NICD) translocates to the nucleus leading to the transcription of multiple genes such as Hes and Hey families (Iso et al., 2003; Kopan, 2002). The extracellular domain of the Notch receptor and ligands contain EGF-like repeats, some of which serve as substrates for O-fucosylation by POFUT1 (Rampal et al., 2007; Wang et al., 2001). The fucosylated product may be further modified within the Golgi network by Fringe proteins: Lunatic (LFNG), Manic (MFNG) and Radical Fringe (RFNG) (Haines and Irvine, 2003; Moloney et al., 2000). Fringe proteins are typically expressed in receptor-expressing ‘receiver’ cells (Haines and Irvine, 2003). Glycosylation of NOTCH1 by LFNG and MFNG increases its activation by DLL1 but decreases its activation by JAG1 (Haines and Irvine, 2003;

Hicks et al., 2000; Panin et al., 1997). In contrast, glycosylation by RFNG increases the activation of NOTCH1 by both DLL1 and JAG1 (LeBon et al., 2014).

Notch pathway provides for spatial and context specific decision making in the intestinal epithelium. At the bottom of the crypt, Notch signalling is important for the maintenance of CBCs (Pellegrinet et al., 2011). In the upper crypt however, Notch activity, mainly through Hes1, is essential for the enterocyte differentiation (Fre et al., 2005; van der Flier and Clevers, 2009). Inhibition of Notch signalling results in the loss of proliferative CBCs and progenitor cells and leads to their differentiation into goblet cells in the upper crypt and villus, indicating the importance of the pathway in maintaining the tissue (Pellegrinet et al., 2011; Riccio et al., 2008; VanDussen et al., 2012; Wu et al., 2010). Notch1, 2 and Dll1, 4 are known to be the necessary receptors and ligands in the intestine (Pellegrinet et al., 2011; Riccio et al., 2008; Schröder and Gossler, 2002). Although, the fringe proteins are known to be expressed in the intestine, their function has not been studied (Schröder and Gossler, 2002).

Here we show that Rfng and Lfng are expressed by the ligand-presenting secretory lineages, but at different locations. At the crypt base, Rfng expressed in Paneth cells modulates DLL1 and DLL4, which enhances Notch signalling and self-renewal of neighbouring CBCs. In the upper crypt and villus, Lfng is expressed by secretory cells including enteroendocrine, Tuft and goblet cells. LFNG promotes Notch signalling in the transit amplifying cells and impedes their differentiation into secretory cells. MFNG does not play any noticeable role in intestinal epithelial homeostasis.

RESULTS

Rfng supports *Lgr5*+ stem cell self-renewal

Rfng transcripts have been detected in the crypt by in situ hybridization (Schröder and Gossler, 2002). We analysed previously published microarray data on *Lgr5*+ CBCs and Paneth cells (Sato et al., 2010) and found that Rfng is significantly upregulated in Paneth cells (Figure 1—figure supplement 1A). We isolated CBCs and Paneth cells ($CD24^{high}/SSC^{high}$) from *Lgr5*-GFP mice by FACS using an established protocol (Sato et al., 2010; Sato et al., 2009) and found that the Paneth cells are enriched for *Rfng* (Figure 1A). We validated that the isolated cells are indeed Paneth cells and CBCs by confirming their Lysozyme and GFP expression respectively (Figure 1—figure supplement 1B,C). We also confirmed that Rfng is enriched in the Paneth cells by RNA in situ hybridisation (ISH) (Figure 1B). We validated the specificity of ISH probes using Rfng null mouse intestinal sections (Figure 1—figure supplement 1D,E). We then established an in vitro knockdown (KD) model using organoid cultures of epithelial cells obtained from *Lgr5*-GFP mice. Single *Lgr5*-GFP+ CBCs were transduced with scrambled (Sc.) shRNA (control) or *Rfng* shRNA and propagated as organoids (Figure 1—figure supplement 1F). The colony formation efficiency of the Rfng KD CBCs was reduced compared to the control (Figure 1C). Flow cytometric analysis showed that the number of *Lgr5*+ CBCs decreased after *Rfng* loss, whereas the number of Paneth cells remained relatively unchanged (Figure 1D).

We confirmed the observation *in vivo* using previously published *Rfng* deficient (*Rfng*−/−) mice (Moran et al., 2009). Crypt cells were isolated from *Rfng*−/− and *Rfng*+/+ control mice and analysed by flow cytometry using a combination of surface

markers to identify CBCs (CD24loCD44+CD166+GRP78-) (Wang et al., 2013). Analysis revealed that the *Rfng*−/− mice had almost a two-fold depletion in CBCs (Figure 1E). A reduction of CBC marker *Lgr5* transcripts in the crypts harvested from *Rfng*−/− mouse intestines was observed by RT-qPCR measurement when compared to the control (Figure 1F). The number of Paneth and goblet cells remain largely unchanged and no other significant phenotype was detected in the epithelium (Figure 1—figure supplement 2A–F). Loss of *Rfng* in organoids seems to show a more significant phenotype in CBC reduction than its loss in vivo. This may be because CBCs *in vivo* also receive cues from the mesenchyme and not just the Paneth cells as in case of organoids.

To confirm that the loss of *Rfng* only in the Paneth cells can affect the CBCs, we performed the Organoid Reconstitution Assay (ORA) described previously (Rodríguez-Colman et al., 2017). FACS sorted Lgr5-GFP+ CBCs were incubated with Paneth cells from wild type or *Rfng* null mice for 10 min at room temperature and plated in Matrigel. We find that the colony formation ability of CBCs incubated with Paneth cells lacking *Rfng* was significantly lower than the control (Figure 1—figure supplement 3). It should be noted that not all CBCs associate with a Paneth cell during the incubation. Also, the *Lgr5*-GFP+ CBCs divide and give rise to Paneth cells with *Rfng*. Hence the result of this assay is not as significant as that shown in Figure 1C.

Rfng promotes Notch signalling in CBCs

We isolate, by FACS, the CBCs and Paneth cells from the *Rfng* KD and control seven days old organoids described earlier. Western blotting shows that Notch target genes *Hes1* and *Hey1* have reduced expression in the CBCs upon loss of *Rfng*.

However, the levels of ligands DLL1, four and JAG1 on Paneth cells were not significantly altered, consistent with the post-translational modifying role of Fringe (Figure 2A). *Rfng* KD and control CBCs were then transfected with an RBPJκ-dsRed reporter (Hansson et al., 2006) as an indicator of Notch activity, cultured overnight and analysed by flow cytometry (Figure 2B). *Rfng* shRNA decreased the mean RBPJκ-dsRed fluorescent intensity, indicating reduced overall Notch signalling in CBCs.

Fringes are known to modulate Notch signalling when present in receptor expressing cell (Haines and Irvine, 2003). But here we find *Rfng* in the ligand presenting cell promoting Notch activity in the neighbouring CBCs. We confirmed that the Paneth cells express the ligand *Dll1* by RNA-ISH (Figure 2—figure supplement 1). To understand the mechanism behind this, we examined ligand availability and concentration on the cell surface according to a previously established method using flow cytometry (LeBon et al., 2014; Taylor et al., 2014; Yang et al., 2005). Seven days old *Rfng* KD and control organoids were dissociated and single unpermeabilised cells were labelled with CD24 to mark Paneth cells and NOTCH1-Fc to quantify ligand binding to NOTCH1 (Figure 2C). Mean fluorescent intensity (MFI) of NOTCH1 binding to Paneth cells with *Rfng* knockdown was reduced compared to the scrambled control. We further confirmed that the ligands available on the Paneth cell surface have reduced by using specific antibodies for DLL1, DLL4 and JAG1. Flow cytometry showed that DLL1 and DLL4 levels on the Paneth cell surface reduced after the loss of *Rfng* (Figure 2D), although the total expression level of DLL1 and DLL4 within Paneth cells is not changed by *Rfng* knockdown after the cells were

permeabilised (Figure 2E). Western blotting also confirmed that the total DLL1 and DLL4 expression in Paneth cells does not change significantly after the loss of *Rfng* (Figure 2A). The outcomes of the ligand availability assays suggest that the DLLs available on the Paneth cell surface for NOTCH1 to bind to have been reduced after the loss of RFNG, which could decrease Notch activity in the adjacent CBCs.

Mfng plays an insignificant role

Mfng is expressed by scattered cells in the intestinal epithelium (Schröder and Gossler, 2002). To understand its potential function in maintaining the epithelium, we established an in vitro shRNA based *Mfng* knockdown model as before. Western blotting and RT-qPCR analysis validated *Mfng* knockdown (Figure 3A, B). Gene expression levels of *Lgr5* and Notch components were comparable between Sc. Control and *Mfng* KD organoids (Figure 3B). Additionally, the colony forming efficiency (Figure 3C) and the expression pattern of *Lgr5*-GFP (CBC) and MUC2 (goblet cells) (Figure 3D) of *Mfng* shRNA-expressing CBCs were similar to the scrambled control. We quantified this observation using flow cytometry, which confirmed no significant change in the number of *Lgr5*-GFP+ CBCs and goblet cells (Figure 3E, F). Finally, the percentage of differentiated cells, identified by CK20 expression, was not significantly altered between Sc. control and *Mfng* KD organoids (Figure 3G).

We then analysed intestinal tissues from *Mfng* deficient ($Mfng^{-/-}$) mice (Moran et al., 2009) (Figure 3—figure supplement 1A–C). IF microscopy showed similar MUC2 staining in intestinal sections of $Mfng^{-/-}$ and wild-type ($Mfng^{+/+}$) mice (Figure 3—figure supplement 1D). Quantification in intestinal tissues based on IF expression

indicated that the number of goblet cells was not significantly altered in *Mfng*$^{+/+}$mice compared to *Mfng*$^{-/-}$ mice. Finally, we examined the total number of CK20+ cells in intestines, which was similar in wild-type and *Mfng* null mice (Figure 3—figure supplement 2A).

We observed that goblet cells were slightly enriched in *Mfng* when compared to the CBCs (Figure 3—figure supplement 2B). We found no significant change in the cell surface expression of DLLs on goblet cells after the loss of *Mfng* (Figure 3—figure supplement 2C–F). Overall, these data suggest *Mfng* plays an insignificant role in intestinal tissues.

Lfng deletion leads to increased goblet cell differentiation

Lunatic Fringe is known to be expressed in the crypts and scattered cells in the villous epithelium (Schröder and Gossler, 2002). Immunofluorescence analysis of intestines from *Lfng*-GFP reporter mice confirmed that *Lfng* is expressed by a subset of cells in the upper crypt (transient-amplifying cell region) and villus (Figure 4A). We observe that *Lfng*-GFP+ cells are post-mitotic in the upper-crypt. (Figure 4A). Further analysis showed that these *Lfng*-GFP+ cells express *ChgA*, *Dclk1* or *Muc2* which are markers for enteroendocrine, Tuft or goblet cells respectively (Figure 4B–D). Secretory cells in the intestine, mainly enteroendocrine and goblet cells, are known to express Notch ligands, especially DLL1 (Van Es et al., 2012). In the upper crypt, immunofluorescence analysis showed Notch1 activity in the cells adjacent to *Lfng*-GFP+ cells but not in themselves (Figure 4E). We performed RT-qPCR measurement of *Lfng* using goblet cells and CBCs isolated from *Lgr5*-GFP mice and confirmed that *Lfng* is in goblet cells and not CBCs (Figure 4—figure supplement 1A).

We established an in vitro shRNA based *Lfng* knockdown model as before (Figure 4—figure supplement 1B). We observed only a slight decrease in colony forming efficiency of CBCs after *Lfng* knockdown and no significant change in the level of Notch activity in the CBCs (Figure 4—figure supplement 1C,D). However, we find that the number of goblet cells (MUC2+) increased after the loss of *Lfng* (Figure 4C). Quantification by flow cytometry showed that number of goblet cells (UEA-1+/CD24-) (Wong et al., 2012) was increased significantly in *Lfng* KD organoids (5.5% of the total population) when compared to the scrambled control (1.9%) (Figure 4D). Accordingly, the ratio of the number of goblet cells to *Lgr5*+ CBCs increased approximately three times in *Lfng* shRNA-expressing organoids (Figure 4E).

We confirmed these observations in vivo by examining the intestinal tissues from *Lfng* deficient ($Lfng^{-/-}$) mice (Moran et al., 2009). We observed an increase in the number of goblet cells in the *Lfng* null mice as expected (Figure 5A and Figure 5—figure supplement 1A, B). Goblet cells were quantified in villus crypt units (VCU) of the small intestine (Ishikawa et al., 1997). Immunofluorescence analysis based on MUC2 expression in small intestinal tissues from $Lfng^{-/-}$ mice showed an increase in the number of goblet cells when compared to the control ($Lfng^{+/+}$) mice (Figure 5B). Finally, using flow cytometry we quantified goblet cell numbers in $Lfng^{-/-}$mice: 14.1% of small intestinal cells, which is significantly higher than the 7.9% goblet cells in the small intestine of wild-type litter-mate control mice (Figure 5C). We observed no change in the Paneth cell numbers after loss of *Lfng* (Figure 5—figure supplement 1C).

Lfng deletion reduces Notch signalling

Suppression of Notch signalling is known to increase the goblet cell numbers (Van Es et al., 2012). We isolated and analysed intestinal progenitor cells ($CD24^{lo}CD44^{+}CD166^{+}GRP78^{+}$) from $Lfng^{+/+}$ and $Lfng^{-/-}$ mice using an established protocol (Wang et al., 2013) (Figure 5D). RT-qPCR measurements indicated decreased *Hes1* and increased *Atoh1* (transcriptional factor essential for generating secretory lineage (Shroyer et al., 2007) expression in $Lfng^{-/-}$progenitors compared to the control. We also confirmed reduced activated Notch1 (NICD) in the upper crypts of *Lfng* null mice intestines (Figure 5—figure supplement 2A, B).

Therefore, *Lfng* silencing appears to lower Notch activity in the progenitors and promotes the secretory lineage leading to an increase in goblet cell numbers (Kim and Shivdasani, 2011b; Zheng et al., 2011).

Secreted LFNG plays no apparent function

Previous reports have indicated that *Lfng* may be secreted into the extracellular space (Shifley and Cole, 2008; Williams et al., 2016). We first examined the medium from intestinal organoid cultures derived from *Lgr5*-GFP mice using solid-phase ELISA. Secreted *Lfng* was detected at a concentration of approximately 315–325 ng/mL using two independent LFNG primary antibodies (Figure 6A and Figure 6—figure supplement 1A). The other Fringes, RFNG and MFNG, were not detected in the culture medium (Figure 6—figure supplement 1B, C). We tried to understand if secreted LFNG influences Notch signalling by affecting receptors in a non-cell autonomous manner. As before, single *Lgr5*-GFP + CBCs were transduced with *Lfng* shRNA and propagated as organoids followed by incubation with conditioned medium harvested from wild-type organoids that contained soluble form of secreted LFNG

(sLFNG). After 24 hr, organoid cultures were analysed using flow cytometry, which showed that the percentage of goblet cells remained similar to the *Lfng* knockdown (shRNA) condition and significantly higher than scrambled shRNA-expressing organoids which express endogenous LFNG (Figure 6B). RT-qPCR revealed that the expression levels of Notch ligands DLL1 and DLL4 were similar between *Lfng* knockdown with and without soluble LFNG (Figure 6—figure supplement 1D).

We then examined intestinal tissues from mutant *Lfng* mice ($Lfng^{RLFNG/+}$ or *RLfng*) in which the N-terminal sequence of LFNG, which normally allows for protein processing and secretion, is replaced with the N-terminus of Radical fringe (a Golgi resident protein) (Williams et al., 2016) (Figure 6—figure supplement 1E). IF analysis based on MUC2 expression in small intestinal tissues from *RLfng* mice showed similar goblet cell numbers in villus crypt units compared to wild-type ($Lfng^{+/+}$) mice (Figure 6C). Taken together, our in vitro and in vivo findings suggest that the effect of LFNG on goblet cell numbers and intestinal homeostasis is not dependent on its secreted form.

LFNG promotes DLL expression on the cell surface

In order to understand if LFNG, like RFNG, can affect the cell surface expression of DLL, we examined ligand availability and concentration on the cell surface. Seven days old *Lfng* KD and control organoids were dissociated and single unpermeabilised cells were labelled with CD24 and UEA-1 to mark goblet cells and NOTCH1-Fc to quantify ligand binding to NOTCH1 (Figure 7A). Mean fluorescent intensity of NOTCH1 binding to goblet cells with *Lfng* knockdown was reduced compared to the

control, suggesting that the ligands available on the goblet cell surface for NOTCH1 to bind to were reduced after the loss of *Lfng*. Flow cytometry shows that DLL1 and DLL4 levels, detected using specific antibodies, on the goblet cell surface reduced after the loss of *Lfng* (Figure 7B), although the total expression of DLL1 and DLL4 by the goblet cells measured after the permeabilising the cells remained almost the same (Figure 7C). Western blotting also confirmed that the total DLL1 and DLL4 expression in goblet cells does not change significantly after the loss of *Lfng* (Figure 4—figure supplement 1D).

DISCUSSION

We report that *Rfng* is enriched in the Paneth cells and promotes cell surface expression of DLL1 and DLL4. This promotes Notch activity in the neighbouring *Lgr5*+ CBCs assisting their self-renewal. *Mfng* does not appear to contribute significantly in maintaining the epithelium. *Lfng* on the other hand is expressed by enteroendocrine, Tuft and goblet cells and suppresses the secretory lineage (Figure 8). Even though Fringe proteins do not appear to be essential, they provide another layer of spatial and lineage-specific modulation that might enhance robustness of intestinal homeostasis. This is consistent with the highly robust regulation of Notch activity in the intestinal epithelium as inhibition of *Notch1* or *Dll1* only causes minor defective phenotype, while inhibition of *Notch2*, *Dll4*, *Jag1*, *Hes1*, *Hes3* or *Hes5* causes no significant phenotype (Pellegrinet et al., 2011; Ueo et al., 2012; Wu et al., 2010).

We have observed that both RFNG and LFNG can increase the presence of DLL1 and DLL4 on the plasma membrane. This can potentially contribute to the increase in cis-inhibition of NOTCH1 by DLL1 in the presence of fringe (LeBon et al., 2014). Fringe modulation of ligands will be of significance in understanding Notch activity in cancer stem cell asymmetric division where LFNG, DLL1 and NOTCH1 are present in the same cell (Bu et al., 2013; Bu et al., 2016). However, the mechanism behind the increase in cell-surface expression of the ligands still needs to be understood. The glycosylation state of proteins has been known to affect their intracellular trafficking (Huet et al., 2003; Ohtsubo and Marth, 2006). It raises the possibility that fringe mediated glycosylation or the addition of Galactose and Sialic acid post fringe activity

might affect the trafficking of DLLs to the cell surface. *Lfng* in the *Dll1* expressing cell, in the presence of *Dll3*, is known to reduce Notch activity in the neighbouring cell (Okubo et al., 2012). This raises the possibility that ligands interact with each other in the presence of *Lfng* which might explain our observation. *In vitro* reductionist studies may need to be conducted in systems expressing single ligand and fringe to understand the mechanism in detail. Also, our experiments cannot completely rule out the possibility that low levels of *Rfng* expression in CBCs (in comparison to Paneth cells) can also contribute to the phenotype by directly modulating Notch receptors. We have observed some mesenchymal cells also express *Rfng* detectable by RNA ISH. We also observe that some of the mesenchymal cells also express *Dll1* (Figure 2—figure supplement 1). Further studies are necessary to map the expression of all the Notch ligands in different mesenchymal cell types. This raises the possibility that the mesenchyme can also provide Notch ligands to the *Lgr5* + CBCs *in vivo*. In case that is true, our proposed mechanism that *Rfng* promotes cell surface expression of *Dll1* might be applicable to the mesenchymal cells too. Upon loss *Rfng*, reduced *Dll1* expression on the cell surface of Paneth cells and mesenchymal cells would result in reduced Notch activity in the CBCs, as observed. However, the crypt cells are separated from the mesenchyme by the basement membrane. The efficacy of DLL mediated Notch signalling across the intestinal basement membrane needs to be explored.

Although we have observed that the *Lfng* expressing cells are found both in the upper crypt and in the villus, our data suggests that LFNG in NICD- post-mitotic secretory cells of the upper crypt promotes Notch activity in the neighbouring enterocyte

progenitors. As Notch signalling is not active in the villus, the *Lfng*+ cells of the villi likely do not impact epithelial cell differentiation. It would be interesting to explore the reason secretory cells expressing Notch ligands and *Lfng* are found in the villi. The differences, other than functional consequence, between the *Lfng*+ cells of the upper crypt and villus needs to be explored.

Notch pathway is a potential therapeutic target, but blocking the pathway leads to serious GI related side effects (Van Es et al., 2012). Targeting the Notch pathway through fringe appears to be a potentially viable strategy to exclusively modulate intestinal epithelial regeneration or its functions, absorption and mucus secretion, as Notch activity in the stem cells or progenitors can be specifically targeted by blocking *Rfng* or *Lfng* respectively.

MATERIALS AND METHODS

Mice

Lgr5-GFP (Jackson Lab #8875, RRID:IMSR_JAX:008875) strain has been described previously (Sato et al., 2009). *Lfng* null ($Lfng^{tm1Rjo}$), *Mfng* null ($Mfng^{tm1Seco}$, RRID:MGI:3849430) and *Rfng* null ($Rfng^{tm1Tfv}$) mice were maintained as described here (Moran et al., 2009; Ryan et al., 2008). $Lfng^{RLFNG/+}$ mice were maintained as previously described (Williams et al., 2016). Littermates of Fringe mutants with wild-type gene expression were used as controls. *Lfng*-GFP (GENSAT # RRID:MMRRC_015881-UCD) were received as FVB/N - C57BL/6 hybrids and crossed to C57BL/6 mice for at least 10 generations (Gong et al., 2003; Maletic-Savatic et al., 2017). All procedures were conducted under protocols approved by the appropriate Institutional Animal Care and Use Committees at Ohio State University (# 2012A00000036-R1), Duke University (# A286-15-11), Baylor College of Medicine (# AN-5004), Cornell University (# 2010–0100) or Research Animal Resource Center of Weill Cornell Medical College (# 2009–0029).

Organoid culture and flow cytometry

Organoids from *Lgr5*-GFP mouse intestines were cultured as previously described with minor modifications (Sato et al., 2010; Sato et al., 2009). Briefly, small intestines were harvested, washed with PBS and opened up longitudinally to expose luminal surface. A glass coverslip was then gently applied to scrape off villi and the tissue was cut into 2–3 mm fragments and incubated with 2 mM EDTA for 45 min on a rocking platform at 4°C. EDTA solution was then decanted and replaced with cold PBS. The tissues were vigorously agitated to release the crypts. Next, single cell dissociation

was achieved by incubating purified crypts at 37°C with Trypsin-EDTA solution containing 0.8KU/ml DNase, 10 μM Y-27632 for 30 min. To isolate *Lgr5*-GFP+ cells, single cells were resuspended in cold PBS with 0.5% BSA and GFP^{high} cells were sorted by FACS (Beckman Coulter/BD FACSAria).

Dissociated cells were also stained with anti-CD24 antibody and UEA-1. Paneth cells were sorted based on side scatter and CD24 expression ($CD24^{high}/SSC^{high}$) and goblet cells were identified as $UEA\text{-}1^{+}/CD24^{-}$ (Sato et al., 2010; Wong et al., 2012). Viable cells were gated based on 7-AAD or Sytox blue staining. Data analysis was performed using FlowJo software.

Single *Lgr5*-GFP+ CBCs were plated in Matrigel and cultured in medium containing: Advanced DMEM/F12 supplemented with Glutamax, 10 mM HEPES, N2, B27 without vitamin A, 1 μM N-acetylcysteine, 50 ng/mL EGF, 100 ng/mL Noggin, and 10% R-SPONDIN1 conditioned medium.

Lentiviral constructs containing *Lfng* shRNA (sc-39491-SH), *Mfng* shRNA (sc-39493-SH), *Rfng* shRNA (sc-39495-SH), or scrambled shRNA (sc-108060) were purchased from Santa Cruz Biotechnology. Lentiviral transduction of *Lgr5*-GFP CBCs were performed by 'spinoculation' method described previously (Koo et al., 2012). Transduced CBCs were cultured as organoids and analysed after 7 days. RBPJκ-dsRed reporter (Addgene #47683) was transfected into single Sc. shRNA-expressing or *Rfng* shRNA-expressing sorted *Lgr5*-GFP CBCs using Lipofectamine-2000 as described earlier (Schwank et al., 2013).

Organoid Reconstitution Assay was performed as described previously (Rodríguez-Colman et al., 2017). Briefly, FACS sorted Paneth cells and *Lgr5*-GFP+ CBCs were

mixed, spun down and incubated at room temperature for 10 min. The pellet was then plated in Matrigel.

RT-qPCR and protein analysis

A Qiagen RNeasy kit was used to extract total RNA. RT-PCR primers from Genecopoeia were used for the following genes: β-*Actin*, *Lgr5*, *Lfng*, *Mfng* and *Rfng*. Taqman primers (ABI) were used for: *Lgr5*, *Notch1*, *Hes1*, *Hes5*, *Dll1*, *Dll4*, and *Jag1*. *Gapdh* or β-*Actin* was used as internal control. Protein isolation and western blotting were performed as previously described, using β-ACTIN for normalisation (Pan et al., 2008). ELISA kits for LNFG, RFNG, and MFNG were purchased from MyBioSource and assays were performed according to the manufacturer's instructions similar to the following protocol. Solid-phase ELISA assays were independently conducted using LFNG, RFNG, and MFNG antibodies (referred to as antibody-2) purchased from Santa Cruz Biotechnology for verification of results obtained from the corresponding kits.

Ligand availability assay

Ligand availability assays were performed as previously described (LeBon et al., 2014). Briefly, blocking buffer (PBS, 2% FBS, 100 μg/mL $CaCl_2$) and binding Buffer (PBS, 2% BSA, 100 μg/mL $CaCl_2$) were prepared. Subsequently, cells were incubated in blocking buffer for 30 min at 37°C followed by incubation with 0.5 μg/mL NOTCH1-Fc (R and D #5267) diluted in binding buffer for 1 hr at 4°C. Cells were then washed three times in blocking buffer and incubated in secondary antibody diluted in binding buffer for 40 min at room temperature. Finally, cells were washed three times in blocking buffer and analysed by flow cytometry.

Immunofluorescence (IF) and immunohistochemistry (IHC)

Sections of paraffin embedded tissues were deparaffinised using Xylene and rehydrated. Antigen retrieval was performed using Proteinase K (Dako) or 10 mM Tris buffer at pH9. The sections were incubated in Protein Block (Dako) for 10 min at room temperature (RT). Primary antibodies diluted in Antibody Diluent (Dako) were added and incubated overnight at 4°C. Slides were then washed in PBS and incubated in secondary antibodies diluted in Antibody Diluent for 1 hr at RT and washed in PBS. The slides were then mounted using Vectashield mounting medium containing DAPI. Intestinal sections were stained with Haematoxylin and Eosin (H and E), Periodic Acid-Schiff (PAS), Alcian Blue (AB) or Nuclear Fast Red according to standard methods. Intestinal organoids embedded in Matrigel were fixed with 3% PFA for 15 min at room temperature and permeabilised using 0.2% Triton-X for IF according to the protocol described above. Antibodies used are listed in supplementary methods. Antibodies used are listed in Supplementary file 1.

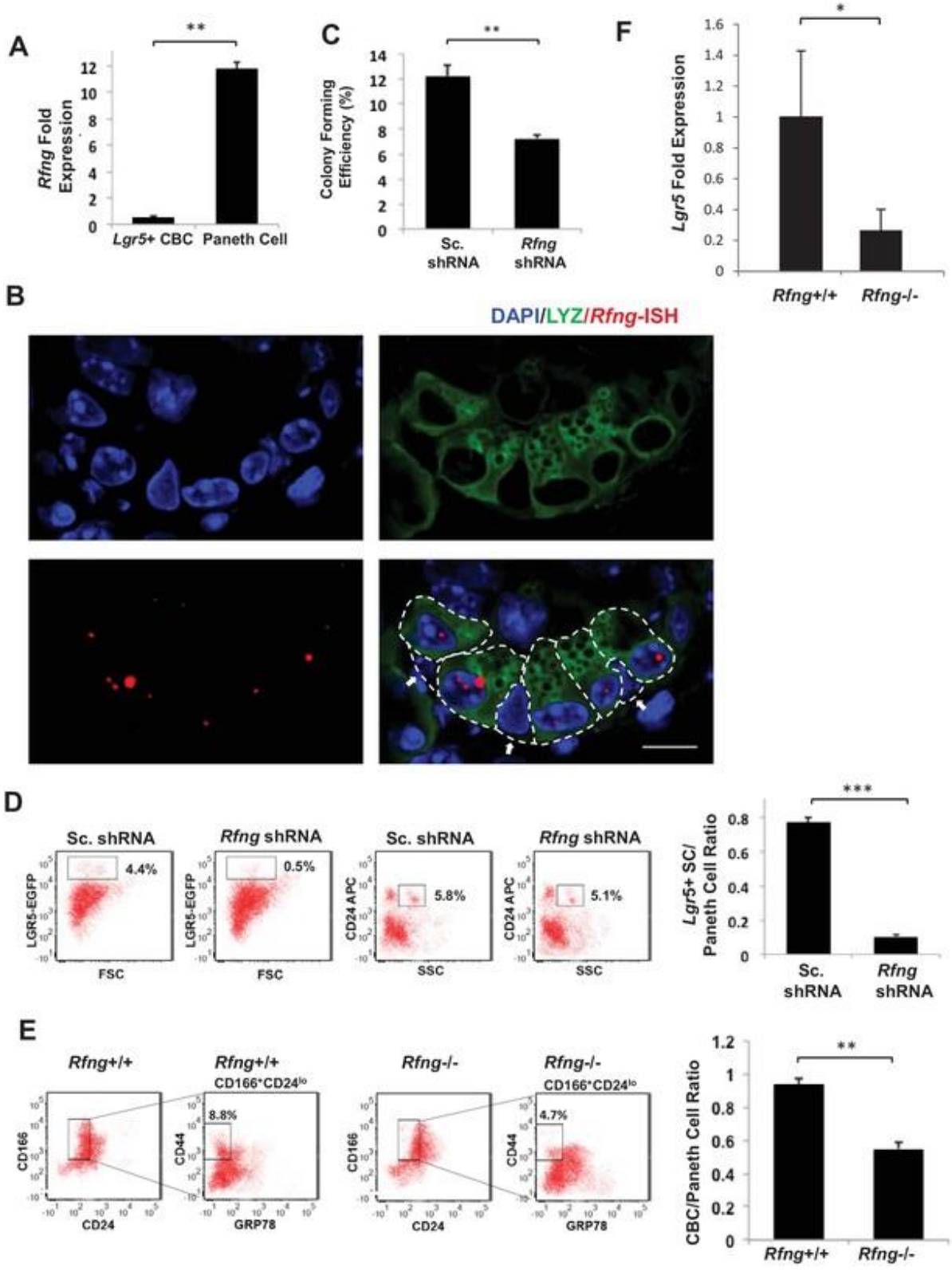
Protocol was modified while staining for Notch1 intracellular domain (NICD). Antigen retrieval was performed using Trilogy (Cell Marque). Sections were then incubated in 3% hydrogen peroxide diluted in PBS for 10 min. Protein blocking, primary antibody and secondary antibody incubation were performed as described above. Signal was further amplified using TSA Plus Fluorescein kit (Perkin Elmer). To quantify, NICD+ nuclei and total number of nuclei (based on DAPI signal) were counted in each crypt (35 to 50 crypts from each section) to obtain the fraction of NICD+ nuclei.

0.5 mg EdU in 150 µl PBS (~16.66 mg/kg) was injected intraperitoneally into *Lfng*-GFP mice two hours prior to euthanasia (Kabiri et al., 2014). Incorporated EdU was detected using Click-It EdU imaging kit (Thermo Fisher #C10640).

Statistical analysis

The data is represented as mean ± S.E.M (standard error of mean) unless otherwise indicated. A Student t-test was applied to compare two groups using $p<0.05$ to establish statistical significance.

Figure 1


A
**
Rfng Fold Expression
12
10
8
6
4
2
0
Lgr5+ CBC
Paneth Cell
C
**
Colony Forming Efficiency (%)
14
12
10
8
6
4
2
0
Sc. shRNA
Rfng shRNA
F
*
Lgr5 Fold Expression
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Rfng+/+
Rfng-/-
B
DAPI/LYZ/Rfng-ISH
D
Sc. shRNA
Rfng shRNA
Sc. shRNA
Rfng shRNA
4.4%
0.5%
5.8%
5.1%
LGR5-EGFP
FSC
CD24 APC
SSC
***
Lgr5+ SC/ Paneth Cell Ratio
0.8
0.6
0.4
0.2
0
Sc. shRNA
Rfng shRNA
E
Rfng+/+
Rfng+/+
CD166+CD24lo
Rfng-/-
Rfng-/-
CD166+CD24lo
8.8%
4.7%
CD166
CD24
CD44
GRP78
**
CBC/Paneth Cell Ratio
1.2
1
0.8
0.6
0.4
0.2
0
Rfng+/+
Rfng-/-

Rfng supports Lgr5+ stem cell self-renewal.

(**A**) RT-qPCR quantification of *Rfng* in *Lgr5*+ CBC and Paneth cells isolated from *Lgr5*-GFP mouse intestines. The experiment was performed in triplicate and presented as mean ± s.d. (standard deviation)

(**B**) Representative image showing *Rfng* transcripts (red) and Lysozyme protein (green) expression at the bottom of the crypt of *Lgr5*-GFP mouse intestine. DAPI (Blue) labels the nuclei and scale bar represents 10 μm. Arrows point to CBCs.

(**C–D**) Single *Lgr5*-GFP CBCs were transduced with either Sc. shRNA or *Rfng* shRNA. The experiment was performed in triplicate.

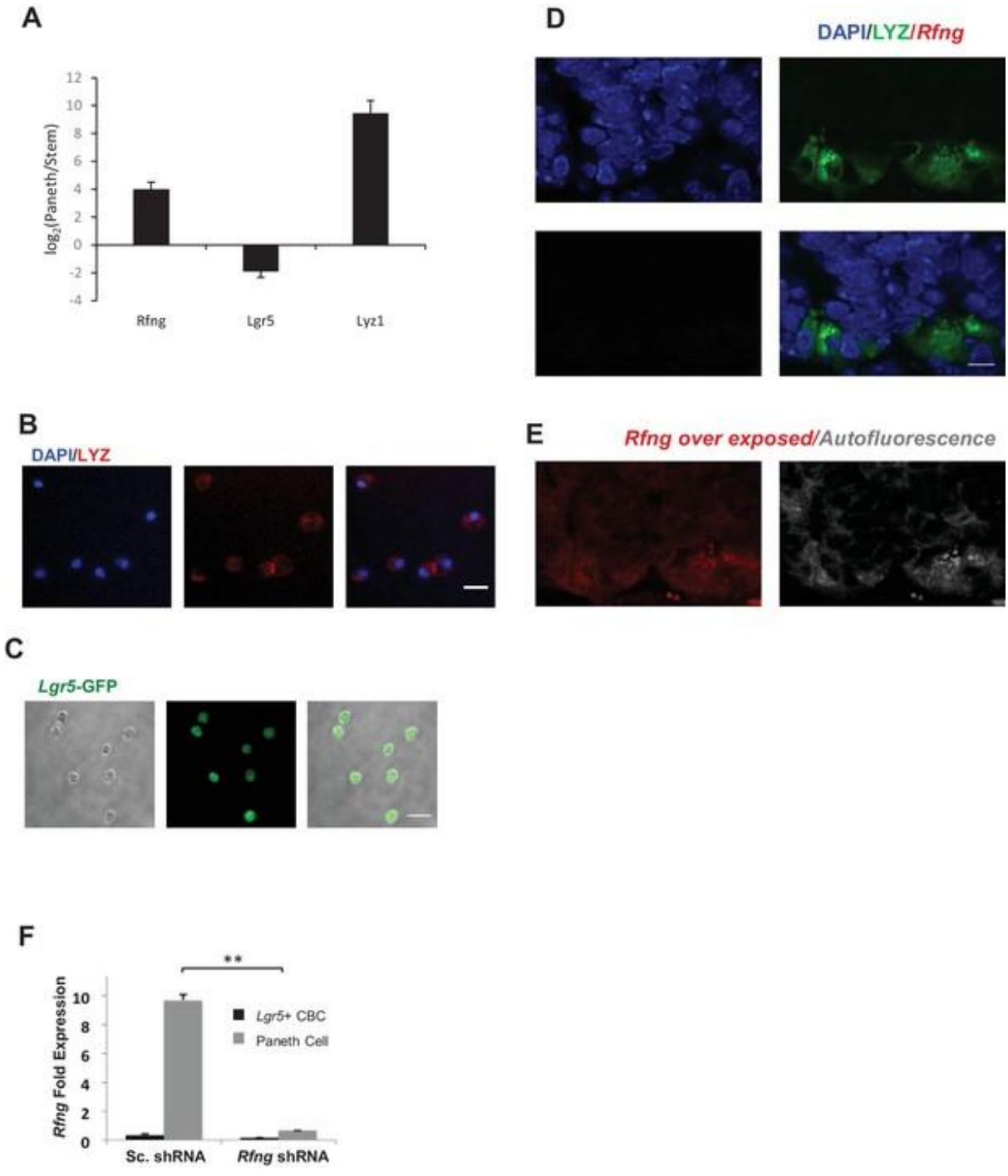
(**C**) Colony forming efficiency measured after 7 days. Quantitative analysis calculated from 1000 cells/replicate presented as mean ± s.d.

(**D**) Left: Representative flow cytometry plots indicating gated percentage of *Lgr5*+ (GFP^{high}) and Paneth cells ($CD24^{high}/SSC^{high}$). Right: Ratio of *Lgr5*-GFP+ CBCs/Paneth cells as determined by flow cytometry and presented as mean ± s.d.

(**E**) Left: Representative plots indicating gated population of CBCs ($CD166^{+}CD24^{lo}CD44^{+}GRP78^{-}$) from the intestine of $Rfng^{+/+}$ and $Rfng^{-/-}$ mice. Percentage reflects fraction of total population. Right: Ratio of number of CBCs to Paneth cells of n = 3 mice and presented as mean ± s.d.

(**F**) RT-qPCR quantification of *Lgr5* in crypts extracted from $Rfng^{+/+}$ and $Rfng^{-/-}$ mice. n = 3 mice. Data is presented as mean ± s.d. (*p<0.05; **p<0.01; ***p<0.001).

Figure 1 – Figure Supplement 1



Paneth Cells express *Rfng*.

(**A**) Ratio of gene expression in *Lgr5*+ CBC and Paneth cells from two independent microarrays published previously (Sato et al., 2011). Data presented as mean ± s.d.

(**B**) FACS Sorted Paneth cells were plated in Matrigel and stained for Lysozyme (red). DAPI (Blue) shows the nuclei and scale bar represents 10 μm.

(**C**) Sorted *Lgr5*-GFP cells were imaged. Endogenous GFP expression is shown in green. Scale bar represents 20 μm.

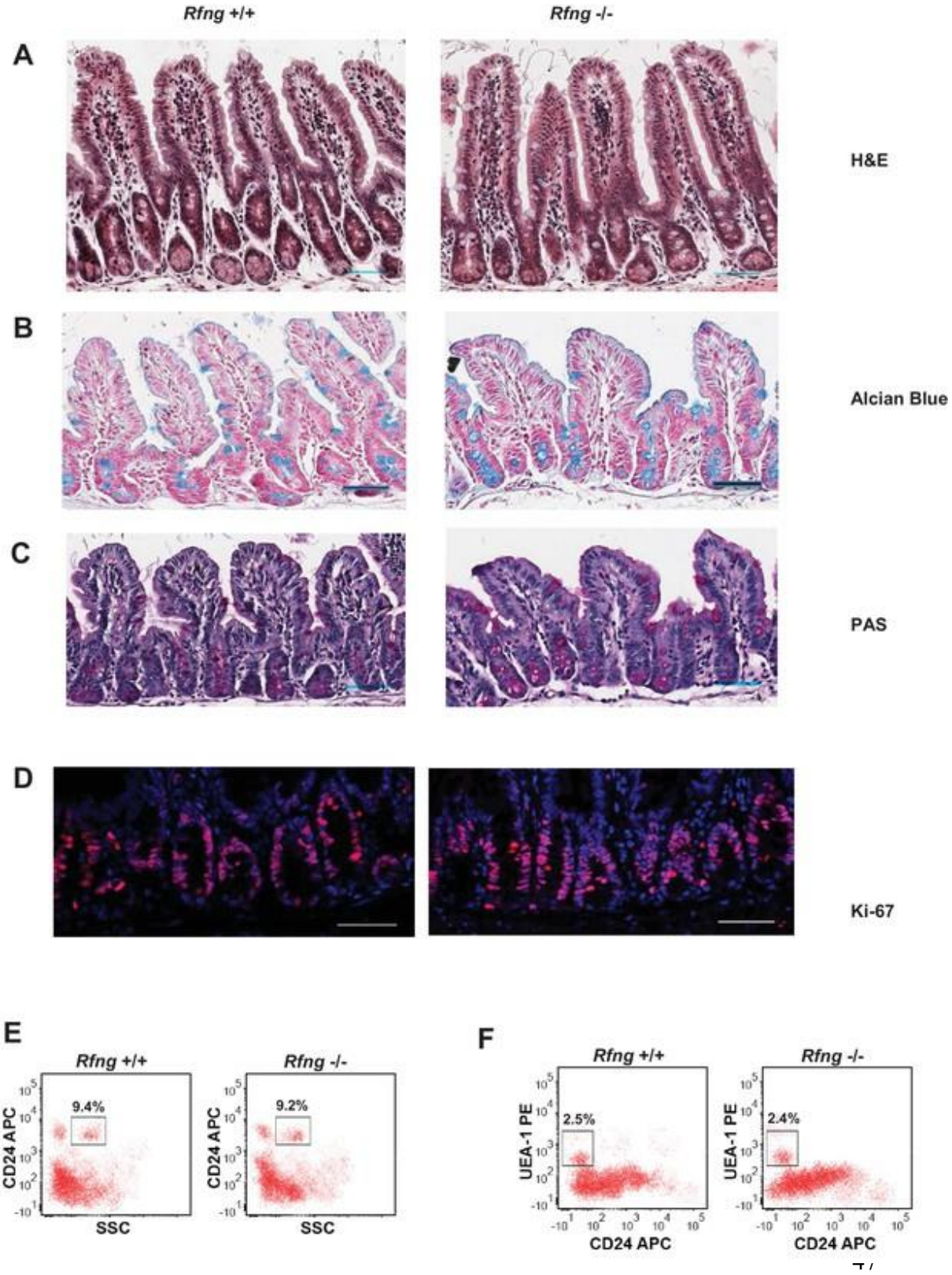
(**D–E**) Representative image from RNAscope assay for *Rfng* on *Rfng*$^{-/-}$ small intestinal section. DAPI (blue) shows nuclei and Lysozyme protein (green) marks the Paneth cells. Scale bar represents 10 μm.

(**D**) No signal is seen in the red channel (*Rfng* channel) confirming the specificity of *Rfng* probes.

(**E**) The region was then significantly overexposed to show the background signal. An additional image was taken by over exposing the region in the far-red channel (shown in grey; no probe/antibody present in this channel) to show tissue auto-fluorescence.

(**F**) *Lgr5* + CBC and Paneth cell populations were isolated from organoids that were infected with scrambled or *Rfng* shRNA. The experiment was performed in triplicate. RT-qPCR quantification of *Rfng* presented as mean ± s.d. in CBC and Paneth cells. (**p<0.01).

Figure 1 – Figure Supplement 2



Histological and flow cytometric analysis of *Rfng* null intestines.

(**A–D**) Representative images from the small intestine of *Rfng*$^{+/+}$ and *Rfng*$^{-/-}$ mouse strains stained using

(**A**) Haematoxylin and Eosin (H and E)

(**B**) Alcian Blue and Nuclear Fast Red

(**C**) Periodic Acid-Schiff (PAS) and Haematoxylin and

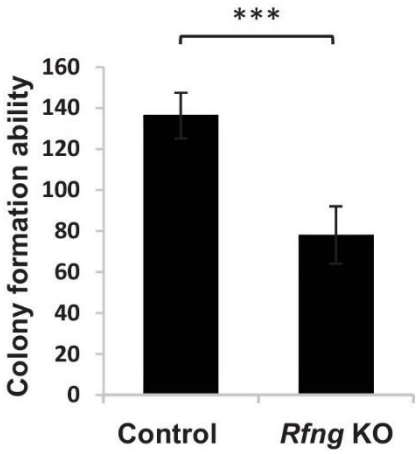
(**D**) Ki-67 (red) and DAPI (blue). Scale bar represents 50 μm.

(**E–F**) Representative flow cytometry plots indicating gated percentage of Paneth cells

(**E**) or goblet cells

(**F**) from the small intestine of *Rfng*$^{+/+}$ and *Rfng*$^{-/-}$ mice.

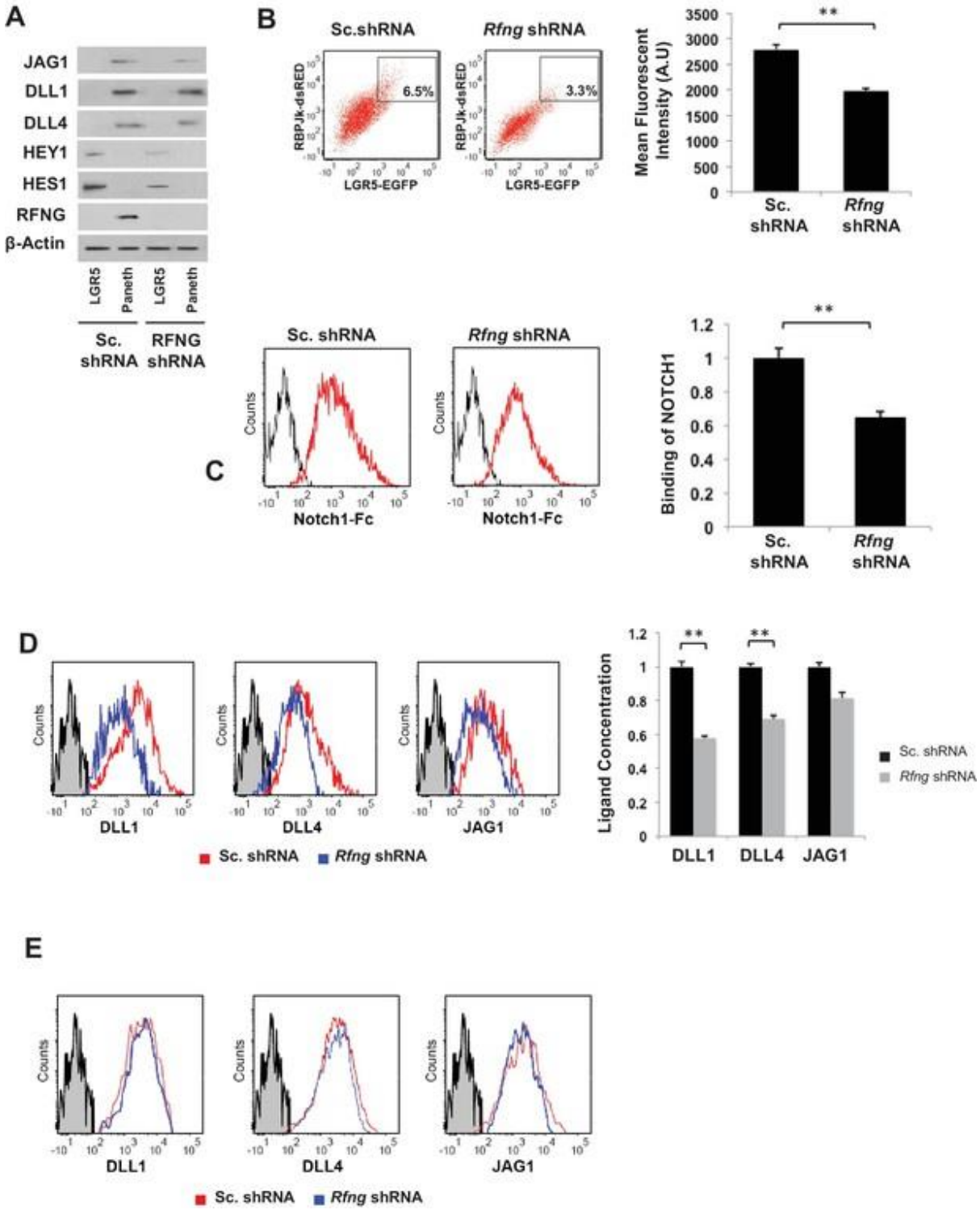
Figure 1 – Figure Supplement 3



Colony formation ability of *Lgr5*+ CBCs when mixed with Paneth cells from control or *Rfng* null mice.

n = 4 replicates with 8000 CBCs per replicate. Data is presented as mean ± s.d. (***p<0.001).

Figure 2


A
JAG1
DLL1
DLL4
HEY1
HES1
RFNG
β-Actin
LGR5
Paneth
LGR5
Paneth
Sc. shRNA
RFNG shRNA
B
Sc.shRNA
Rfng shRNA
RBPJk-dsRED
LGR5-EGFP
6.5%
3.3%
Mean Fluorescent Intensity (A.U)
3500
3000
2500
2000
1500
1000
500
0
**
Sc. shRNA
Rfng shRNA
C
Sc. shRNA
Rfng shRNA
Counts
Notch1-Fc
Binding of NOTCH1
1.2
1
0.8
0.6
0.4
0.2
0
**
Sc. shRNA
Rfng shRNA
D
Counts
DLL1
DLL4
JAG1
Sc. shRNA
Rfng shRNA
Ligand Concentration
1.2
1
0.8
0.6
0.4
0.2
0
**
**
DLL1
DLL4
JAG1
Sc. shRNA
Rfng shRNA
E
Counts
DLL1
DLL4
JAG1
Sc. shRNA
Rfng shRNA

***Rfng* promotes Notch signaling in *Lgr5*+ CBC.**

(**A**) Western blot analysis of Notch signaling components in CBCs and Paneth cells FACS sorted from *Rfng* KD and control organoids.

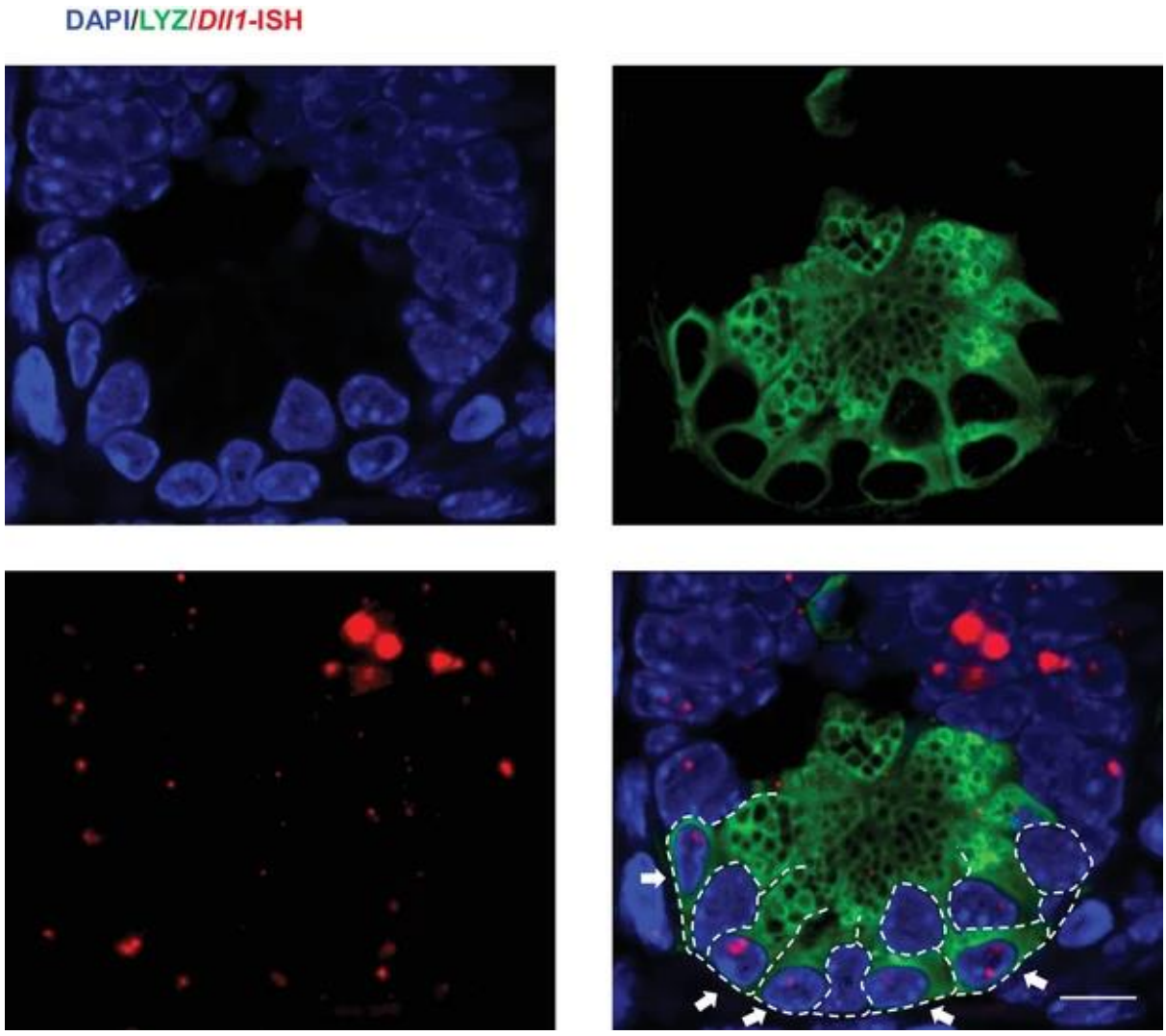
(**B**) Left: Representative plots for RBPJκ-dsRed and *Lgr5*-GFP expression indicating a gated double positive fraction for *Rfng* KD and control CBCs transfected with RBPJκ-dsRed reporter. Right: Mean fluorescence intensity (MFI) of RBPJκ-dsRed expression. The experiment was performed in triplicate and presented as mean ± s.d.

(**C**) Ligand availability on *Rfng* KD and control Paneth cells. Representative traces (left) and MFI (right) showing ligand binding to NOTCH1 measured by flow cytometry. Unstained Paneth cells were used as a negative control. The experiment was performed in triplicate and presented as mean ± s.d.

(**D**) Cell surface DLL1, DLL4, and JAG1 concentration on *Rfng* KD and control unpermeabilised Paneth cells. Left: Representative traces measured by flow cytometry. Right: MFI measurements. The experiment was performed in triplicate and presented as mean ± s.d.

(**E**) Cell surface DLL1, DLL4, and JAG1 concentration on *Rfng* KD and control permeabilised Paneth cells. (**$p<0.01$).

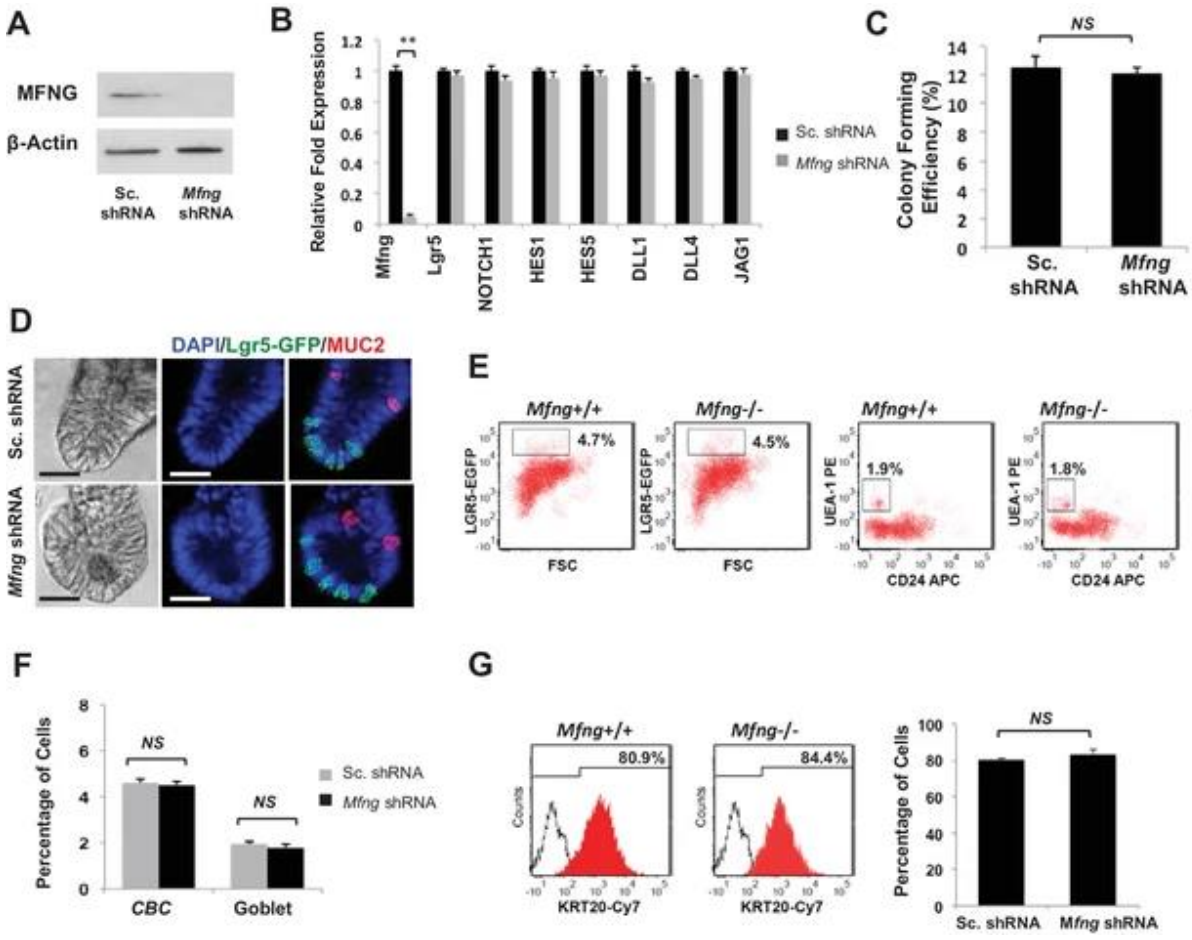
Figure 2 – Figure Supplement 1



***Dll1* expression in the crypts.**

Representative image showing *Dll1* transcripts (red) and Lysozyme protein (green) expression at the bottom of the crypt of *Lgr5*-GFP mouse intestine.DAPI (Blue) labels the nuclei and scale bar represents 10 μm. Arrows point to Paneth cells.

Figure 3


A
MFNG
β-Actin
Sc. shRNA
Mfng shRNA
B
Relative Fold Expression
**
Mfng
Lgr5
NOTCH1
HES1
HES5
DLL1
DLL4
JAG1
Sc. shRNA
Mfng shRNA
C
NS
Colony Forming Efficiency (%)
Sc. shRNA
Mfng shRNA
D
DAPI/Lgr5-GFP/MUC2
Sc. shRNA
Mfng shRNA
E
Mfng+/+
Mfng-/-
4.7%
4.5%
1.9%
1.8%
LGR5-EGFP
FSC
UEA-1 PE
CD24 APC
F
Percentage of Cells
NS
CBC
Goblet
Sc. shRNA
Mfng shRNA
G
Mfng+/+
Mfng-/-
80.9%
84.4%
Counts
KRT20-Cy7
Percentage of Cells
NS
Sc. shRNA
Mfng shRNA

***Mfng* plays an insignificant role.**

Single *Lgr5*-GFP CBCs were transduced with either Sc. shRNA or *Mfng* shRNA. The experiment was performed in triplicate.

(**A**) Western blot for *Mfng* expression.

(**B**) RT-qPCR quantification of *Mfng* and Notch components in organoids.

(**C**) Colony forming efficiency measured after 7 days. Quantitative analysis from 1000 cells/replicate.

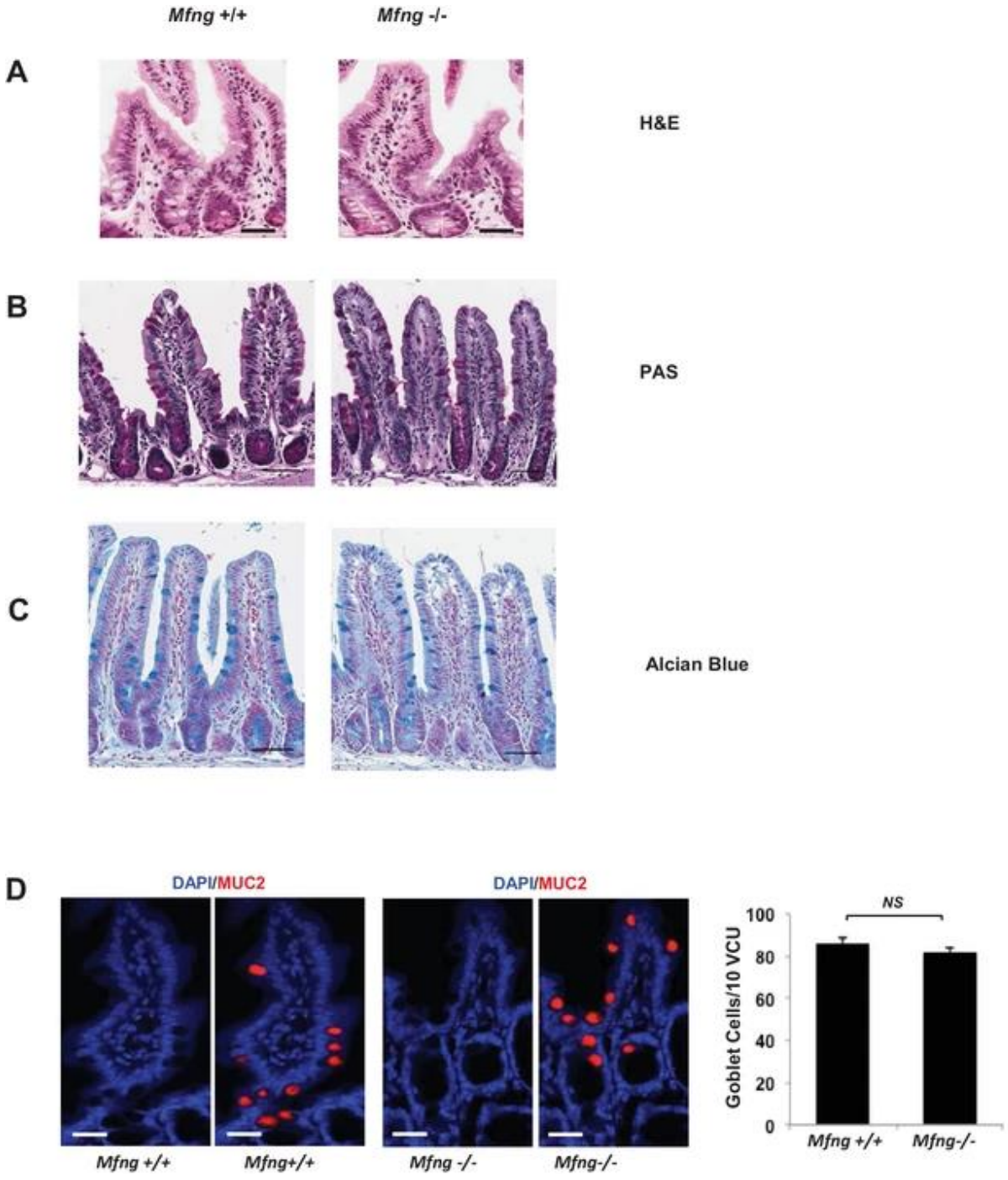
(**D**) Representative bright field and co-IF images indicating *Lgr5*-GFP (green) expression. MUC2 (red) marks Goblet cells. DAPI (blue) labels nuclei and scale bar represents 25 μm.

(**E**) Representative flow cytometry plots indicating gated percentage of *Lgr5*+ CBCs (GFP^{high}) and goblet cells ($UEA\text{-}1^{+}/CD24^{-}$).

(**F**) Percentage of *Lgr5*+ CBCs and goblet cells as determined by flow cytometry and presented as mean ± s.d.

(**G**) Left: Representative flow cytometry histograms indicating KRT20+ (CK20+) cells. Right: Percentage of KRT20+ cells and presented as mean ± s.d.

Figure 3 – Figure Supplement 1



Histological analysis of *Mfng* null intestines.

(**A–C**) Representative images from the small intestine of $Mfng^{+/+}$ and $Mfng^{-/-}$ mouse strains stained using

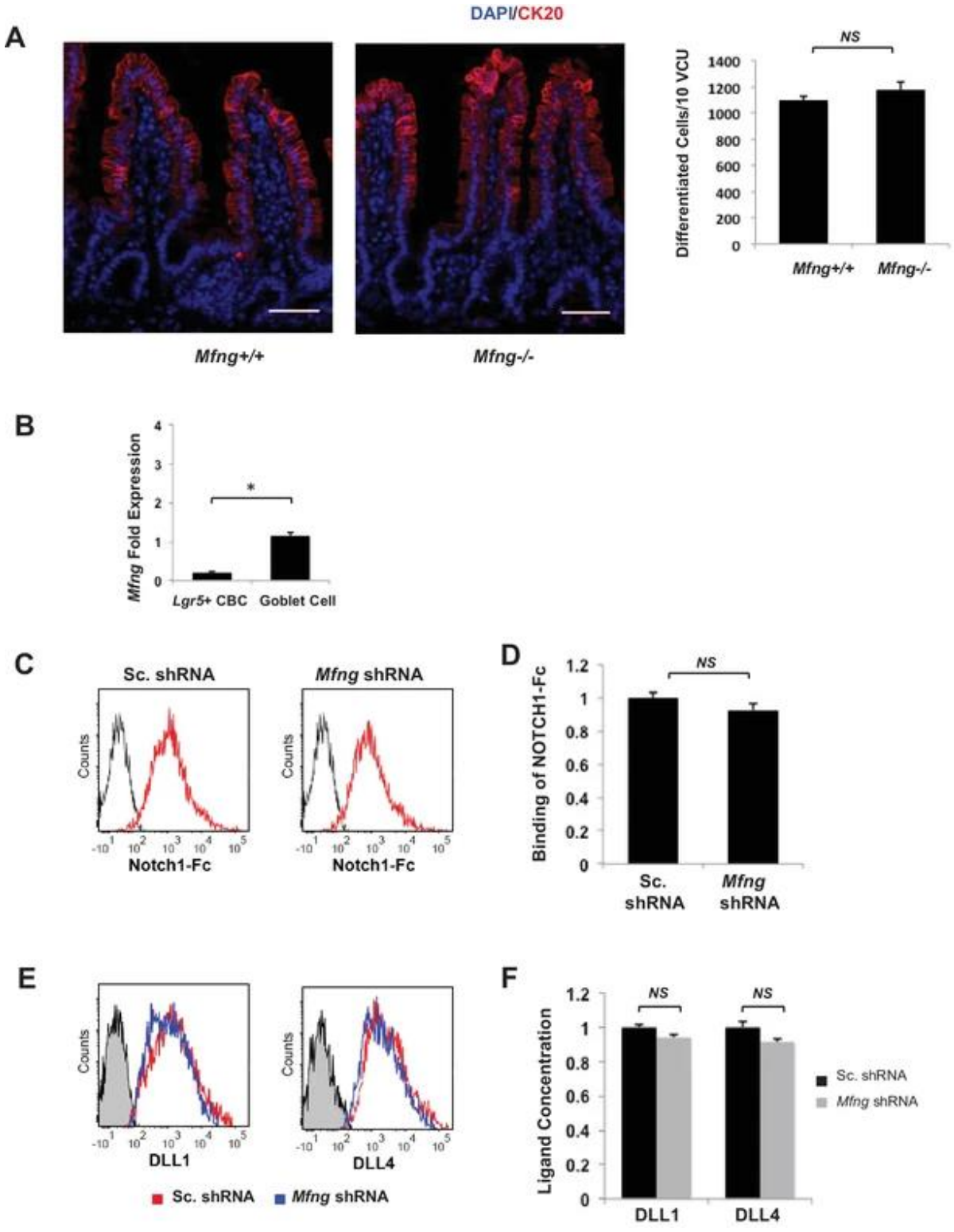
(**A**) Haematoxylin and Eosin (H and E)

(**B**) Periodic Acid Schiff (PAS) and Haematoxylin

(**C**) Alcian Blue and Nuclear Fast Red. Scale bar represents 50 μm.

(**D**) Left: Representative IF images of villus-crypt units (VCU) from the small intestine of $Mfng^{+/+}$ and $Mfng^{-/-}$ mouse strains. MUC2 (red) marks goblet cells. DAPI (blue) labels nuclei. Right: Quantification of the number of goblet cells of n = 4 mice/condition. Data represents mean ± s.d of n = 500 VCU per mouse. (NS (not significant)).

Figure 3 – Figure Supplement 2



No significant phenotype detected upon loss of Mfng.

(**A**) Left: Representative IF images of villus-crypt units (VCU) from the small intestine of *Mfng*$^{+/+}$ and *Mfng*$^{-/-}$ mouse strains. CK20 (red); DAPI (blue) labels nuclei; scale bar represents 50 μm. Right: Quantification of CK20+ differentiated cells in VCU of *Mfng*$^{+/+}$ and *Mfng*$^{-/-}$ mice based on IF. The data represents 500 VCU/mouse of n = 4 mice/condition and is presented as mean ± s.d. (NS) (not significant); Student t-test).

(**B**) RT-qPCR quantification of *Mfng* in *Lgr5*+ CBC and goblet cell populations from small intestinal tissue of *Lgr5*-GFP mice. The experiment was performed in triplicate and presented mean ± s.d.

(**C**) Single *Lgr5*-GFP CBCs were transduced with either Sc. shRNA or *Mfng* shRNA and propagated as organoids for seven days. Isolated goblet cells were incubated with 0.5 μg/ml NOTCH1-Fc. Shown are representative traces indicating ligand binding to Notch1 measured by flow cytometry in goblet cells. Unstained goblet cells were used as a negative control.

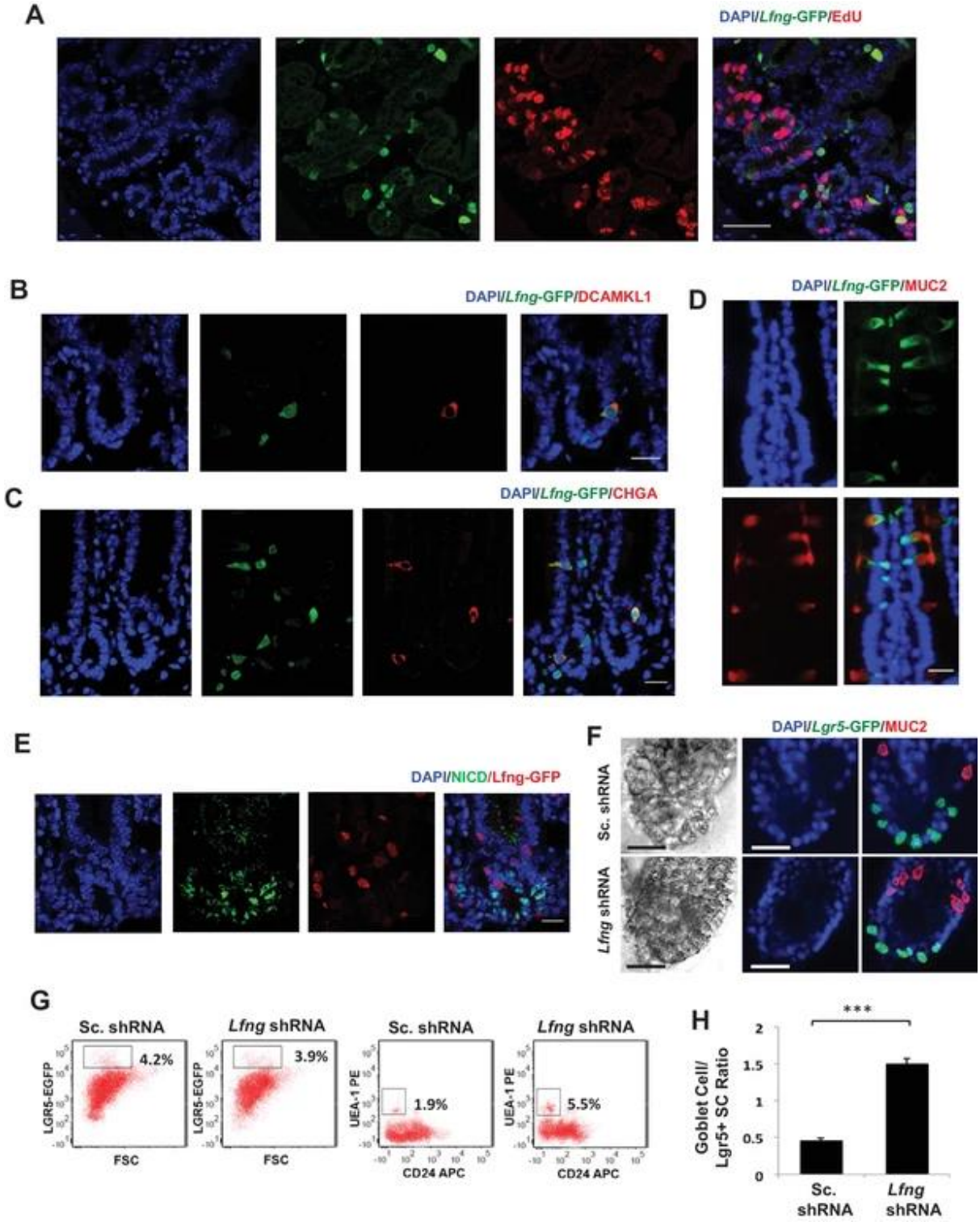
(**D**) Ligand binding data to NOTCH1-Fc in *Mfng* shRNA-expressing goblet cells normalized to Sc. shRNA-transfected goblet cells using flow cytometry mean fluorescent intensity (MFI) measurements. The experiment was performed in triplicate and presented mean ± s.d. (***NS***).

(**E**) Single *Lgr5*-GFP CBCs were transduced with either Sc. shRNA or *Mfng* shRNA and propagated as organoids for 7 days. Isolated goblet cells were incubated with antibodies directed against DLL1 and DLL4 ligands. Shown are representative traces

showing surface ligand concentration measured by flow cytometry on goblet cells. Unstained goblet cells were used as a negative control.

(**F**) Surface DLL1 and DLL4 ligand concentration on *Mfng* shRNA-expressing goblet cells normalized to Sc. shRNA-transfected goblet cells using flow cytometry mean fluorescent intensity (MFI) measurements. The experiment was performed in triplicate and presented mean ± s.d. (NS). (*$p<0.05$; **$p<0.01$).

Figure 4


A
DAPI/Lfng-GFP/EdU
B
DAPI/Lfng-GFP/DCAMKL1
C
DAPI/Lfng-GFP/CHGA
D
DAPI/Lfng-GFP/MUC2
E
DAPI/NICD/Lfng-GFP
F
DAPI/Lgr5-GFP/MUC2
Sc. shRNA
Lfng shRNA
G
Sc. shRNA
Lfng shRNA
Sc. shRNA
Lfng shRNA
4.2%
3.9%
1.9%
5.5%
LGR5-EGFP
FSC
UEA-1 PE
CD24 APC
H
***
Goblet Cell/ Lgr5+ SC Ratio
2
1.5
1
0.5
0
Sc. shRNA
Lfng shRNA

***Lfng* loss results in increased goblet cell differentiation in vitro.**

(**A–E**) Representative IF images of the small intestine of *Lfng*-GFP reporter mice. (**A**) GFP (green) shows the *Lfng*expression and EdU (red) marks the proliferating cells. DAPI (blue) labels nuclei. Scale bar represents 50 μm.

(**B**) GFP (green) shows the *Lfng* expression and DCAMKL1 (red) marks the Tuft cells. Scale bar represents 20 μm.

(**C**) GFP (green) shows the *Lfng* expression and CHGA (red) marks the enteroendocrine cells. Scale bar represents 20 μm.

(**D**) GFP (green) shows the *Lfng* expression and MUC2 (red) marks the goblet cells. Scale bar represents 20 μm.

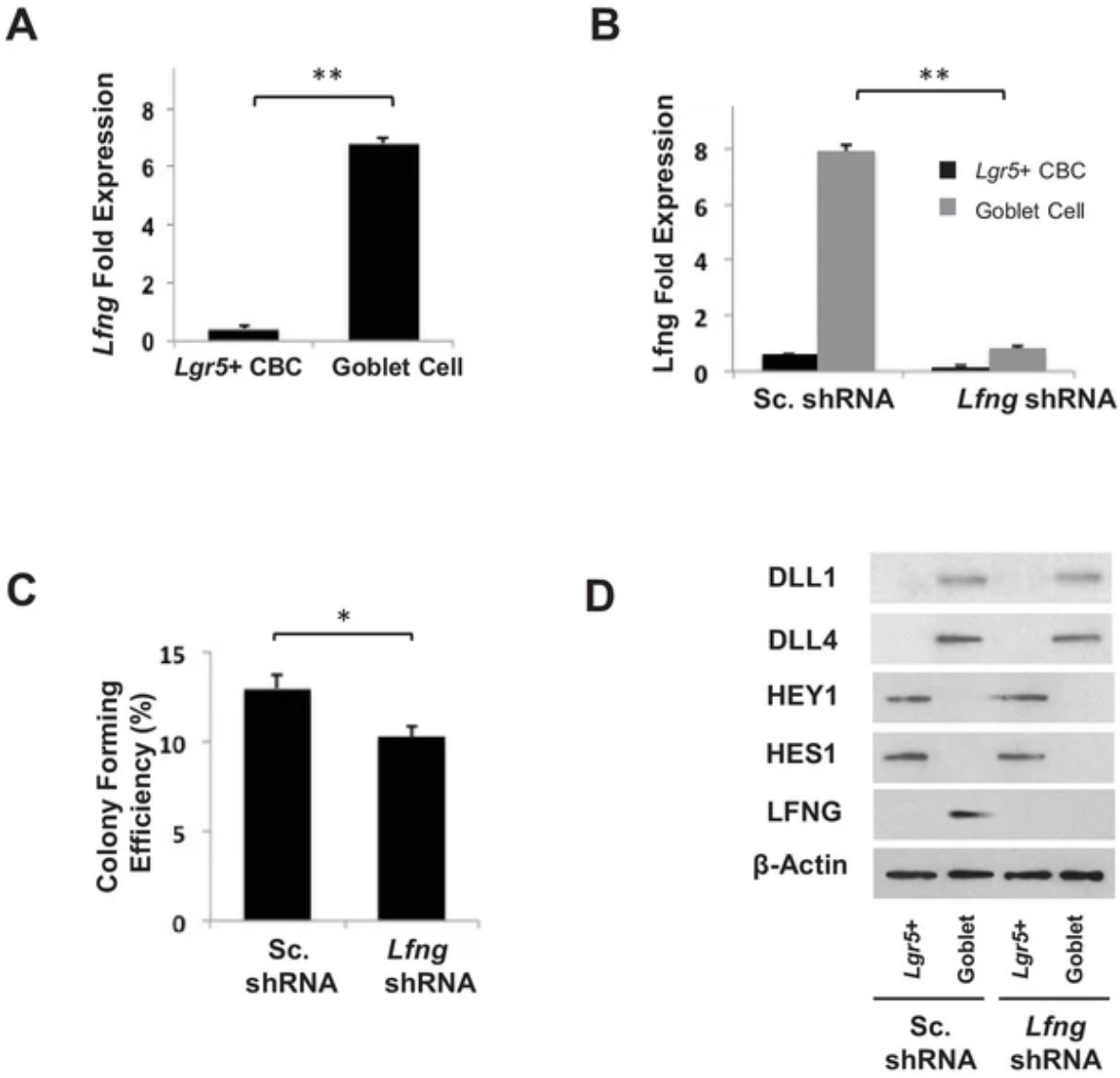
(**E**) GFP (red) shows the *Lfng* expression and NICD (green) identifies the cells with NOTCH1 activity. Scale bar represents 20 μm.

(**F**) Representative bright field and co-IF images of *Lfng* KD and control organoids indicating *Lgr5*-GFP (green) expression. MUC2 (red) marks goblet cells. DAPI (blue) labels nuclei and scale bar represents 25 μm.

(**G**) Representative plots indicating gated percentage of *Lgr5*+ (GFP^{high}) and goblet cells ($UEA\text{-}1^{+}/CD24^{-}$) of *Lfng* KD and control organoids.

(**H**) Ratio of goblet cells to *Lgr5*-GFP + CBCs as determined by flow cytometry. The experiment was performed in triplicate and presented mean ± s.d. (***$p<0.001$).

Figure 4 – Figure Supplement 1



Characterisation of *Lfng* KD organoids.

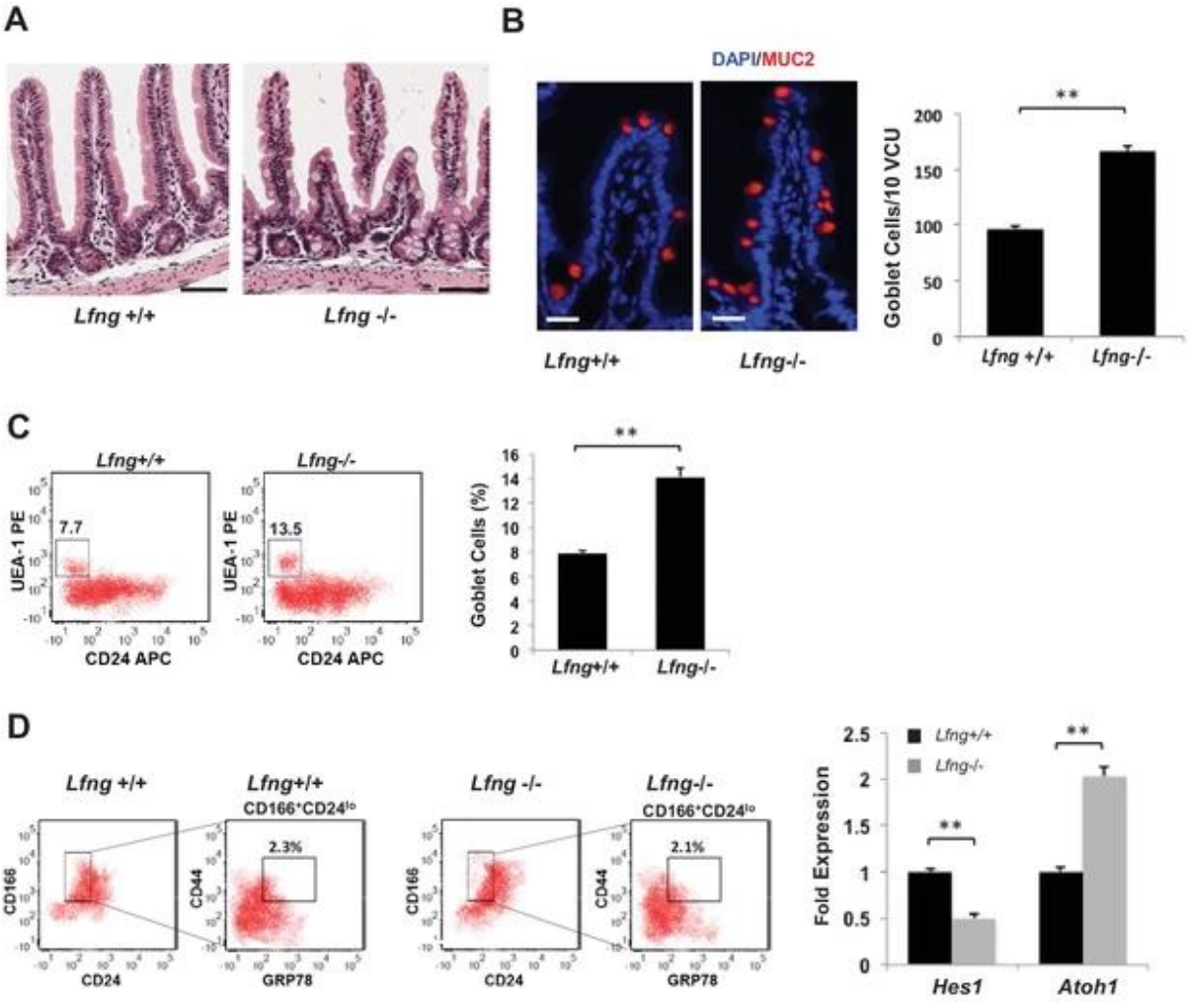
(**A**) RT-qPCR quantification of *Rfng* in *Lgr5*+ CBC and goblet cell populations from *Lgr5*-GFP murine intestinal tissue. The experiment was performed in triplicate and presented mean ± s.d.

(**B**) RT-qPCR quantification of *Lfng* in *Lgr5*+ CBC and goblet cell populations after organoids were infected with Scrambled shRNA (Sc. shRNA) or *Lfng* shRNA. The experiment was performed in triplicate and presented mean ± s.d.

(**C**) Single *Lgr5*-GFP CBCs were transduced with either Sc. shRNA or *Lfng* shRNA. Shown is colony forming efficiency measured after seven days. Quantitative analysis calculated from 1000 cells/replicate. The experiment was performed in triplicate and presented mean ± s.d.

(**D**) Single *Lgr5*-GFP CBCs were transduced with either Sc. shRNA or *Lfng* shRNA and propagated as organoids for seven days. Shown is Western blot analysis for LFNG and Notch signaling components. Actin was used as a loading control. (*$p<0.01$; **$p<0.01$).

Figure 5


A
Lfng +/+
Lfng -/-
B
DAPI/MUC2
Lfng+/+
Lfng-/-
Goblet Cells/10 VCU
**
Lfng +/+
Lfng-/-
C
Lfng+/+
Lfng-/-
UEA-1 PE
7.7
13.5
CD24 APC
Goblet Cells (%)
**
D
Lfng +/+
Lfng+/+
CD166+CD24lo
Lfng -/-
Lfng-/-
CD166+CD24lo
CD166
CD44
CD24
GRP78
2.3%
2.1%
Fold Expression
Lfng+/+
Lfng-/-
**
Hes1
Atoh1

***Lfng* loss results in increased goblet cell differentiation in vivo.**

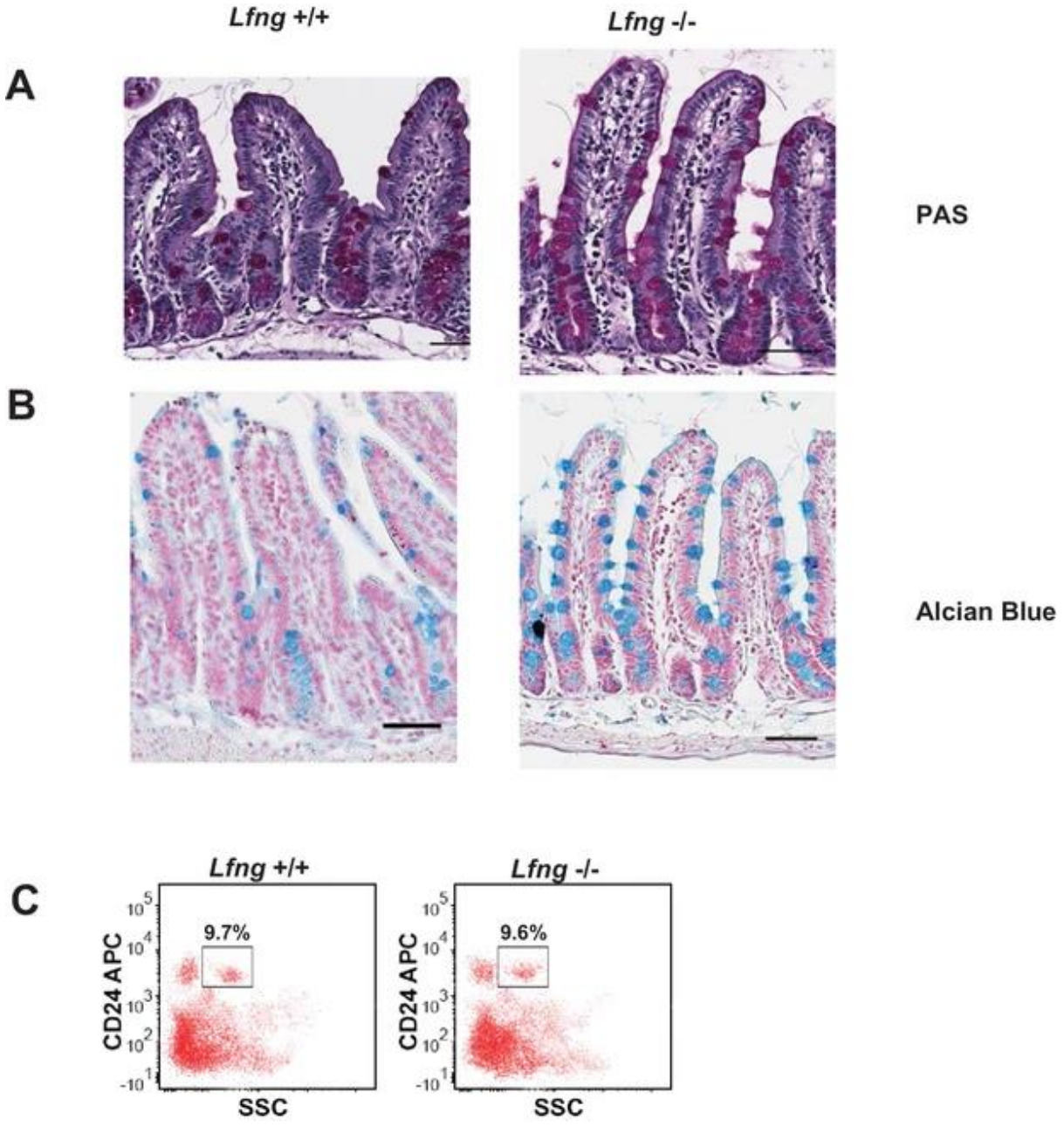
(**A**) Representative H and E sections from the small intestine of *Lfng*$^{+/+}$ and *Lfng*$^{-/-}$ mice. Scale bar represents 50 μm.

(**B**) Left: Representative IF images of intestine of *Lfng*$^{+/+}$ and *Lfng*$^{-/-}$ mice. MUC2 (red) marks goblet cells. DAPI (blue) labels nuclei. Right: Quantification of the number of goblet cells of n = 4 mice/condition and n = 500 VCU per mouse presented as mean ± s.d.

(**C**) Left: Representative plots indicating gated percentage of goblet cells (UEA-1^{+}/CD24^{-}) from intestinal tissue derived from *Lfng*$^{+/+}$ or *Lfng*$^{-/-}$ mice. Right: Percentage of goblet cells presented as mean ± s.d. The data represent n = 3 mice/condition.

(**D**) Left: Representative plots indicating gated population of intestinal progenitors from the intestine of *Lfng*$^{+/+}$ and *Lfng*$^{-/-}$ mice. Percentage reflects fraction of total population. Right: RT-qPCR measurements in progenitor cells from *Lfng*$^{+/+}$ and *Lfng*$^{-/-}$ mice. The experiment was performed in triplicate presented as mean ± s.d. (**p<0.01).

Figure 5 – Figure Supplement 1



Histological and flow cytometric analysis of *Lfng* null intestines.

(A-B) Representative images from the small intestine of *Lfng*$^{+/+}$ and *Lfng*$^{-/-}$ mouse strains stained using

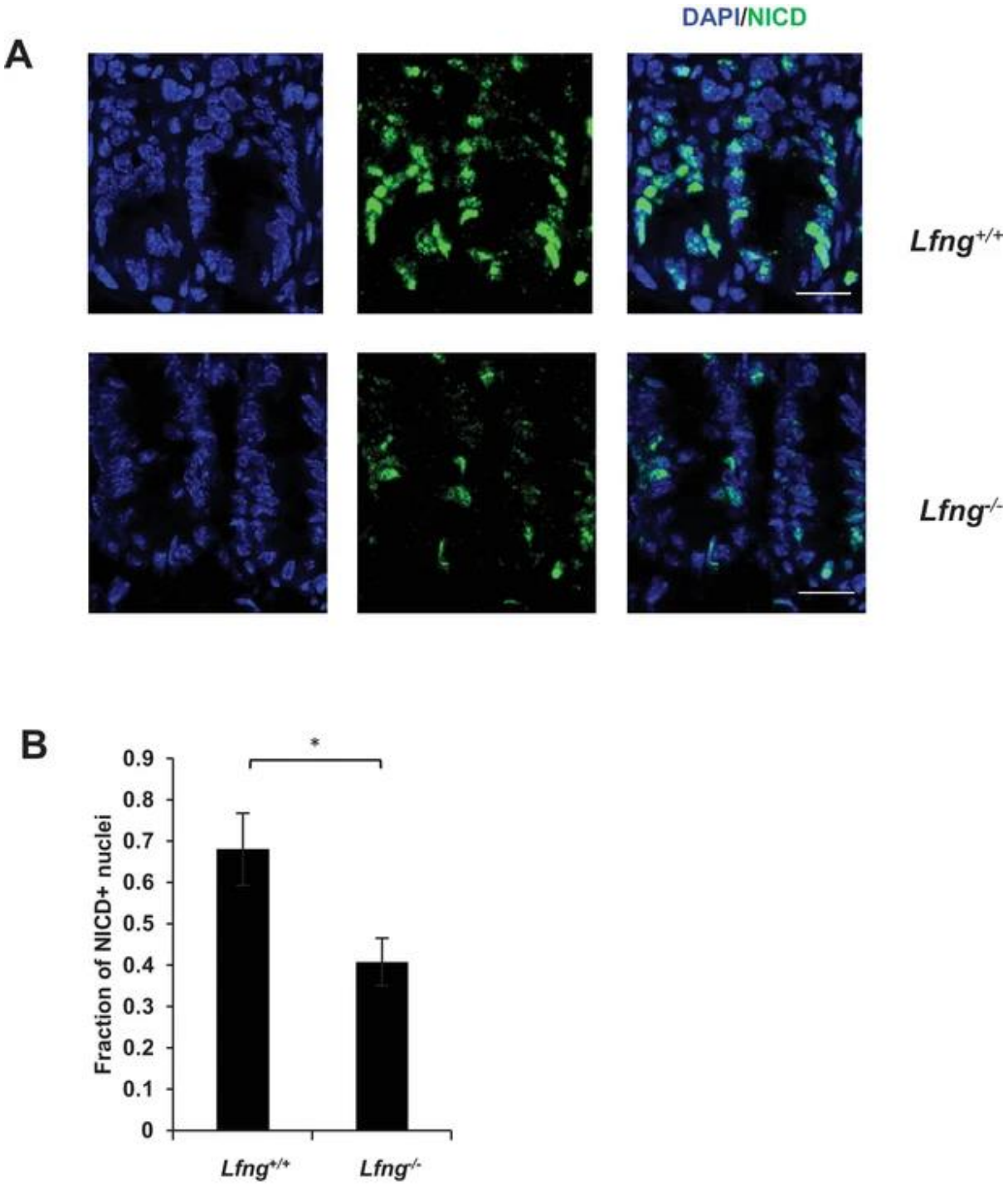
(**A**) PAS and Haematoxylin

(**B**) Alcian Blue and Nuclear Fast Red.

Scale bar represents 50 μm.

(**C**) Representative flow cytometry plots indicating gated percentage of Paneth cells from the small intestine of *Lfng*$^{+/+}$ and *Lfng*$^{-/-}$ mice.

Figure 5 – Figure Supplement 2

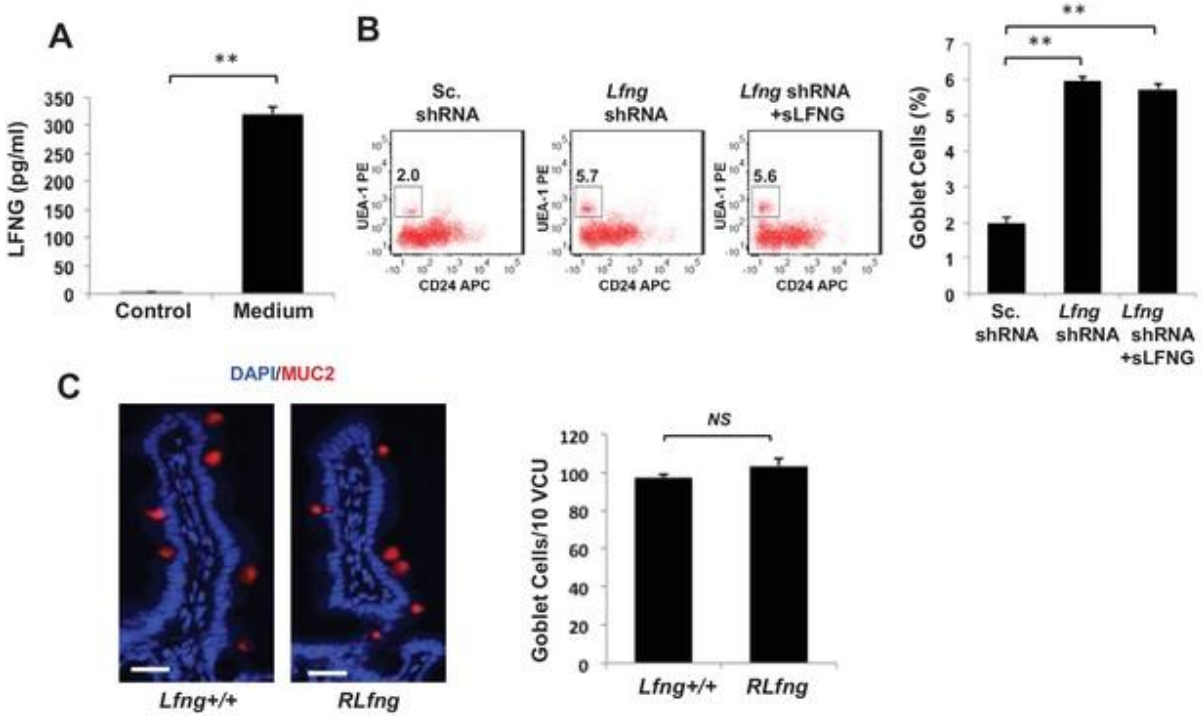


***Lfng* loss results in reduced Notch activity.**

(**A**) Representative images from the small intestine of *Lfng*$^{+/+}$ and *Lfng*$^{-/-}$ mouse strains stained for NICD. Scale bar represents 20 μm.

(**B**) Fraction of NICD + nuclei per crypt quantified from small intestines of *Lfng*$^{+/+}$ and *Lfng*$^{-/-}$ mice presented as mean ± s.d. Data represents n = 3 mice/condition. (*, $p<0.05$).

Figure 6


A
**
LFNG (pg/ml)
350
300
250
200
150
100
50
0
Control
Medium
B
Sc. shRNA
Lfng shRNA
Lfng shRNA +sLFNG
UEA-1 PE
2.0
5.7
5.6
CD24 APC
**
**
Goblet Cells (%)
7
6
5
4
3
2
1
0
Sc. shRNA
Lfng shRNA
Lfng shRNA +sLFNG
C
DAPI/MUC2
Lfng+/+
RLfng
NS
Goblet Cells/10 VCU
120
100
80
60
40
20
0
Lfng+/+
RLfng

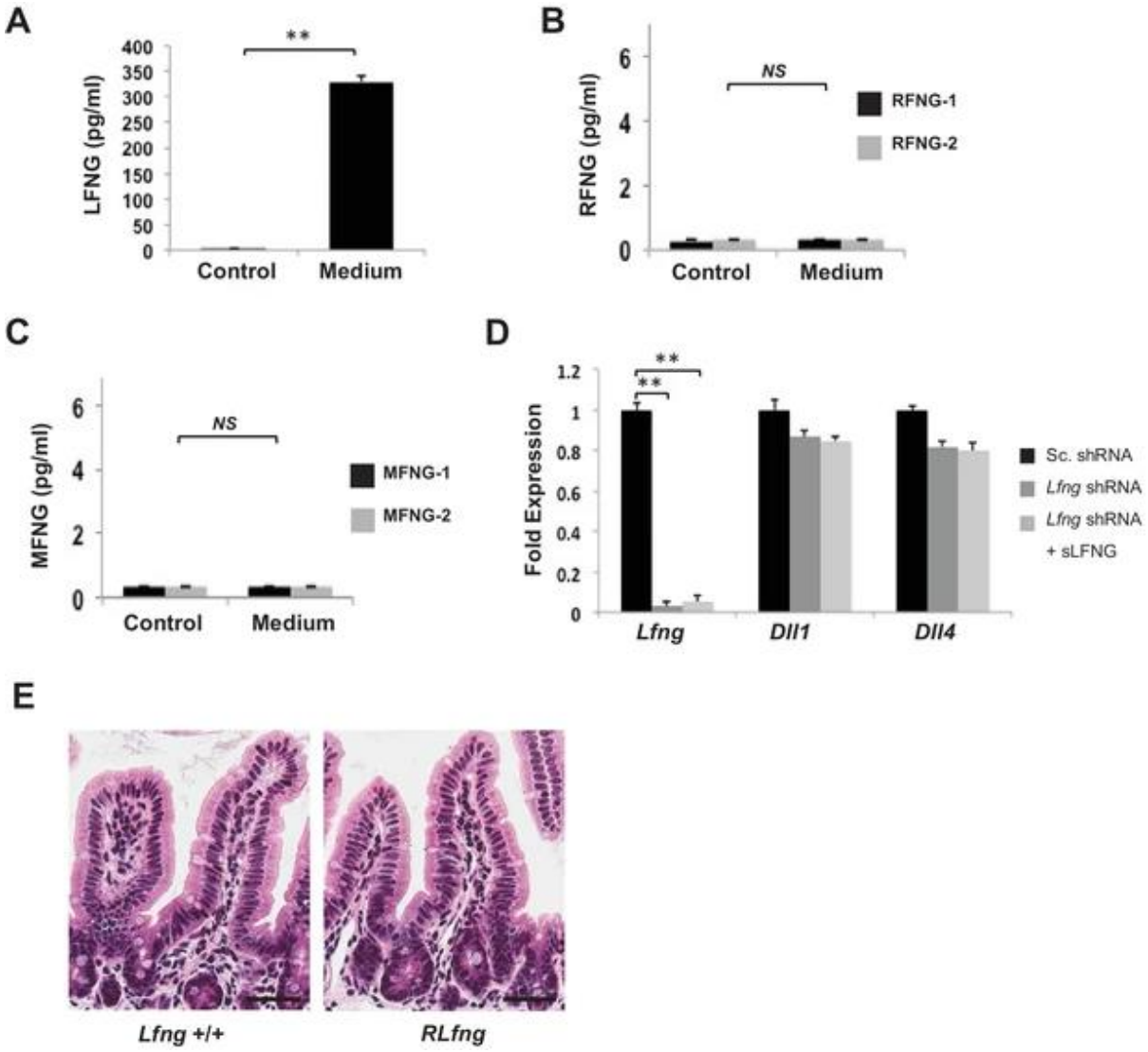
Secreted LFNG plays no apparent function.

(**A**) ELISA of the secretion of LFNG in culture medium from *Lgr5*-GFP organoids. Culture medium (T = 0 days) was used as a control. The experiment was performed in triplicate and presented as mean ± s.d.

(**B**) Left: Representative plots indicating gated percentage of goblet cells (UEA-1^{+}/$CD24^{-}$) for organoids under Sc. shRNA control, *Lfng* KD and *Lfng* KD treated with sLFNG conditions. Right: Percentage of goblet cells in each condition. The experiment was performed in triplicate and presented as mean ± s.d.

(**C**) Left: Representative IF images of intestine of $Lfng^{+/+}$ and $Lfng^{RLfng/+}$ mice. MUC2 (red) marks goblet cells. DAPI (blue) labels nuclei. Right: Quantification of the number of goblet cells of n = 4 mice/condition and n = 500 VCU/mouse. Data presented as mean ± s.d. (**p<0.01).

Figure 6 – Figure Supplement 1



Secretion of Fringe proteins.

(**A**) ELISA of the secretion of LFNG in culture medium from murine *Lgr5*-GFP+ CBCs propagated as organoids for 7 days using an additional primary LFNG antibody (LFNG-2). Culture medium (T = 0 days) was used as a control. The experiment was performed in triplicate and presented mean ± s.d.

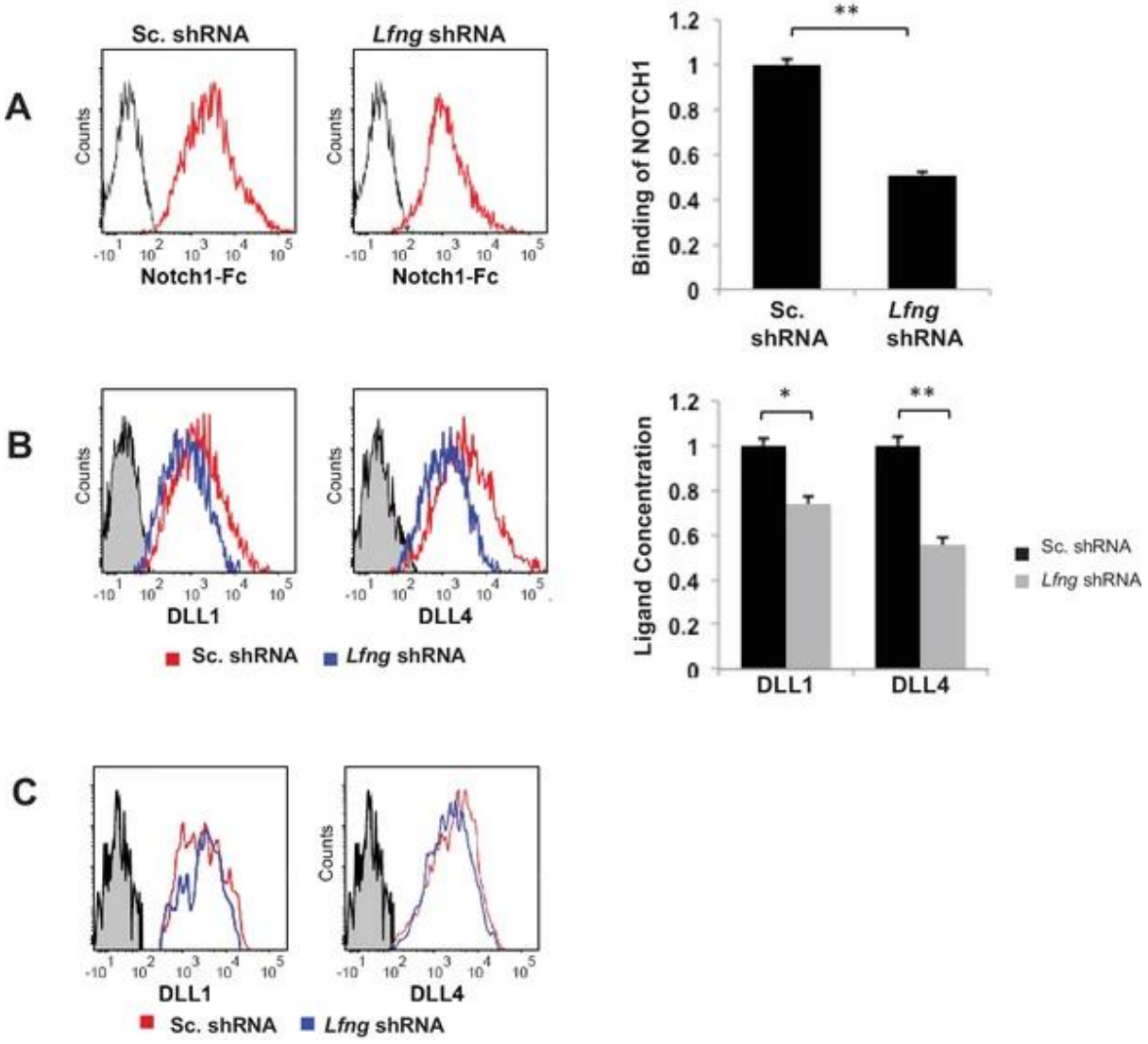
(**B**) ELISA of the secretion of RFNG in culture medium from murine *Lgr5*-GFP+ CBCs propagated as organoids for 7 days using two separate primary RFNG antibodies (RFNG-1, RFNG-2). Culture medium (T = 0 days) was used as a control. The experiment was performed in triplicate and presented mean ± s.d.

(**C**) ELISA of the secretion of MFNG in culture medium from murine *Lgr5*-GFP+ CBCs propagated as organoids for 7 days using two separate primary MFNG antibodies (MFNG-1, MFNG-2). Culture medium (T = 0 days) was used as a control. The experiment was performed in triplicate and presented mean ± s.d. (NS).

(**D**) RT-PCR quantification of *Lfng* and Notch ligands (DLL1 and DLL4) in *Lgr5*-GFP+ organoids that were infected with Scrambled shRNA (Sc. shRNA) or *Lfng* shRNA. *Lfng* shRNA-expressing CBC organoids were subsequently incubated in conditioned medium containing soluble LFNG (sLFNG) for 24 hr. The experiment was performed in triplicate and presented mean ± s.d.

(**E**) Representative H and E sections from the small intestine of *Lfng*+/+and *RLfng* mice. Scale bar represents 50 μm. (**p<0.01).

Figure 7


A
Sc. shRNA
Lfng shRNA
Counts
Notch1-Fc
Notch1-Fc
Binding of NOTCH1
**
Sc. shRNA
Lfng shRNA
B
Counts
DLL1
DLL4
Sc. shRNA
Lfng shRNA
Ligand Concentration
*
**
DLL1
DLL4
Sc. shRNA
Lfng shRNA
C
Counts
DLL1
DLL4
Sc. shRNA
Lfng shRNA

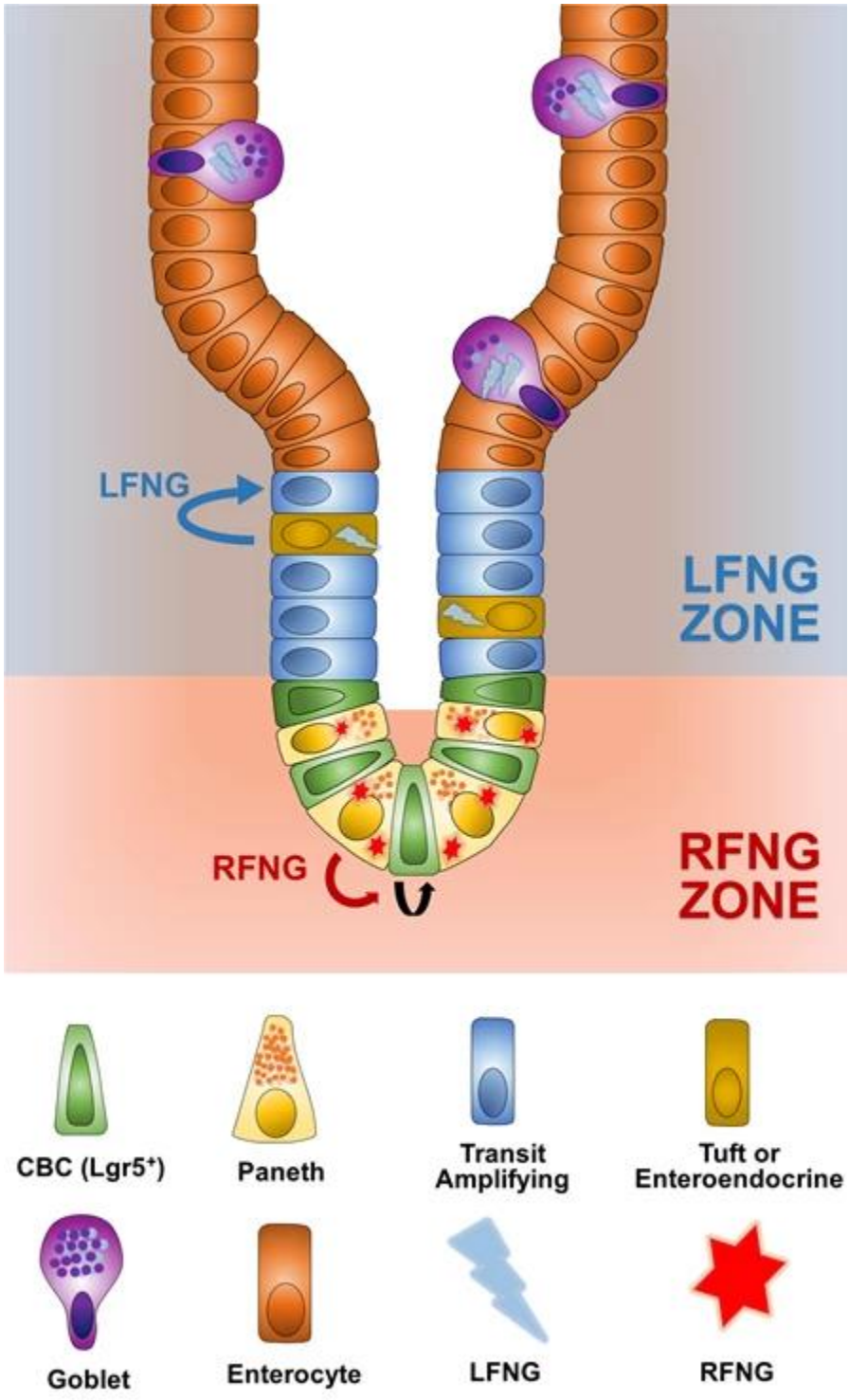
LFNG promotes cell surface expression of DLL.

(**A**) Ligand availability on *Lfng* KD and Sc. Control goblet cells. Representative traces (left) and MFI (right) showing ligand binding to NOTCH1 measured by flow cytometry. Unstained goblet cells were used as a negative control. The experiment was performed in triplicate and presented as mean ± s.d.

(**B**) Cell surface DLL1 and DLL4 concentration on *Lfng* KD and Sc. Control unpermeabilised goblet cells. Left: Representative traces measured by flow cytometry. Right: MFI measurements. The experiment was performed in triplicate and presented as mean ± s.d.

(**C**) Cell surface DLL1 and DLL4 concentration on *Lfng* KD and Sc. Control permeabilised goblet cells. (**$p<0.01$).

Figure 8


LFNG
LFNG ZONE
RFNG
RFNG ZONE
CBC (Lgr5*)
Paneth
Transit Amplifying
Tuft or Enteroendocrine
Goblet
Enterocyte
LFNG
RFNG

Summary.

Rfng is enriched in the Paneth cells and promotes cell surface expression of DLL1 and DLL4. This promotes Notch activity in the neighbouring *Lgr5*+ CBCs assisting their self-renewal. *Mfng* does not appear to contribute significantly in maintaining the epithelium. *Lfng* on the other hand is expressed by enteroendocrine, Tuft, and goblet cells and suppresses the secretory lineage.

Supplementary File 1

Antibodies used:

Primary Antibody	Supplier	Catalogue #	RRID #	Dilution*
anti-β-ACTIN	Abcam	ab6276	RRID:AB_2223210	1:4000 (WB)
Anti-CD24 (APC)	Abcam	ab51535	RRID:AB_2072741	1:500 (FC)
Anti-DLL1	Abcam	ab85346	RRID:AB_1860332	1:500 (WB)
Anti-DLL4	Abcam	ab7280	RRID:AB_449562	1:1000 (WB)
Anti-GFP	Abcam	ab5450	RRID:AB_304897	1:200 (IF) (IF-p)
Anti-HES1	Abcam	ab108937	RRID:AB_10862625	1:1000 (WB)
Anti-HES5	Santa Cruz Biotechnology	sc-25395	RRID:AB_2118099	1:500 (WB)
Anti-HEY1	Abcam	ab154077		1:1000 (WB)
Anti-JAG1	Santa Cruz Biotechnology	sc-6011	RRID:AB_649689	1:500 (WB)
Anti-KRT20	Origene	TA300936	RRID:AB_2265515	1:1000 (FC)
Anti-LFNG	Santa Cruz Biotechnology	sc-324092		1:1000 (WB) 1:500 (ELISA)
Anti-LYSOZYME	Abcam	ab108508	RRID:AB_10861277	1:100 (IF)

Anti-MFNG	Santa Cruz Biotechnology	sc-292668	RRID:AB_11149012	1:1000 (WB) 1:500 (ELISA)
Anti-MUC2	Abcam	ab76774	RRID:AB_1523987	1:100 (IF)
Anti-MUC2	Santa Cruz Biotechnology	sc-15334	RRID:AB_2146667	1:200 (IF-p)
Anti-RFNG	Santa Cruz Biotechnology	sc-8240	RRID:AB_2178618	1:1000 (WB) 1:500 (ELISA)
Anti-CK20	Abcam	ab118574	RRID:AB_10901782	1:200 (IF-p)
Anti-Ki-67	Abcam	ab15580	RRID:AB_443209	1:200 (IF-p)
Anti-CHGA	Abcam	ab45179		1:50 (IF-p)
Anti-DCAMKL1	Abcam	ab37994		1:200 (IF-p)
Anti-NICD (Val1744)	Cell Signalling Technology	4147	RRID:AB_2153348	1:600 (IF-p)

***Application: IF: (Immunofluorescence); WB: (Western Blotting); FC: (Flow Cytometry); IF-p: (IF on paraffin embedded sections)**

REFERENCES

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., *et al.* (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature *449*, 1003-1007.

Bray, S.J. (2006). Notch signalling: a simple pathway becomes complex. Nature reviews Molecular cell biology *7*, 678-689.

Bu, P., Chen, K.-Y., Chen, J.H., Wang, L., Walters, J., Shin, Y.J., Goerger, J.P., Sun, J., Witherspoon, M., and Rakhilin, N. (2013). A microRNA miR-34a-regulated bimodal switch targets Notch in colon cancer stem cells. Cell stem cell *12*, 602-615.

Bu, P., Wang, L., Chen, K.-Y., Srinivasan, T., Murthy, P.K.L., Tung, K.-L., Varanko, A.K., Chen, H.J., Ai, Y., and King, S. (2016). A miR-34a-Numb feedforward loop triggered by inflammation regulates asymmetric stem cell division in intestine and colon cancer. Cell stem cell *18*, 189-202.

Fre, S., Huyghe, M., Mourikis, P., Robine, S., Louvard, D., and Artavanis-Tsakonas, S. (2005). Notch signals control the fate of immature progenitor cells in the intestine. Nature *435*, 964-968.

Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., and Hatten, M.E. (2003). A gene expression

atlas of the central nervous system based on bacterial artificial chromosomes. Nature *425*, 917-925.

Haines, N., and Irvine, K.D. (2003). Glycosylation regulates Notch signalling. Nature reviews Molecular cell biology *4*, 786-797.

Hansson, E.M., Teixeira, A.I., Gustafsson, M.V., Dohda, T., Chapman, G., Meletis, K., Muhr, J., and Lendahl, U. (2006). Recording Notch signaling in real time. Developmental neuroscience *28*, 118-127.

Hicks, C., Johnston, S.H., diSibio, G., Collazo, A., Vogt, T.F., and Weinmaster, G. (2000). Fringe differentially modulates Jagged1 and Delta1 signalling through Notch1 and Notch2. Nature Cell Biology *2*, 515-520.

Huet, G., Gouyer, V., Delacour, D., Richet, C., and Zanetta, J.P. (2003). Involvement of glycosylation in the intracellular trafficking of glycoproteins in polarized epithelial cells. Biochimie.

Ishikawa, N., Wakelin, D., and Mahida, Y.R. (1997). Role of T helper 2 cells in intestinal goblet cell hyperplasia in mice infected with Trichinella spiralis. Gastroenterology *113*, 542-549.

Iso, T., Kedes, L., and Hamamori, Y. (2003). HES and HERP families: Multiple effectors of the notch signaling pathway. Journal of Cellular Physiology *194*, 237-255.

Kabiri, Z., Greicius, G., Madan, B., Biechele, S., Zhong, Z., Zaribafzadeh, H., Aliyev, J., Wu, Y., Bunte, R., and Williams, B.O. (2014). Stroma provides an intestinal stem cell niche in the absence of epithelial Wnts. Development *141*, 2206-2215.

Kim, T.H., and Shivdasani, R.A. (2011). Genetic evidence that intestinal Notch functions vary regionally and operate through a common mechanism of Math1 repression. Journal of Biological Chemistry.

Koo, B.-K., Stange, D.E., Sato, T., Karthaus, W., Farin, H.F., Huch, M., Van Es, J.H., and Clevers, H. (2012). Controlled gene expression in primary Lgr5 organoid cultures. Nature methods *9*, 81.

Kopan, R. (2002). Notch: a membrane-bound transcription factor. Journal of cell science *115*, 1095-1097.

LeBon, L., Lee, T.V., Sprinzak, D., Jafar-Nejad, H., and Elowitz, M.B. (2014). Fringe proteins modulate Notch-ligand cis and trans interactions to specify signaling states. eLife *3*.

Maletic-Savatic, M., Semerci, F., Choi, W.T.-S., Bajic, A., Thakkar, A., Encinas, J.M., Depreux, F., Segil, N., and Groves, A.K. (2017). Lunatic fringe-mediated Notch signaling regulates adult hippocampal neural stem cell maintenance. eLife *6*, e24660.

Moloney, D.J., Panin, V.M., Johnston, S.H., Chen, J., Shao, L., Wilson, R., Wang, Y., Stanley, P., Irvine, K.D., Haltiwanger, R.S., *et al.* (2000). Fringe is a glycosyltransferase that modifies Notch. Nature *406*, 369-375.

Moran, J.L., Shifley, E.T., Levorse, J.M., Mani, S., Ostmann, K., Perez-Balaguer, A., Walker, D.M., Vogt, T.F., and Cole, S.E. (2009). Manic fringe is not required for embryonic development, and fringe family members do not exhibit redundant functions in the axial skeleton, limb, or hindbrain. Developmental Dynamics *238*, 1803-1812.

Ohtsubo, K., and Marth, J.D. (2006). Glycosylation in cellular mechanisms of health and disease. Cell *126*, 855-867.

Okubo, Y., Sugawara, T., Abe-Koduka, N., Kanno, J., Kimura, A., and Saga, Y. (2012). Lfng regulates the synchronized oscillation of the mouse segmentation clock via trans-repression of Notch signalling. Nature Communications *3*, 1141.

Pan, Z., Sikandar, S., Witherspoon, M., Dizon, D., Nguyen, T., Benirschke, K., Wiley, C., Vrana, P., and Lipkin, S.M. (2008). Impaired placental trophoblast lineage differentiation in Alkbh1−/− mice. Developmental dynamics *237*, 316-327.

Panin, V.M., Papayannopoulos, V., Wilson, R., and Irvine, K.D. (1997). Fringe modulates Notch–ligand interactions. Nature *387*, 908-912.

Pellegrinet, L., Rodilla, V., Liu, Z., Chen, S., Koch, U., Espinosa, L., Kaestner, K.H., Kopan, R., Lewis, J., and Radtke, F. (2011). Dll1- and dll4-mediated notch signaling are required for homeostasis of intestinal stem cells. Gastroenterology *140*, 1230.

Rampal, R., Luther, K.B., and Haltiwanger, R.S. (2007). Notch signaling in normal and disease States: possible therapies related to glycosylation. Current molecular medicine *7*, 427-445.

Riccio, O., van Gijn, M.E., Bezdek, A.C., Pellegrinet, L., van Es, J.H., Zimber-Strobl, U., Strobl, L.J., Honjo, T., Clevers, H., and Radtke, F. (2008). Loss of intestinal crypt progenitor cells owing to inactivation of both Notch1 and Notch2 is accompanied by derepression of CDK inhibitors p27Kip1 and p57Kip2. EMBO reports *9*, 377-383.

Rodríguez-Colman, M.J., Schewe, M., Meerlo, M., Stigter, E., Gerrits, J., Pras-Raves, M., Sacchetti, A., Hornsveld, M., Oost, K.C., and Snippert, H.J. (2017). Interplay between metabolic identities in the intestinal crypt supports stem cell function. Nature *543*, 424-427.

Ryan, M.J., Bales, C., Nelson, A., Gonzalez, D.M., Underkoffler, L., Segalov, M., Wilson-Rawls, J., Cole, S.E., Moran, J.L., Russo, P., *et al.* (2008). Bile duct proliferation in Jag1/fringe heterozygous mice identifies candidate modifiers of the alagille syndrome hepatic phenotype. Hepatology *48*, 1989-1997.

Sato, T., van Es, J.H., Snippert, H.J., Stange, D.E., Vries, R.G., van den Born, M., Barker, N., Shroyer, N.F., van de Wetering, M., and Clevers, H. (2010). Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. Nature *469*, 415-418.

Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., *et al.* (2009). Single Lgr5 stem cells

build crypt–villus structures in vitro without a mesenchymal niche. Nature *459*, 262-265.

Schröder, N., and Gossler, A. (2002). Expression of Notch pathway components in fetal and adult mouse small intestine. Gene expression patterns : GEP *2*, 247-250.

Schwank, G., Andersson-Rolf, A., Koo, B.-K., Sasaki, N., and Clevers, H. (2013). Generation of BAC transgenic epithelial organoids. PloS one *8*, e76871.

Shifley, E.T., and Cole, S.E. (2008). Lunatic fringe protein processing by proprotein convertases may contribute to the short protein half-life in the segmentation clock. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1783*, 2384-2390.

Shroyer, N.F., Helmrath, M.A., Wang, V., Antalffy, B., Henning, S.J., and Zoghbi, H.Y. (2007). Intestine-Specific Ablation of Mouse atonal homolog 1 (Math1) Reveals a Role in Cellular Homeostasis. Gastroenterology *132*, 2478-2488.

Taylor, P., Takeuchi, H., Sheppard, D., Chillakuri, C., Lea, S.M., Haltiwanger, R.S., and Handford, P.A. (2014). Fringe-mediated extension of O-linked fucose in the ligand-binding region of Notch1 increases binding to mammalian Notch ligands. Proceedings of the National Academy of Sciences of the United States of America *111*, 7290-7295.

Ueo, T., Imayoshi, I., Kobayashi, T., Ohtsuka, T., Seno, H., Nakase, H., Chiba, T., and Kageyama, R. (2012). The role of Hes genes in intestinal development, homeostasis and tumor formation. Development *139*, 1071-1082.

van der Flier, L.G., and Clevers, H. (2009). Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium. Annual Review of Physiology *71*, 241-260.

Van Es, J.H., Sato, T., Van De Wetering, M., Lyubimova, A., Nee, A., Gregorieff, A., Sasaki, N., Zeinstra, L., Van Den Born, M., and Korving, J. (2012). Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. Nature cell biology *14*, 1099-1104.

VanDussen, K.L., Carulli, A.J., Keeley, T.M., Patel, S.R., Puthoff, B.J., Magness, S.T., Tran, I.T., Maillard, I., Siebel, C., Kolterud, Å., *et al.* (2012). Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells. Development *139*, 488-497.

Wang, F., Scoville, D., He, X.C., Mahe, M.M., Box, A., Perry, J.M., Smith, N.R., Lei, N., Davies, P.S., Fuller, M.K., *et al.* (2013). Isolation and Characterization of Intestinal Stem Cells Based on Surface Marker Combinations and Colony-Formation Assay. Gastroenterology *145*, 383.

Wang, Y., Shao, L., Shi, S., Harris, R.J., Spellman, M.W., Stanley, P., and Haltiwanger, R.S. (2001). Modification of Epidermal Growth Factor-like Repeats withO-Fucose MOLECULAR CLONING AND EXPRESSION OF A NOVEL GDP-FUCOSE PROTEINO-FUCOSYLTRANSFERASE. Journal of Biological Chemistry *276*, 40338-40345.

Williams, D.R., Shifley, E.T., Braunreiter, K.M., and Cole, S.E. (2016). Disruption of somitogenesis by a novel dominant allele of Lfng suggests important roles for protein processing and secretion. Development *143*, 822-830.

Wong, V.W., Stange, D.E., Page, M.E., Buczacki, S., Wabik, A., Itami, S., Van De Wetering, M., Poulsom, R., Wright, N.A., and Trotter, M.W. (2012). Lrig1 controls intestinal stem-cell homeostasis by negative regulation of ErbB signalling. Nature cell biology *14*, 401.

Wu, Y., Cain-Hom, C., Choy, L., Hagenbeek, T.J., de Leon, G.P., Chen, Y., Finkle, D., Venook, R., Wu, X., Ridgway, J., *et al.* (2010). Therapeutic antibody targeting of individual Notch receptors. Nature *464*, 1052-1057.

Yang, L.-T., Nichols, J.T., Yao, C., Manilay, J.O., Robey, E.A., and Weinmaster, G. (2005). Fringe glycosyltransferases differentially modulate Notch1 proteolysis induced by Delta1 and Jagged1. Molecular biology of the cell *16*, 927-942.

Zheng, X., Tsuchiya, K., and Okamoto, R. (2011). Suppression of hath1 gene expression directly regulated by hes1 via notch signaling is associated with goblet cell depletion in ulcerative colitis. Inflammatory bowel ….

CHAPTER 3

AP-1 MEDIATED EPIGENETIC REPROGRAMMING OF LUNG EPITHELIAL CELLS BY MUTANT KRAS

Preetish Kadur Lakshminarasimha Murthy, Rui Xi, Diana Arguijo, Jeffrey I. Everitt, Dewran D. Kocak, Silvestre Vicent, Gregory E. Crawford, Timothy Reddy and Xiling Shen

AUTHOR CONTRIBUTIONS

PKLM and RX designed and performed all the experiments, unless stated otherwise, and interpreted the data. **PKLM** and DA performed bioinformatic analysis. **PKLM** and JIE performed histopathological assessment of tissue sections. GEC and SV contributed to data not shown in the manuscript. **PKLM** wrote the manuscript with inputs from RX, DA and XS. DDK and TR made the FOSL1-Flag tagged A549 cell-line. XS initiated and supervised the study, and secured funding.

INTRODUCTION

Mutant *KRAS* is one of the most frequent drivers of epithelial cancers. KRAS mutations are found in approximately 97% of pancreatic ductal adenocarcinomas (PDAC), 45% of colorectal adenocarcinomas (COAD) and 30% of lung adenocarcinomas (LUAD) (Cox et al., 2014). Oncogenic mutation keeps KRAS in its active GTP-bound state leading to constant downstream signalling through RAF, PI3K and other kinases leading to uncontrolled cellular proliferation. Research over the years has shown that the role of mutant KRAS in neoplasia involves not only the self-sufficiency of growth factors but other 'hallmarks of cancer' as well (Hanahan and Weinberg, 2011; Pylayeva-Gupta et al., 2011). KRAS-mutant cells that undergo transformation are able to suppress apoptosis, reprogram their metabolism, remodel the microenvironment by inducing angiogenesis and evade an immune response (Pylayeva-Gupta et al., 2011). However, it has also been shown that in many primary cell-types with unmutated P53 and $P16^{Ink4a}$ - $P19^{Arf}$ pathways, expression of mutant *Kras* leads to a senescent phenotype (Kamijo et al., 1997; Sarkisian et al., 2007; Serrano et al., 1997). And there are reports of colonic epithelial cells that are seemingly unaffected by mutant KRAS (Aivado et al., 2000; Zhu et al., 1997). These suggest that the ability of mutant *Kras* depends on how pliable a cell is to be transformed into a neoplastic one.

Cellular identity which, for a long time, was defined by the expression of a handful, if not one, of marker genes, is now - with the advent of single-cell transcriptomics - being defined by the overall mRNA composition (Shapiro et al., 2013). However, recent studies have shown how the organisation of chromatin affects the function of transcription factors and the need to include epigenetic information to define

cellular identity (Bell et al., 2011; Buenrostro et al., 2018; Morris, 2019). This made us consider whether the chromatin organisation in a cell is a determinant of its response to a KRAS mutation. More importantly, if a drastic remodelling of chromatin is critical for KRAS driven neoplastic transformation (Ge et al., 2017).

Adult epithelial cells can often be scored based on their stemness, that is the ability to self-renew and differentiate into other cell-types, and categorised into stem, progenitor and differentiated cells (Rawlins and Hogan, 2006). Resident stem cells have been shown to be the origin of tumours in their respective tissues (Barker et al., 2009; Latil et al., 2017). However, many cell-types and sub-types have been shown to demonstrate stemness in the lung (Barkauskas et al., 2013; Kumar et al., 2011; Nabhan et al., 2018; Zacharias et al., 2018). Cells in the pulmonary epithelial lining also show remarkable plasticity (Tata and Rajagopal, 2017). The two main cell types of the alveolus, alveolar type I (AEC1) and alveolar type II (AEC2) can give rise to each other under the right context (Barkauskas et al., 2013; Jain et al., 2015). Club and basal cells of the airway show a similar behaviour (Tata and Rajagopal, 2017). Accordingly, many cell types have been considered to be cells of origin of lung tumours (Sutherland et al., 2014). Hence, the pulmonary epithelium is a good system to study the effect of mutant *Kras* in various cell types and follow changes to the chromatin architecture in each case.

In this study we show that mutant *Kras* remodels the epigenetic landscape of Club and AEC2 cells of the lung. We find that FOSL1 based AP-1 transcription factor directly binds to and guides the nucleosome remodelling SWI/SNF complex to increase chromatin accessibility at genomic loci controlling the expression of genes necessary for neoplastic transformation. We show that AP-1 mediated epigenetic reprogramming

is not restricted to pulmonary epithelium and can also be seen in intestinal stem cells, suggesting that it is a general mechanism downstream of mutant *Kras*.

RESULTS

Neoplastic progression in SK and CK lung adenocarcinoma models

Intranasal delivery of adenovirus containing Cre recombinase to *Kras*$^{LSL\text{-}G12D/+}$; *Trp53*$^{fl/fl}$ mice is one of the most common methods used to model lung adenocarcinoma (Kwon and Berns, 2013; Tammela et al., 2017; Vallejo et al., 2017; Winslow et al., 2011). However, as mutant *Kras* is expressed in multiple cell types in this model, we chose the methods described in (Xu et al., 2012), where Cre recombinase is expressed in a cell-type specific manner. Unlike (Xu et al., 2012), tumour suppressor *Trp53* was not deleted in our study as it has been shown to be not important during the early stages of lung tumorigenesis (Muzumdar et al., 2016). This allows us to delineate, specifically, the effects of mutant *Kras* in pulmonary tumours.

We studied the initial stages of KRAS mediated lung adenocarcinogenesis using two mouse models: *Sftpc*-CreER; *Kras*$^{LSL\text{-}G12D/+}$; R26R-tdTomato and *Scgb1a1*-CreER; *Kras*$^{LSL\text{-}G12D/+}$; R26R-fGFP, henceforth referred to as SK and CK models respectively (Figure 1A). In the SK model, mutant *Kras* is expressed by the alveolar type II (AEC2) cells after tamoxifen doses leading to their proliferation (Xu et al., 2012). Hyperplastic lesions are found throughout the lung four weeks post Cre induction (Figure 1B). Several atypical adenomatous hyperplastic (AAH) lesions, adenomas and very rarely adenocarcinoma can also be observed (Figure S1A). Significant macrophage infiltration is seen in the hyperplastic regions likely due to the overproduction of surfactants (Wright, 1990) (Figures 1B and S1A).

In the CK model, mutant *Kras* is expressed in the Club cells, putative Bronchioalveolar stem cells (BASCs) and a small fraction (~1%) of AEC2 (CC10+

AEC2) (Rawlins et al., 2009; Xu et al., 2012). Neoplastic transformation is slower in this model and terminal bronchiolar hypertrophy is mostly observed 4 weeks post the tamoxifen doses. Alveoli appear largely normal at this stage. Bronchiolar hyperplasia can be observed after 10 weeks, while adenomas and adenocarcinomas can be found after 16 weeks. Tumours appear to arise from the bronchioles and protrude into the alveolar space (Figures 1C, 1D and S1B).

For further analysis, we isolated the cells of interest by FACS. AEC2 cells were isolated from *Sftpc*-CreER; R26R-tdTomato mice (henceforth referred to as SC mice) based on the tdTomato label (Figure S1C-S1D). Normal cells - Club, AEC2 and CC10+ AEC2 cells - were isolated from *Scgb1a1*-CreER; R26R-fGFP mice (henceforth referred to as CC mice) by a combination of GFP label and lysotracker dye staining (Figures 1E, 1F and S1E) (der Velden et al., 2013). SK Neoplastic cells were isolated at 4 weeks post tamoxifen doses by dissociating the whole lung and sorting based on the lineage label tdTomato (Figure 1G). CK neoplastic lungs were similarly isolated at 4, 10, 16- and 22-weeks post *Kras*^{G12D} activation using lineage label fGFP (Figure S1F). Additionally, visible tumours (larger than 3 mm in diameter) were excised from 22w CK lungs, dissociated and fGFP labelled cells were isolated for further analysis.

Mutant KRAS remodels the open chromatin landscape

To identify changes to the chromatin state, we performed Assay for Transposase Accessible Chromatin and Sequencing or ATAC-seq (Buenrostro et al., 2013) on the normal and neoplastic cells isolated from both models (Figures S2A and S2B, and Supplementary Table 1). Principal component analysis (PCA) shows that the neoplastic cells from the CK mice cluster with the Club cells and that the neoplastic cells from the

SK mice are closer to AEC2 and CC10+ AEC2 cells in their epigenetic landscapes. This suggests that Club and AEC2 cells are the major contributors to neoplasia in the CK and SK models respectively. Then, we separated the dataset into two groups: one with AEC2 and SK neoplastic cells and the other with Club and CK neoplastic cells (Figure 2A). We identified differentially accessible regions in both the groups using DiffBind (Ross-Innes et al., 2012) (Figures 2B, S2C and S2D). Following the convention in (Denny et al., 2016), we shall henceforth refer to the regions with increased accessibility in neoplastic cells as "newly open" regions and those with reduced accessibility in neoplastic cells as "newly closed". We find that most of the differentially accessible regions are away from the Transcription Start Sites (TSS) and are in intronic or intragenic regions of the genome (putative *cis*-regulatory elements or enhancers) (Figures 2C and S2E-G). To understand their relevance to gene regulation, using GREAT, we mapped the newly open regions in the neoplastic cells from both SK and CK models to the genes whose expression they are likely to regulate (McLean et al., 2010). The gene sets thus obtained strongly correlate with the genes upregulated in the Kras2LA lung adenocarcinoma model (Sweet-Cordero et al., 2005). Interestingly, the newly open regions also appear to regulate apoptotic pathway genes in both SK and CK models (Figure 2D).

We observed increased chromatin accessibility at enhancer loci of genes that have been implicated in *Kras* mediated tumorigenesis. We find that enhancers near *Ccnd1* locus and 'super-enhancers' of *Myc* are more accessible in the neoplastic cells (Pylayeva-Gupta et al., 2011; Zhang et al., 2016) (Figures 2E and 2F). This suggests that there are mechanisms, other than the known growth promoting kinase cascade, that

control neoplastic transformation. We observed that the changes to chromatin are very similar in the two models (Figures 2E, 2F, S2H, S2I). Upon quantification, we find that about 60% of newly open regions in the 22w neoplastic cells from CK mice were also newly open in the SK neoplastic cells (Figure 2G). This suggests that there might be a common mechanism by which the genomic regions are made accessible to transcription factors in the presence of constitutively active KRAS signalling.

Newly open regions in neoplastic cells are linked to AP-1 activity

To delineate the factors responsible for the epigenetic changes observed, we identified the transcription factor (TF) motifs enriched in the differentially accessible regions using HOMER (Heinz et al., 2010a). In the AEC2 derived neoplastic cells, regions with increased accessibility are enriched for AP-1 motif sites and those with reduced accessibility harbour ETS and FOX family motifs (Figures 3A and S3A). Using PIQ (Protein Interaction Quantification), a machine-learning tool to identify binding probabilities of TFs, we found that AP-1 factors are more likely to be bound to the genome in *Kras* mutant cells than in normal AEC2 (Sherwood et al., 2014a) (Figure S3B). We then performed bivariate genomic footprinting using the Bagfoot algorithm which identifies differential TF activity based on changes to the TF footprint and TF motif flanking accessibility (Baek et al., 2017). AP-1 TFs showed significantly higher accessibility and deeper footprints in the SK neoplastic cells than in AEC2 cells suggesting their increased activity (Figures 3B and 3C). These data raise the possibility that AP-1 is involved ‘opening’ the right regions of chromatin to transform normal cells into neoplastic ones.

Even in the CK neoplastic cells collected at 4, 10, 16- and 22-weeks post tamoxifen doses, regions with increased accessibility are enriched for AP-1 motif sites (data not shown). CK model provides us with the advantage of tracking the epigenetic changes with tumour development. By hierarchical clustering we identify three classes of differentially accessible regions. The first among them encompasses regions that have high accessibility in Club cells but low in all the 4 stages of CK neoplasia. Similar to the SK model, motif sites of FOX family TFs are enriched here. The third cluster has regions with low accessibility in the Club cells but high in all the neoplastic stages. Lung epithelial lineage specific TF, NKX2.1, has its motif sites enriched here. Interestingly, the second cluster has regions that become more accessible with time and is strongly enriched for AP-1 motif sites (Figure 3D). We also see a linear increase with time in the ratio of percentage of sites harbouring an AP-1 motif to that in background sequences (Figure 3E). Bagfoot analysis also shows an increased AP-1 motif flanking accessibility with time (Figures S3C-G). Increase in flanking accessibility at the AP-1 motif in CK neoplastic cells is shown in (Figures 3F and S3H-K). These observations suggest that AP-1 might be involved in continuous remodelling of chromatin.

Transcription factors can displace nucleosomes to increase accessibility to genomic sites (Bell et al., 2011; Denny et al., 2016). To understand whether AP-1 complex is involved in nucleosome remodelling, we evaluate changes to nucleosome occupancy at AP-1 motif occurrences in the newly accessible regions in the SK neoplastic cells using NucleoATAC software (Schep et al., 2015). We find that in newly open regions in neoplastic cells, nucleosomes have been depleted at putative AP-1 binding sites, whereas no such change is observed in the constitutively open regions

(Figure 3G). To ensure that the observation is specific to AP-1 and not newly open peaks, we perform a similar analysis at CTCF, a TF which was not identified as an outlier by BagFoot analysis in the SK model, motif occurrences. We find that nucleosome occupancy at CTCF motif occurrences in both AEC2 and SK neoplastic cells is nearly identical in the newly open regions (Figure S3L). We observed that the reduction in nucleosome occupancy at AP-1, but not CTCF, motif centre in the newly open regions of SK neoplastic cells correlates with the fold change of newly open regions (Figure S3M). Similarly, in the CK model, we observe a reduction in the nucleosome occupancy at the AP-1 motif centre in the newly open peaks at 4, 10, 16- and 22-weeks post tamoxifen doses (Figures S4A-E). These data suggest that AP-1 complex is directing the nucleosome remodelling factors to various region of the genome to increase their accessibility to TF binding (Vierbuchen et al., 2017).

AP-1 activity is necessary for tumour initiation *in vitro*

To test whether AP-1 mediated changes to the open chromatin are necessary for tumour initiation, we performed an organoid based tumour initiation assay where mutant *Kras* is expressed in AEC2 cells seeded in Matrigel and their ability to form tumour spheroids is assayed under various conditions (Figure 4A). *Sftpc*-CreER mouse model is 'leaky' with a fraction of AEC2 cells showing CreER activity in the absence of any tamoxifen doses. So, we isolated normal AEC2 cells from *Sftpc*-CreER; LSL-*Kras*^{G12D}; R26R-tdTomato mice by sorting for the LysoTracker+ tdTomato- population and excluded the tdTomato+ cells that were present due to the "leakiness" of CreER (Figure 4B). Normal AEC2 were plated and treated with 4-OHT to express mutant *Kras* and simultaneously AP-1 activity was inhibited by a small molecule SR 11302 (Fanjul et

al., 1994; Huang et al., 1997). After 8 days we observed that *Kras* mutant AEC2 cells form smaller and fewer spheroids when AP-1 activity was inhibited in comparison to those treated by DMSO as control (Figures 4C and 4D).

AP-1 has been shown to control cell proliferation and apoptosis in a context dependent manner (Shaulian and Karin, 2001). To ensure that blockage of AP-1 activity is specifically affecting *Kras* mutant AEC2, we treated normal AEC2 from *Sftpc*-CreER; R26R-tdTomato mice with SR 11302 or DMSO (vehicle control) for 8 days. Blocking AP-1 activity does not significantly affect the number or size of normal AEC2 derived organoids (Figures 4E and 4F).

Inhibition of AP-1 activity leads to reduced proliferation of *Kras*-mutant AEC2 *in vivo*

To investigate whether blocking AP-1 activity would affect the *Kras*-mutant AEC2 cells *in vivo*, we injected SR 11302 or DMSO to SK mice along with tamoxifen doses and continued to administer the drug or vehicle control for 3 weeks. *Kras* mutant AEC2 cells from the lungs of mice that received the AP-1 inhibitor showed an almost 3-fold reduction in proliferation (Figure 4G).

FOSL1 is the AP-1 protein of interest in *Kras* mutant cells.

AP-1 complex consists of various permutations of dimers of FOS (FOS, FOSB, FOSL1 and FOSL2) and JUN (JUN, JUNB and JUND) family proteins all of which have a similar DNA binding motif making them difficult to distinguish bioinformatically (Eferl and Wagner, 2003; Khan et al., 2017; Vierbuchen et al., 2017). We performed RNA-seq of AEC2 and 4w SK neoplastic cells to see if changes in gene expression could offer clues to identify the AP-1 complex composition (Figures S5A-

C). We find that only *Fosl1*, among the *Fos* and *Jun* family genes, is upregulated in *Kras* mutant neoplastic cells (Figure 5A). We also observe increased FOSL1 protein levels in 4w SK lungs when compared to normal (Figure 5B). FOSL1 protein could also be detected in CK neoplastic cells (Figures 5C, S5D and S5E). In addition to that, recent publications have also shown that *Fosl1* is important in *Kras* driven lung adenocarcinomas (Elangovan et al., 2018; Vallejo et al., 2017).

Nucleosome remodelling SWI/SNF complex directly binds to AP-1 in *KRAS* mutant cells

Heterodimers of FOS and JUN family proteins have been shown to open the chromatin at enhancer loci by recruiting SWI/SNF complexes and aid in the differentiation of mouse embryonic fibroblasts (MEFs) (Vierbuchen et al., 2017). Similarly, we hypothesised that AP-1 complex consisting of FOSL1 is involved in chromatin remodelling and directing the development of mutant *Kras* driven tumours.

We FLAG tagged FOSL1 gene in KRAS mutant human lung adenocarcinoma cell line, A549, using CRISPR-Cas9 based tools. We extracted nuclear protein from these cells and immunoprecipitated (IP) the FLAG tagged protein while using the extract from WT A549 cells as a control (Figure 5D). We used mass spectrometry to identify proteins that co-precipitated with FOSL1. Known binding-partners of FOSL1, JUNB, JUN and JUND were detected (Eferl and Wagner, 2003). We also detected SWI/SNF complex proteins SMARCC2, ACTL6A, ARID1A etc. confirming that AP-1 directly binds to the nucleosome remodelling BAF complex in KRAS mutant cells (Figure 5E).

To test whether loss of *Fosl1* prevents the changes to the accessible chromatin in *Kras* mutant cells, we crossed SK mice with *Fosl1*$^{fl/fl}$ mice (Eferl et al., 2004). After Cre mediated recombination, exons 3 and 4 of *Fosl1* gene are deleted and FOSL1-GFP fusion protein is expressed in these mice (Eferl et al., 2004). 4 weeks after tamoxifen doses to *Sftpc*-CreER; *Kras*$^{LSL\text{-}G12D/+}$; *Fosl1*$^{fl/fl}$; R26R-tdTomato mice (referred to as SFK mice) we observed that only about 1% of lineage labelled (tdTomato+) cells expressed GFP (Figures 5F and S5F). It appears that AEC2 cells in which Cre mediated recombination did not occur at the *Fosl1* locus out competed the cells in which *Fosl1* was deleted. This is consistent with previous reports which have shown that loss of *Fosl1* leads to reduced lung tumour burden (Elangovan et al., 2018; Vallejo et al., 2017). ATAC-seq performed on these tdTomato+ GFP+ cells yielded noisy data in which no peaks could be identified (data not shown). *Fosl1* has been shown to promote the survival of Kras mutant lung adenocarcinoma cells (Elangovan et al., 2018). It is possible that these tdTomato+ GFP+ cells underwent apoptosis, owing to the high pressure in the FACS sorter or because they were outside the body for a few hours, leading to very high background in ATAC-seq data.

Signature of the cell of origin remains in neoplastic cells

AP-1 has been shown to select enhancer loci while cooperating with cell type specific transcription factors (Vierbuchen et al., 2017). To identify the collaborating TFs in lung tumours, we took advantage of the two models we have studied. We observe that, out of 5871 newly open regions that harbour an AP-1 motif in CK 22w neoplastic cells, 3713 (63%) are also newly open in the SK 4w neoplastic cells. SK 4w also has another 16,713 regions containing an AP-1 motif specific to them. (Figure 6A). We used

Centrimo package to identify TF motifs enriched in the neighbourhood (~50 bp away) of putative AP-1 binding sites in regions open only in CK 22w cells or SK 4w neoplastic cells, and the overlapping regions (Bailey et al., 2009). We rank ordered the enriched TF motifs in each of the three cases based on P-value and summarised a few of the top hits in a "bubble-plot" (Figure 6B). Transcription factors ranked high in all the three conditions include TCF7 and LEF1 which are downstream effectors of Wnt signalling. Wnt activity has previously been shown to be important in lung adenocarcinoma (Pacheco-Pinedo and Morrisey, 2011; Tammela et al., 2017). We hypothesise that the factor ranked very high in CK specific peak and low in the other two cases could be a CK specific collaborating TF. We identify HES1, TF downstream of Notch signalling, as one such factor. Detection of cleaved Notch1 (NICD) shows that Notch signalling is active in CK lung tumours (Figure 6C). Gene set enrichment analysis shows that Notch activity has decreased in SK neoplastic cells when compared to AEC2 (Figure 6D) (Subramanian et al., 2005). Notch signalling has been shown to be necessary for maintaining Club cell identity (Lafkas et al., 2015). Blocking Notch activity results in reduced tumour formation in the CK model (Xu et al., 2014). This, along with the data in Figure 2A, suggests that tumour progression is influenced by its cell of origin.

We performed BagFoot analysis to detect differential TF activity in CC10-AEC2 and Club cells isolated from normal (CC) mouse lungs (Figure 6E). And we ran a similar analysis comparing SK 4w neoplastic cells and CK (16w and 22w) neoplastic cells (Figures S6A and 6F). We find that the factors, like FOX family TFs, that show increased activity in Club cells also show a similar activity in the CK neoplastic cells.

HNF, NKX and GATA family TFs show higher activity in both AEC2 and AEC2 derived neoplastic cells.

AP-1 signature in *Kras*^{G12D} expressing *Lgr5*+ stem cells of the intestine

Expression of *Kras*^{G12D} in intestinal epithelial cells alone does not lead to the formation of tumours (Snippert et al., 2014). And, field cancerisation can be observed in human colonic epithelium where KRAS mutations are found in histologically normal tissue (Aivado et al., 2000; Zhu et al., 1997). However, It is possible that the *Lgr5*+ stem cells in this region are more likely to develop tumours than the normal ones after acquiring an APC mutation leading to constitutively active Wnt signalling (Barker et al., 2009; Snippert et al., 2014).

Mutant KRAS has been shown to upregulate AP-1 transcription factors in many cell types such as mouse embryonic fibroblasts and pancreatic and colonic epithelial cells (Pylayeva-Gupta et al., 2011; Shao et al., 2014; Vallejo et al., 2017). And, on observing a similar AP-1 mediated epigenetic remodelling in both Club and AEC2 derived neoplasia, we hypothesised that the mechanism is dependent on mutant *Kras* activity but not restricted to the cell-type. We decided to test the hypothesis by expressing mutant *Kras* in the intestinal epithelial cells using *Lgr5*-CreER-EGFP; LSL-*Kras*^{G12D}; *Rosa26*-tdTomato mice. Intestinal tissue was harvested 5 weeks after tamoxifen doses and appeared to be histologically normal (Figure 7A). Ribbons of lineage-labelled cells can be seen (Figure 7B). We sorted *Kras*-mutant *Lgr5*+ stem cells based on GFP expression and tdTomato label and the *Kras* WT cells based on GFP expression and the absence of tdTomato expression from the same mouse (Figure 7C)

and performed ATAC-seq. Bivariate genomic footprinting shows an increased AP-1 activity in the Kras mutant *Lgr5*+ stem cells (Figure 7D).

DISCUSSION

In this study, we have used ATAC-seq to characterise the changes to accessible chromatin in pulmonary epithelial cells in response to an oncogenic *Kras* mutation. By analysing the initial stages of tumorigenesis in the absence of additional supportive mutations, we find that *Kras* mutation leads to a remodelled chromatin landscape of Club and AEC2 cells. We find that AP-1 complex containing of FOSL1 directly binds to the SWI/SNF complex and directs to regions in the genome to increase accessibility (Figure 7E). We observe that a small but significant increase, about one and a half times, in the *Fosl1* transcript levels leads to a drastic increase in the protein level. This suggests that post-translational modifications to FOSL1, downstream of KRAS signaling, are at play here and need to be further characterised. We have found that all the JUN family proteins, c-JUN, JUNB and JUND bind to FOSL1 in KRAS mutant cells. It would be interesting to characterise the differences in the functions of each of the three AP-1 heterodimers. Similarly, the composition of the SWI/SNF complexes which are made up of multiple sub-units and their differential interaction with AP-1, which may be cell type specific, would be important to characterise.

Kras-mutant AEC2 cells fail to form spheroids *in vitro* when AP-1 activity is blocked using a small molecule SR 11302. *Kras* mutant AEC2 show a significant reduction in proliferation *in vivo* when treated with the same drug suggesting that AP-1 dependent chromatin remodelling is necessary for *Kras* driven tumour initiation. Other than the pulmonary epithelial cells, we also observe an increased AP-1 signature on the accessible chromatin of *Kras* mutant *Lgr5+* stem cell of the intestine. A previous report analysing chromatin accessibility in *Kras* mutant intrafollicular epidermal cells and hair

follicle stem cells shows an enrichment of AP-1 motifs in both the cases (Latil et al., 2017). All these data suggest that the mechanism is cell type independent.

We observe that tumours retain a signature of their cells of origin. This has been observed previously in squamous cell carcinoma of the skin where the cell of origin determines the metastatic ability (Latil et al., 2017). It is likely that, in collaboration with cell-type specific transcription factors, AP-1 can open cell of origin specific regions on the chromatin, such as HES1 in CK tumours. Understanding the chromatin state of the cell of origin can also help identify novel targeted therapeutic targets (Latil et al., 2017).

METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Monoclonal ANTI-FLAG® M2 antibody produced in mouse	Sigma-Aldrich	Cat #: F1804 RRID: AB_262044
Anti-Ki67 antibody	Abcam	Cat #: ab15580 RRID: AB_443209
Ki-67 (D3B5) Rabbit mAb (Mouse Preferred; IHC Formulated)	Cell Signaling Technology	Cat #: 12202 RRID: AB_2620142
β-Actin (13E5) Rabbit mAb	Cell Signaling Technology	Cat #: 4970 RRID: AB_2223172
Cleaved Notch1 (Val1744) (D3B8) Rabbit mAb	Cell Signaling Technology	Cat #: 4147 RRID: AB_2153348
Anti-GFP antibody	Abcam	Cat #: ab13970 RRID: AB_300798
Anti-GFP antibody	Abcam	Cat #: ab5450 RRID: AB_304897
Anti-GFP antibody - ChIP Grade	Abcam	Cat #: ab290 RRID: AB_303395

tdTomato antibody	Scigen	Cat #: AB8181-200 RRID: AB_2722750
Fra-1 Antibody (D-3)	Santa Cruz Biotechnology	Cat #: sc-376148 RRID: AB_11012022
Anti-p63 antibody	Abcam	Cat #: ab735 RRID: AB_305870
Bacterial and Virus Strains		
E. coli: DH5α Competent Cells	Thermo Fisher Scientific	Cat #: 18258012
Chemicals, Peptides, and Recombinant Proteins		
Tamoxifen	Sigma-Aldrich	Cat #: T5648-1G
SR 11302	Tocris	Cat #: 2476
SR 11302	Santa Cruz Biotechnology	Cat #: sc-204295
Cholera Toxin from *Vibrio cholerae*	Sigma-Aldrich	Cat #: C8052-.5MG
Bovine Pituitary Extract (BPE)	Thermo Fisher Scientific	Cat #: 13028014
Insulin-Transferrin-Selenium-Sodium Pyruvate (ITS-A) (100X)	Thermo Fisher Scientific	Cat #: 51300044

Recombinant Mouse EGF Protein, CF	R&D Systems	Cat #: 2028-EG-200
Retinoic acid	Sigma-Aldrich	Cat #: R2625-50MG
3X FLAG Peptide	Sigma-Aldrich	Cat #: F4799-4MG
(Z)-4-Hydroxytamoxifen	Sigma-Aldrich	Cat #: H7904-5MG
SYTOX Blue Dead Cell Stain, for flow cytometry	Thermo Fisher Scientific	Cat #: S34857
LysoTracker Red DND-99	Thermo Fisher Scientific	Cat #: L7528
LysoTracker Green DND-26	Thermo Fisher Scientific	Cat #: L7526
ACCUMAX	STEMCELL Technologies	Cat #: 07921
Y-27632	STEMCELL Technologies	Cat #: 72304
ACK Lysing Buffer	Thermo Fisher Scientific	Cat #: A1049201
RT-PCR Grade Water	Thermo Fisher Scientific	Cat #: AM9935

SYBR™ Green I Nucleic Acid Gel Stain - 10,000X concentrate in DMSO	Thermo Fisher Scientific	Cat #: S7563
RNase-Free DNase Set	Qiagen	Cat #: 79254
Ham's F-12K (Kaighn's) Medium	Thermo Fisher Scientific	Cat #: 21127030
DMEM/F-12, HEPES	Thermo Fisher Scientific	Cat #: 11330057
Halt Protease and Phosphatase Inhibitor Single-Use Cocktail (100X)	Thermo Fisher Scientific	Cat #: 78442
Corn oil (delivery vehicle for fat-soluble compounds)	Sigma-Aldrich	Cat #: C8267-500ML
Cas9 protein with NLS, high concentration	PNA Bio	Cat #: CP02
Paraformaldehyde	Sigma-Aldrich	Cat #: P6148-500G
Matrigel Growth Factor Reduced (GFR) Basement Membrane Matrix	Corning	Cat #: 354230
MEM (Minimum Essential Medium)	Thermo Fisher Scientific	Cat #: 11095098

TrueBlack Lipofuscin Autofluorescence Quencher	Biotium	Cat #: 23007
VECTASHIELD Antifade Mounting Medium with DAPI	Vector Laboratories	Cat #: H-1200
DAKO Protein Block, Serum-Free	Agilent	Cat #: X090930-2
Critical Commercial Assays		
Nextera DNA Library Prep Kit	Illumina	Cat #: FC-121-1030
NEBNext High-Fidelity 2X PCR Master Mix	New England BioLabs	Cat #: M0541S
MinElute PCR Purification Kit	Qiagen	Cat #: 28006
QIAquick PCR Purification Kit	Qiagen	Cat #: 28106
RNeasy Mini Kit	Qiagen	Cat #: 74104
QuantiFast SYBR Green PCR Kit	Qiagen	Cat #: 204054
QuantiTect Reverse Transcription Kit	Qiagen	Cat #: 205310
Tumor Dissociation Kit, mouse	Miltenyi Biotec	Cat #: 130-096-730
Lung Dissociation Kit, mouse	Miltenyi Biotec	Cat #: 130-095-927
CD45 MicroBeads, mouse	Miltenyi Biotec	Cat #: 130-052-301
Anti-FLAG M2 Magnetic Beads	Sigma-Aldrich	Cat #: M8823-1ML

Phusion Hot Start II High-Fidelity PCR Master Mix	Thermo Fisher Scientific	Cat #: F565S
MEGAshortscript™ T7 Transcription Kit	Thermo Fisher Scientific	Cat #: AM1354
MEGAclear™ Transcription Clean-Up Kit	Thermo Fisher Scientific	Cat #: AM1908
TSA Plus Fluorescein Evaluation Kit	PerkinElmer	Cat #: NEL741E001KT
Experimental Models: Cell Lines		
Human: MRC-5 cells	Duke University Cell Culture Facility (Duke CCF)	N/A
Human: A549 cells (FosL1-FLAG)	This paper	N/A
Human: A549 cells	ATCC	CCL-185
Experimental Models: Organisms/Strains		
Mouse: *Sftpc*-CreER^{T2}	(Xu et al., 2012)	N/A
Mouse: *Scgb1a1*-CreER^{TM}	(Rawlins et al., 2009)	N/A
Mouse: *Rosa26R*-tdTomato	(Xu et al., 2012)	N/A
Mouse: *Kras*$^{LSL-G12D/+}$	(Xu et al., 2012)	N/A
Mouse: *Rosa26R*-fGFP	(Xu et al., 2012)	N/A
Mouse: *Fosl1*$^{fl/fl}$	(Eferl et al., 2004)	N/A

Mouse: *Lgr5*-EGFP-IRES-creERT2	The Jackson Laboratory	JAX #: 008875
Oligonucleotides		
qPCR Primer for *Sftpc* (NM_011359.2)	GeneCopoeia	Cat #: MQP094943
qPCR Primer for *Scgb1a1* (NM_011681.2)	GeneCopoeia	Cat #: MQP030442
qPCR Primer for *Actb* (NM_007393.5)	GeneCopoeia	Cat #: MQP026493
Recombinant DNA		
Mouse gRNA pooled sub-library targeting mitochondrial DNA (pUC57kan-T7-mt-gRNA)	(Wu et al., 2016)	Addgene (Pooled Library #82480)
Software and Algorithms		
GREAT	(McLean et al., 2010)	http://great.stanford.edu/public/html/
Aperio Imagescope	N/A	https://www.leicabiosystems.com/digital-pathology/manage/aperio-imagescope/
MACS	(Zhang et al., 2008)	https://github.com/taoliu/MACS

featureCounts	(Liao et al., 2014)	https://bioinf.wehi.edu.au/featureCounts/
TCseq	(Wu and Gu, 2018)	https://bioconductor.org/packages/release/bioc/html/TCseq.html
Diffbind	(Stark and Brown, 2011)	https://bioconductor.org/packages/release/bioc/html/DiffBind.html
Homer	(Heinz et al., 2010b)	https://homer.ucsd.edu/homer/
Bagfoot	(Baek et al., 2017)	N/A
FIMO	(Bailey et al., 2009)	https://meme-suite.org/index.html
Centrimo	(Bailey et al., 2009)	https://meme-suite.org/index.html

NucleoATAC	(Schep et al., 2015)	https://nucleoatac.readthedocs.io/en/latest/
Fiji	(Schindelin et al., 2012)	https://imagej.net/Fiji
DeepTools	(Ramírez et al., 2014)	https://deeptools.readthedocs.io/en/develop/

Experimental Model and Subject Details

All experiments involving animals were approved by the Institutional Animal Care and Use Committee at Duke University and conducted under protocols A286-15-11 and A235-18-10. *Sftpc*-CreER, *Scgb1a1*-CreER, *R26R*-tdTomato, *Kras*$^{LSL\text{-}G12D/+}$ and *R26R*-fGFP mice were provided by Mark Onaitis of Duke University. Aline Bozec from University of Erlangen for provided the *Fosl1*$^{fl/fl}$ mice. *Lgr5*-EGFP-IRES-creERT2 mice were purchased from Jackson Laboratory (#8875). Mice were group housed, and food and water were provided *ad libidum*. Only 6-10 weeks old male mice were used in lung cancer experiments (Histo-pathological analysis of lung sections and isolation of primary cells for ATAC-seq, RNA-seq and qPCR). Both female and males of 6-10 weeks of age were used for the intestinal study (histo-pathological analysis of the tissue and isolation *Lgr5*+ cells for ATAC-seq)

Method details

CreER activation

Tamoxifen was dissolved in corn oil at a concentration of 20 mg/ml and stored at 4^0C for a maximum of one month. 75 mg/kg body weight doses of tamoxifen were administered to mice intraperitoneally (IP) to activate CreER. All mice used in the lung cancer study received four doses, with a gap of 48 hours between each dose, except CK 22w ones which received only two doses. *Lgr5*-CreER mice received a dose of tamoxifen on each day for four consecutive days.

In SK mice, Cre activity is seen in a few AEC2 cells in the absence of any tamoxifen. Owing to the increased surfactant production, female SK mice often succumb before reaching maturity (6 weeks of age). Hence, the study has been restricted to male mice. Surfactant burden on the male mice is high 5-6 weeks after they receive tamoxifen doses, hence samples are collected up to 4 weeks post the doses.

AP-1 inhibition *in vivo*

SK mice were intraperitoneally injected with 2 mg/kg of AP-1 inhibitor SR 11302 followed by a dose of tamoxifen after two hours. The injections were repeated after 2 days after which tamoxifen doses were stopped. Mice continued to receive SR 11302 doses once every 48 hours for 3 weeks. Mice were then euthanised and lungs were harvested.

Lung harvest

Mice were anaesthetised using Avertin. Skin and muscular layers were cut to expose the abdominal and thoracic cavities. Descending aorta was cut to exsanguinate and euthanise the animal. Cold PBS was perfused through the cardiac right ventricle to flush the blood out of the lungs. Lungs were then removed and inflated with 4% PFA and fixed overnight at 4^0C in PFA for histology. For isolation of live primary cells, lungs

were instead inflated with lung/tumour dissociation buffer (Miltenyi Biotec), cut into small pieces and incubated at 37^0C for 30 minutes (normal lung) or 45 minutes (lung with neoplastic growth). A single cell suspension was obtained by mechanically dissociating using gentleMACS (Miltenyi Biotec) dissociator and passing the samples through a 70 µm filter. RBCs were lysed by resuspending the cells in ACK lysis buffer and leucocytes were depleted by using anti-CD45 microbeads (Miltenyi Biotec).

To isolate AEC2 cells from CC mice, Lysotracker dye was used as described in (der Velden et al., 2013). Briefly, dissociated lung cells were resuspended in DMEM/F12 containing Lysotracker (1 µL dye in 14 mL medium) and incubated at 37^0C for 30 minutes.

Primary cells were isolated by FACS using SONY SH800 sorter. Sytox dyes (Blue, Green, Red and AADvanced) were used to exclude dead cells.

NGS library preparation

ATAC-seq

Harvested primary cells were immediately used to prepare ATAC-seq libraries as described in (Buenrostro et al., 2013). Briefly, isolated cells were washed in PBS and resuspended in buffer containing transposase Tn5 and incubated at 37^0C for 30 mins. Number of cells used to prepare each library has been summarised in Table S1. Tagmented DNA was PCR amplified for 10-16 cycles.

Mitochondrial DNA present in the libraries was cleaved using a CRISPR-Cas9 based method described in (Wu et al., 2016). sgRNA library targeting mitochondrial DNA was purchased from Addgene (#82480). *In vitro* transcription was performed to produce the gRNAs using MEGAshortscript Kit (Thermo Fisher Scientific). ATAC-seq

libraries were treated with 10 µg Cas9 protein and 15 µg of sgRNA in a 40 µL reaction for 2 hours at 37^0C. The libraries were purified using Qiaquick PCR purification kit (Qiagen) and sequenced using Illumina Hi-Seq 2000 (50 bp paired end) or Next-Seq (75 bp paired end).

RNA-seq

Total RNA was isolated from primary cells using RNeasy Mini kit (Qiagen). Samples with RIN > 9 were used for library preparation by Duke University's genomics core using KAPA stranded mRNA library prep kit. Libraries were sequenced using Illumia Hi-Seq 4000 (50 bp single end).

Cell culture

A549 cell-line was cultured in F12-K medium (Thermo Fisher) and MRC5 cells in MEM (Thermo Fisher). Culture media were supplemented with 10% FBS (Thermo Fisher) and 1% Antibiotic-Antimycotic (Thermo Fisher).

AEC2 cells were cultured as spheroids as described in (Barkauskas et al., 2013) with minor modifications. Isolated AEC2 cells were mixed with MRC5 cells (passage # 18-25) at a 1:5 ratio and resuspended in Matrigel. 100 µL of the suspension, containing 5000 AEC2, was pipetted to a Transwell (Corning). MTEC plus medium was added to the outer side of the Transwell. Y-27632 at 1 µM was added to the medium for the first 1-2 days. 4-OHT (1.3 µg/mL concentration) was added for the first 24 hours of culture to activate CreER. After 8 days of culture, spheroids were imaged using a Leica fluorescent microscope with 4x objective lens.

FOSL1-FLAG tagged A549 cell line production

Expression plasmids for the Cas effectors and their respective gRNAs were obtained through Addgene (Addgene #41815, 47108); the target sequence for FOSL1 was CTCAGGCGCCTCACAAAGCG. The guide RNA scaffold region was modified to remove a poly T sequence and increase the length of a hairpin structure, these result in more favourable CRISPR function in human cells (Chen et al., 2013). To create the sgRNA plasmid, oligonucleotides containing the target sequences were obtained from IDT, hybridized, phosphorylated and cloned in the appropriate plasmids using BbsI or BsmBI sites. Gibson assembly was used to create the donor plasmid: regions of the A549 genome were amplified by PCR and cloned into a plasmid backbone along with a DNA fragment that contains the FLAG and selection cassette.

A549 cells were electroporated in a 0.2 cm cuvette using Biorad's GenePulser Xcell; 2M cells were resuspended in 200 µL of OptiMEM (ThermoFisher) with 5 µg of plasmid (1 µg of donor, 1 µg of guide RNA, and 3 µg of Cas, expression vectors). Immediately after electroporation, cells were rescued with 1 mL of complete media and transferred into complete media. Media was exchanged every 2 days thereafter. Transfection efficiencies were routinely higher than 60%, as determined by fluorescence microscopy after delivery of a control eGFP expression plasmid. 5 days post electroporation cells were selected with puromycin at a concentration of 1 µg/mL. Cells were then treated with adenoviral Cre recombinase per manufacturers protocol (Univ. Iowa Vector Core). The efficiency of the desired integration was assessed by restriction fragment length polymorphism analysis. Genomic DNA of the cell population was extracted using the DNeasy kit (Qiagen). The C-terminus of FOSL1 was amplified using AccuPrime Taq DNA Polymerase (ThermoFisher) with primers that

bind external to the homology arm region. The PCR product was purified using a PCR purification kit (Qiagen) and the product was digested using PsiI (New England Biolabs), which recognizes a site specific to the FLAG sequence. Digests were run on an agarose gel and the efficiency of gene editing was assessed by densitometry; efficiency was approximately 30%. The desired modification was then confirmed by western blot using the Sigma M2 antibody. Cells were then expanded and stored for all subsequent experiments.

Nuclear protein extraction

Nuclear protein was extracted from A549 cells (WT or FOSL1-FLAG) as described in (Erdoğan et al., 2016). 4 x 10 cm dishes of A549 cells were trypsinised and resuspended in hypotonic lysis buffer (10 mM Hepes pH 7.9, 1.5 mM MgCl2, and 10 mM KCl, 10 mM DTT and 1x protease inhibitor cocktail). After 15 minutes of incubation on ice, Igepal was added to a final concentration of 0.6%. Samples were vortexed and spun down at 10,000 g for 30s. Nuclear buffer (50 mM Tris-HCl, pH 7.5; 150 mM NaCl, 0.5% Triton-X 100 and 1x protease inhibitor cocktail) was added to the pellet and sonicated for 5s twice using a probe-based sonicator.

Immunoprecipitation (IP)

IP was performed to purify FLAG tagged FOSL1 protein immediately after extracting the nuclear protein from A549 cells. Magnetic anti-Flag M2 beads (Sigma) were washed thrice using TBS and incubated with nuclear extract for 60 minutes at room temperature. Beads were then washed three times in TBS. Bead bound proteins were eluted by competition with 3X-Flag peptides.

Mass spectrometric analysis of isolated protein was performed by the proteomics core facility at Duke University GCB.

Western Blotting

The assay was performed as described in (Xi et al., 2019). Immunoprecipitated samples were electrophoresed on Mini Protean Stain free gels (Bio-rad). Proteins were then transferred to low fluorescence PVDF membrane (Bio-rad). Membrane was blocked using 5% non-fat milk in TBS for an hour at room temperature. Incubation with primary antibodies was performed overnight at 4^0C followed by three washes in TBST (TBS + 0.1% Tween-20). HRP conjugated secondary antibodies were added for an hour at room temperature and the membrane was washed thrice in TBST. Chemiluminescent signal was produced at antibody bound sites using SuperSignal West femto (Thermoo Fisher) and detected using Chemi-Doc (Bio-rad).

Histology

Fixed samples were dehydrated using a series of increasing concentrations of ethanol. Ethanol was then removed by incubation with Xylene and then embedded in paraffin. 4-8 µm sections were cut from embedded blocks and placed on slides. Slides were baked at 60^0C for an hour and deparaffinised using xylene and dried using ethanol and rehydrated. H&E staining was performed using standard methods.

Immunostaining was performed as described in (Murthy et al., 2018). Briefly, antigen retrieval was performed using Tris-EDTA buffer at pH 9 (Vector Labs). Serum free protein block (Dako) was applied for 10 mins at room temperature. Slides were then incubated in endogenous peroxide block (Dako). Sections were then incubated with primary antibodies overnight at 4^0C. Slides were washed three times in PBS and

incubated with secondary antibodies for an hour at room temperature. TSA amplification (Perkin Elmer) was performed when probing for NICD and FOSL1. Autofluorescence was quenched using TrueBlack reagent (Biotium). Mounting medium containing DAPI (Vector Labs) was used and slides were imaged using Zeiss 780i confocal microscope.

qRT-PCR

Total RNA was extracted from isolated cells using RNeasy Mini Kit (Qiagen). Primers to detect *Scgb1a1*, *Sftpc* and *Actin* were purchased from Genecopoeia.

Quantification and Statistical Analysis

General

Unless otherwise specified, data are presented as Mean ± Standard Deviation. Student's t-test was performed and a *P*-value threshold of 0.05 was used to determine statistical significance.

ATAC-seq

ATAC-seq data was processed using the pipeline developed in (Koh et al., 2016) with default parameters. 75 bp PE reads were trimmed to 50 bp PE using FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) before analysing them. Reads were aligned to mm10 genome using Bowtie 2. Duplicate and mitochondrial reads were removed. And peaks were called using MACS2 software. Peaks were annotated using HOMER (Heinz et al., 2010a). Sample quality statistics are provided in Supplementary Table 1.

Differential peak calling and annotation

Differential peaks were identified using Diffbind with a minimum of 2 biological replicates for each case. (Ross-Innes et al., 2012). Unless otherwise stated,

peaks with *P*-value < 0.05 and $\log_2$ |Fold Change| > 1 were considered 'differentially accessible'. Others were considered 'constitutively open' peaks (Denny et al., 2016). Principal component analysis was performed using dba.PlotPCA command in the DiffBind package. Differential peaks were annotated using findMotifsGenome.pl program in HOMER package (Heinz et al., 2010a). Peaks were assigned to their putative effector genes using GREAT (McLean et al., 2010) online tool using the whole genome as background.

Bivariate Genomic Footprinting (BaGFoot)

BaGFoot analysis was performed following the algorithm published in (Baek et al., 2017). mm10 genome was used as reference. Aligned and filtered ATAC-seq reads from biological replicates were pooled to create a single file for each condition and used for pairwise Bagfoot analysis.

Nucleosome occupancy

Broadpeak files of all biological replicates per condition were merged and extended by 100bp on each side. Filtered Bam files were similarly pooled and used as inputs to the NucleoATAC program (Schep et al., 2015) to obtain nucleosome occupancy scores.

Identification of collaborating transcription factors

Cell type specific transcription factors collaborating with AP-1 were identified as described in (Vierbuchen et al., 2017). 220 bp regions centred around AP-1 motif occurrences (pooled occurrences of FOS, FOSB, FOSL1, FOSL2, JUN, JUNB and JUND motifs) were used as input to the Centrimo package to identify the TF motif enrichment in the neighbourhood of putative AP-1 binding sites (Bailey et al., 2009).

RNA-Seq

RNA-seq data were analysed using the RNA Express package available on Basespace (Illumina). Briefly, reads were aligned to the mm10 genome using STAR and differentially expressed genes were identified using DESeq2 (Dobin et al., 2013; Love et al., 2014). Transcripts with $|FC| > 1.5$ and q value < 0.05 were considered to be differentially expressed.

Histopathology

H&E stained images of lung lobes were scanned using a 40x objective on Aperio Scanscope slide scanner. Various histopathological features were annotated using Aperio Imagescope software based on the guidelines in (Nikitin et al., 2004). Hypertrophic bronchiolar regions were clubbed with the normal bronchioles. No distinction was made between adenoma and adenocarcinoma. Area occupied by each feature was measured and normalised to the total lobe area. A minimum of two lobes per mouse and 3 mice per time point were analysed.

Tumour initiation assay

Images of spheroids were analysed using ImageJ. Spheroids greater than 35 µm in diameter were quantified. A minimum of 3 technical replicates were analysed for each experiment.

Figure 1

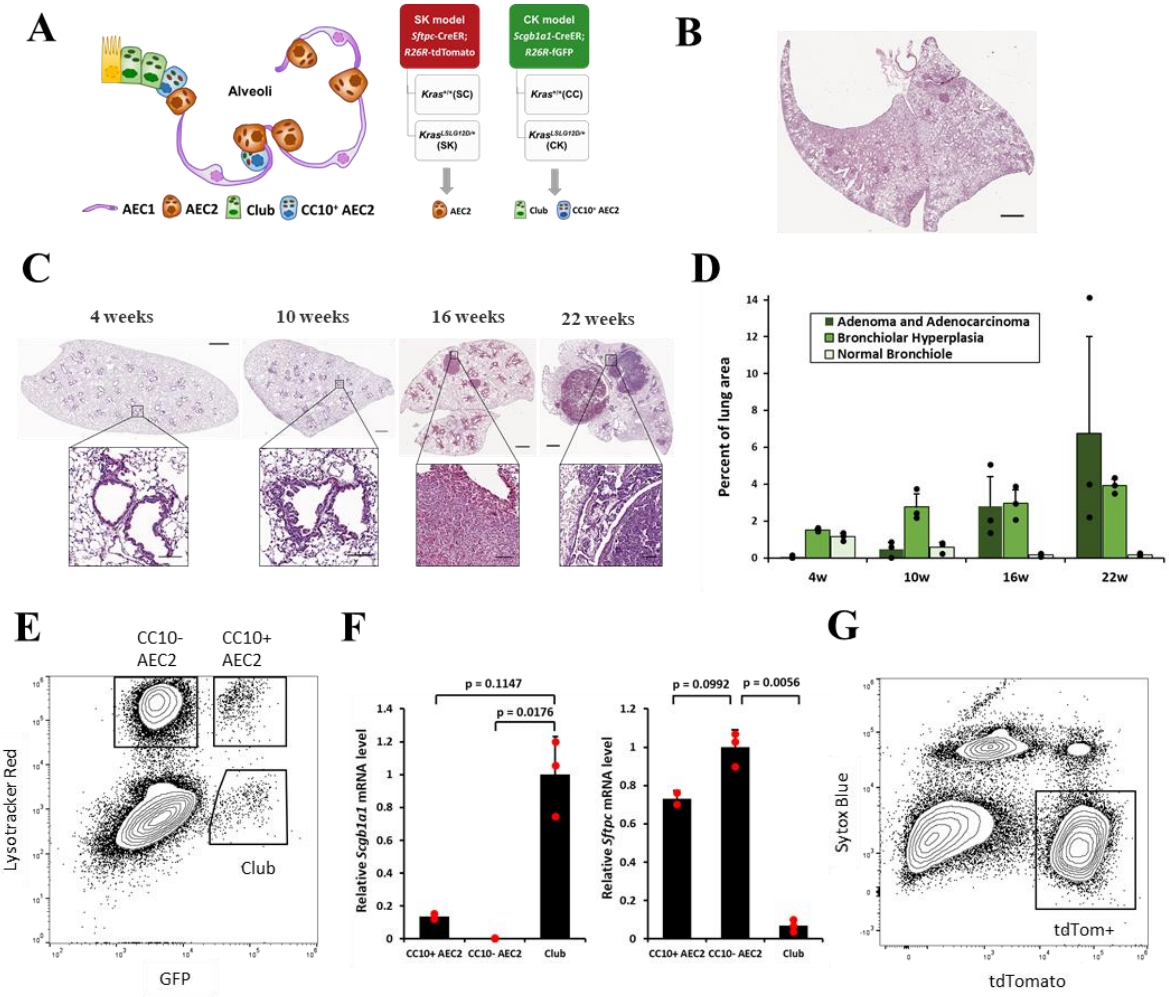

A
Alveoli
AEC1
AEC2
Club
CC10+ AEC2
SK model
Sftpc-CreER;
R26R-tdTomato
CK model
Scgb1a1-CreER;
R26R-fGFP
(SC)
(SK)
(CC)
(CK)
B
C
4 weeks
10 weeks
16 weeks
22 weeks
D
Adenoma and Adenocarcinoma
Bronchiolar Hyperplasia
Normal Bronchiole
Percent of lung area
4w
10w
16w
22w
E
CC10-
AEC2
CC10+
AEC2
Lysotracker Red
Club
GFP
F
p = 0.1147
p = 0.0176
Relative Scgb1a1 mRNA level
CC10+ AEC2
CC10- AEC2
Club
p = 0.0992
p = 0.0056
Relative Sftpc mRNA level
G
Sytox Blue
tdTom+
tdTomato

Figure 1. Neoplastic progression in SK and CK lung adenocarcinoma models

A. Schematics showing the epithelial lining of distal bronchioles and alveoli (left) and the mouse models used (right). In the *Sftpc*-CreER model, mutant *Kras* is expressed in the AEC2 cells of the alveoli and in the *Scgb1a1*-CreER model, it is expressed in the club cells of bronchioles and a subset of AEC2 (CC10+ AEC2).
B. Representative H&E image of a lung section from an SK mouse at 4 weeks after tamoxifen doses. Scale bar represents 1 mm.
C. Representative H&E images of lung sections from CK mice at 4, 10, 16 and 22 weeks after tamoxifen doses. Scale bar represents 1 mm in the whole-lobe images and 100 μm in the zoomed-in regions.
D. Bar plot showing the percent of lung area occupied by normal bronchiolar epithelium, hypertrophic and hyperplastic bronchiolar epithelium, and adenoma and adenocarcinoma in the CK model at 4, 10, 16- and 22-weeks post tamoxifen doses.
E. Representative FACS plot showing the gates used to sort AEC2, Club and CC10+ AEC2 cells from CC mice.
F. RT-qPCR showing the levels of *Scgb1a1* and *Sftpc* expression in AEC2 (n=3), Club (n=3) and CC10+ AEC2 (n=2 biological replicates) cells isolated as shown in Fig. 1E.
G. Representative FACS plot showing the gate used to sort *Sftpc*-CreER lineage labelled (tdTom+) cells from SK mice.

Figure 2

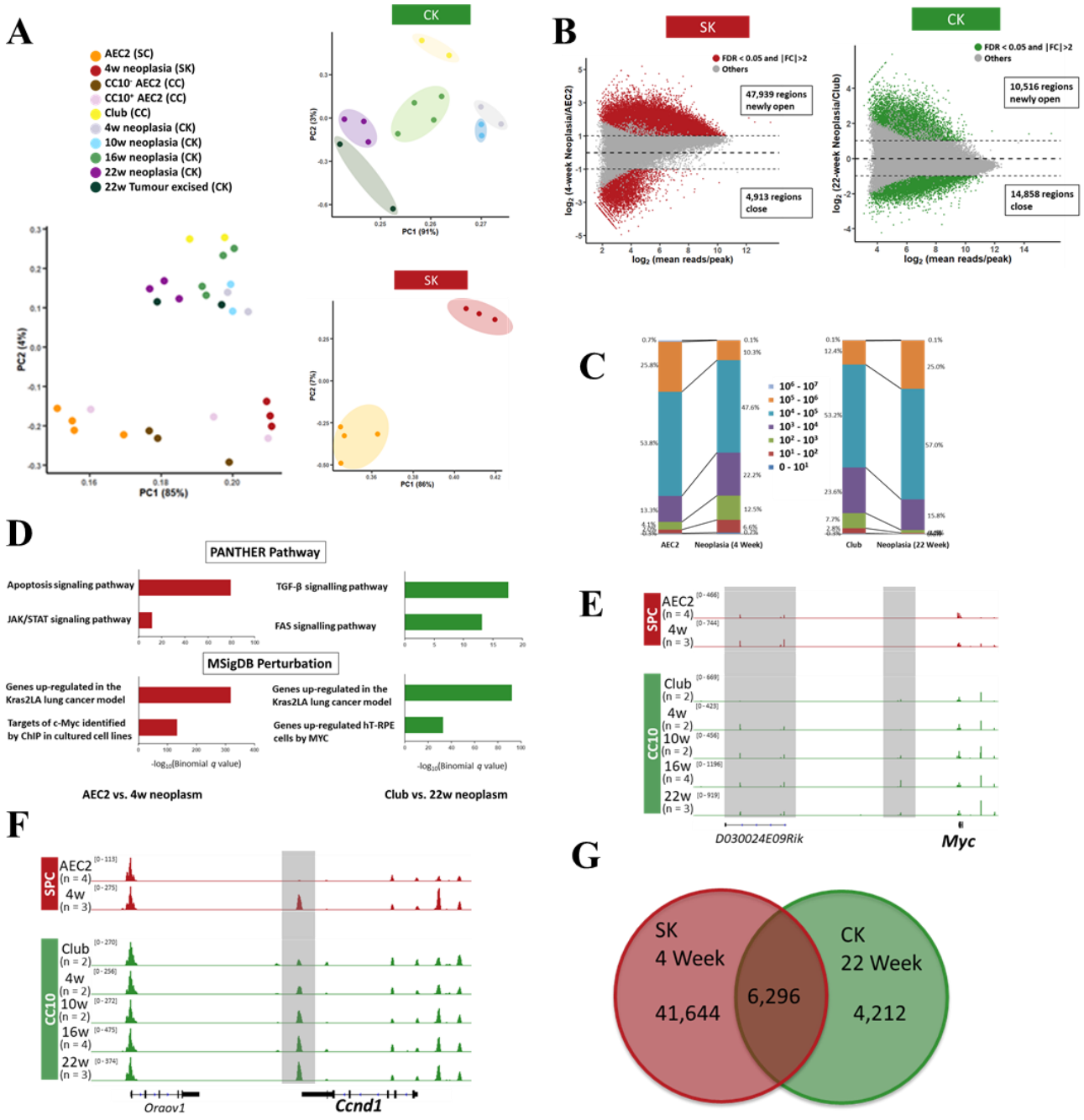

A
AEC2 (SC)
4w neoplasia (SK)
CC10- AEC2 (CC)
CC10+ AEC2 (CC)
Club (CC)
4w neoplasia (CK)
10w neoplasia (CK)
16w neoplasia (CK)
22w neoplasia (CK)
22w Tumour excised (CK)
CK
SK
PC1 (85%)
PC2 (4%)
PC1 (91%)
PC2 (3%)
PC1 (86%)
PC2 (7%)
B
SK
CK
FDR < 0.05 and |FC|>2
Others
47,939 regions newly open
4,913 regions close
10,516 regions newly open
14,858 regions close
log2 (4-week Neoplasia/AEC2)
log2 (22-week Neoplasia/Club)
log2 (mean reads/peak)
C
10^6 - 10^7
10^5 - 10^6
10^4 - 10^5
10^3 - 10^4
10^2 - 10^3
10^1 - 10^2
0 - 10^1
AEC2
Neoplasia (4 Week)
Club
Neoplasia (22 Week)
D
PANTHER Pathway
Apoptosis signaling pathway
JAK/STAT signaling pathway
TGF-β signalling pathway
FAS signalling pathway
MSigDB Perturbation
Genes up-regulated in the Kras2LA lung cancer model
Targets of c-Myc identified by ChIP in cultured cell lines
Genes up-regulated in the Kras2LA lung cancer model
Genes up-regulated hT-RPE cells by MYC
-log10(Binomial q value)
AEC2 vs. 4w neoplasm
Club vs. 22w neoplasm
E
SPC
CC10
AEC2 (n = 4)
4w (n = 3)
Club (n = 2)
4w (n = 2)
10w (n = 2)
16w (n = 4)
22w (n = 3)
D030024E09Rik
Myc
F
SPC
CC10
AEC2 (n = 4)
4w (n = 3)
Club (n = 2)
4w (n = 2)
10w (n = 2)
16w (n = 4)
22w (n = 3)
Oraov1
Ccnd1
G
SK 4 Week
41,644
6,296
CK 22 Week
4,212

Figure 2. Mutant KRAS remodels the open chromatin landscape

A. Principal Component Analysis of ATAC-seq data from normal and neoplastic cells collected from SC (AEC2), CC (Club, CC10+ AEC2 and CC10- AEC2), SK (4w) and CK (4w, 10w, 16w and 22w) mouse lungs.
B. MA-plot showing the differential expression of accessible chromatin regions in SK (4w vs AEC2) and CK (22w vs Club) models.
C. Plot showing the distribution of differential open chromatin regions in SK (4w vs AEC2) and CK (22w vs Club) models by their distance from the nearest TSS.
D. GREAT analysis of newly open regions in CK and SK models.
E. ATAC-seq signal tracks at the *Ccnd1* locus.
F. ATAC-seq signal tracks at the *Myc* locus.
G. Venn diagram showing the overlap of newly open regions in SK (4w) and CK (22w) models.

Figure 3

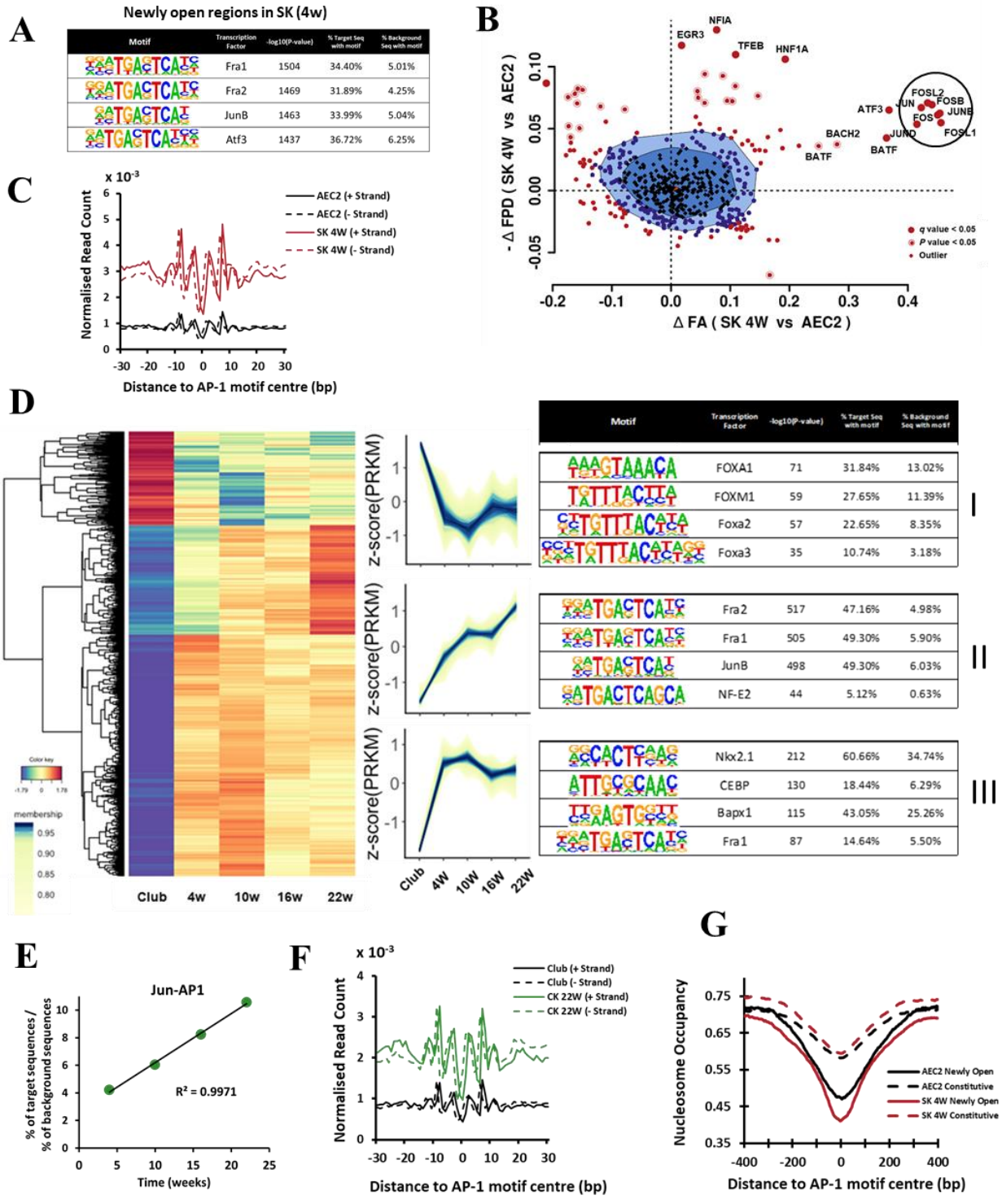

A
Newly open regions in SK (4w)
Motif | Transcription Factor | -log10(P-value) | % Target Seq with motif | % Background Seq with motif
Fra1 | 1504 | 34.40% | 5.01%
Fra2 | 1469 | 31.89% | 4.25%
JunB | 1463 | 33.99% | 5.04%
Atf3 | 1437 | 36.72% | 6.25%
B
NFIA
EGR3
TFEB
HNF1A
ATF3
FOSL2
JUN
FOSB
JUNB
FOS
FOSL1
BACH2
JUND
BATF
BATF
- Δ FPD (SK 4W vs AEC2)
Δ FA (SK 4W vs AEC2)
q value < 0.05
P value < 0.05
Outlier
C
x 10-3
Normalised Read Count
AEC2 (+ Strand)
AEC2 (- Strand)
SK 4W (+ Strand)
SK 4W (- Strand)
Distance to AP-1 motif centre (bp)
D
z-score(PRKM)
Club
4w
10w
16w
22w
Color key
membership
Motif | Transcription Factor | -log10(P-value) | % Target Seq with motif | % Background Seq with motif
FOXA1 | 71 | 31.84% | 13.02%
FOXM1 | 59 | 27.65% | 11.39%
Foxa2 | 57 | 22.65% | 8.35%
Foxa3 | 35 | 10.74% | 3.18%
I
Fra2 | 517 | 47.16% | 4.98%
Fra1 | 505 | 49.30% | 5.90%
JunB | 498 | 49.30% | 6.03%
NF-E2 | 44 | 5.12% | 0.63%
II
Nkx2.1 | 212 | 60.66% | 34.74%
CEBP | 130 | 18.44% | 6.29%
Bapx1 | 115 | 43.05% | 25.26%
Fra1 | 87 | 14.64% | 5.50%
III
E
Jun-AP1
% of target sequences / % of background sequences
R² = 0.9971
Time (weeks)
F
x 10-3
Normalised Read Count
Club (+ Strand)
Club (- Strand)
CK 22W (+ Strand)
CK 22W (- Strand)
Distance to AP-1 motif centre (bp)
G
Nucleosome Occupancy
AEC2 Newly Open
AEC2 Constitutive
SK 4W Newly Open
SK 4W Constitutive
Distance to AP-1 motif centre (bp)

Figure 3. Newly open regions in neoplastic cells are linked to AP-1 activity

A. TF motifs enriched in the newly open regions (with $\log_2$ FC > 2) of SK neoplastic cells.
B. BaGFoot analysis of ATAC-seq data from AEC2 and 4w SK neoplastic cells. AP-1 transcription factors are circled.
C. Plot showing the normalised ATAC-seq read count of SK 4w and AEC2 cells at the AP-1 motif sites present in differentially accessible regions.
D. Heatmap shows the differentially accessible regions on the chromatin of Club and CK neoplastic cells at 4, 10, 16- and 22-weeks post tamoxifen injection. Hierarchical clustering of these regions yields three groups with distinct temporal dynamics. The tables on the right show the top enriched TF motifs found in the respective cluster.
E. Plot showing the ratio of the percentage of newly open regions in CK neoplasia containing an AP-1 motif to that in the background sequences.
F. Plot showing the normalised ATAC-seq read count of CK 22w and Club cells at the AP-1 motif sites present in differentially accessible regions.
G. Plot showing nucleosome occupancy computed from ATAC-seq of AEC2 and SK 4w neoplastic cells at AP-1 motif sites in constitutive and newly open peaks (with $\log_2$ FC > 2).

Figure 4

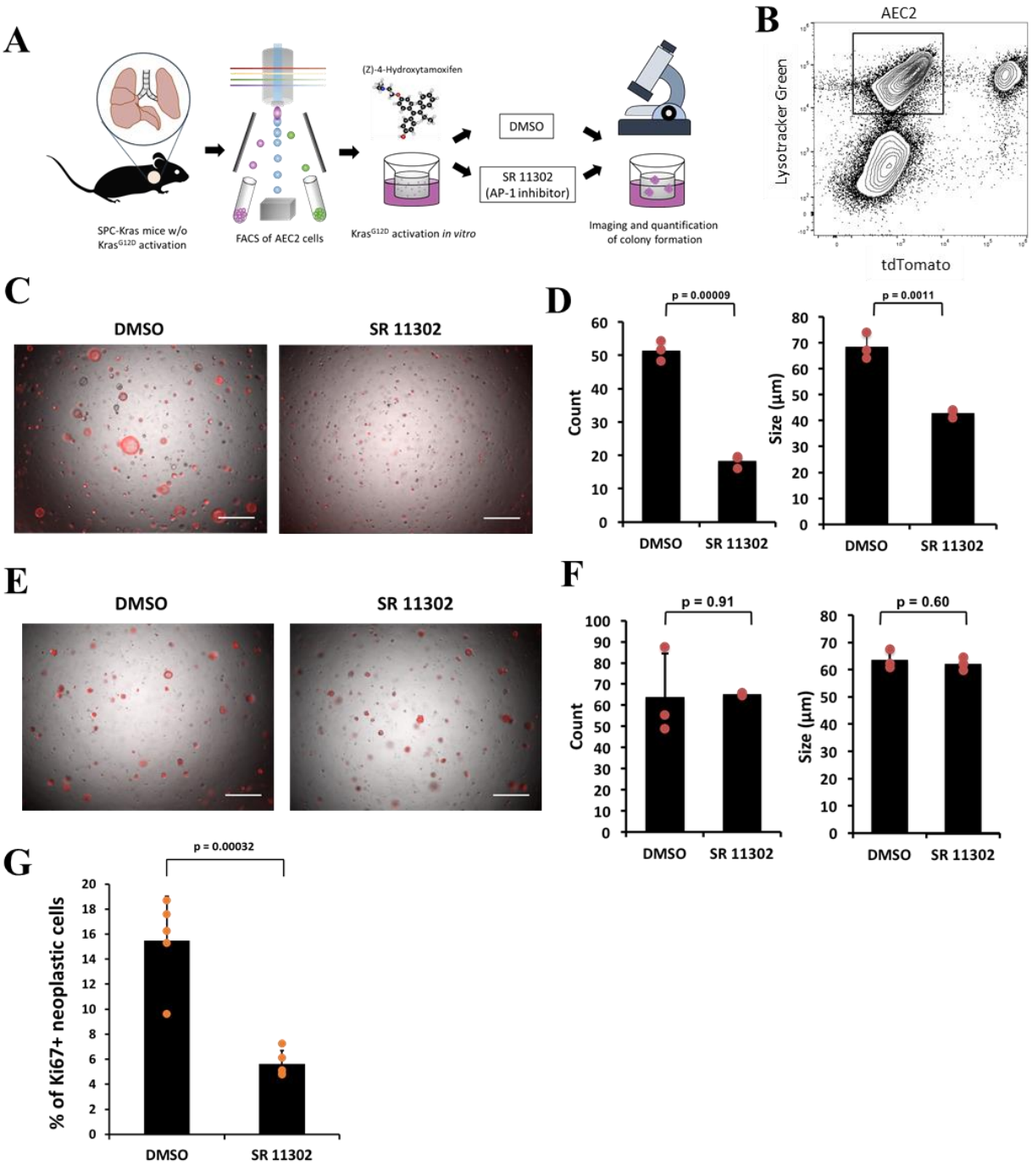

A
SPC-Kras mice w/o
KrasG12D activation
FACS of AEC2 cells
(Z)-4-Hydroxytamoxifen
KrasG12D activation in vitro
DMSO
SR 11302
(AP-1 inhibitor)
Imaging and quantification
of colony formation
B
AEC2
Lysotracker Green
tdTomato
C
DMSO
SR 11302
D
p = 0.00009
Count
DMSO
SR 11302
p = 0.0011
Size (µm)
DMSO
SR 11302
E
DMSO
SR 11302
F
p = 0.91
Count
DMSO
SR 11302
p = 0.60
Size (µm)
DMSO
SR 11302
G
p = 0.00032
% of Ki67+ neoplastic cells
DMSO
SR 11302

Figure 4. AP-1 activity is necessary for tumour initiation *in vitro*

A. Schematic depicting the in vitro tumour initiation assay.
B. Representative FACS plot showing the gate (tdTomato- and Lysotracker Green+) used to isolate normal AEC2 cells (with no 'leaky' CreER activity) from SK mice.

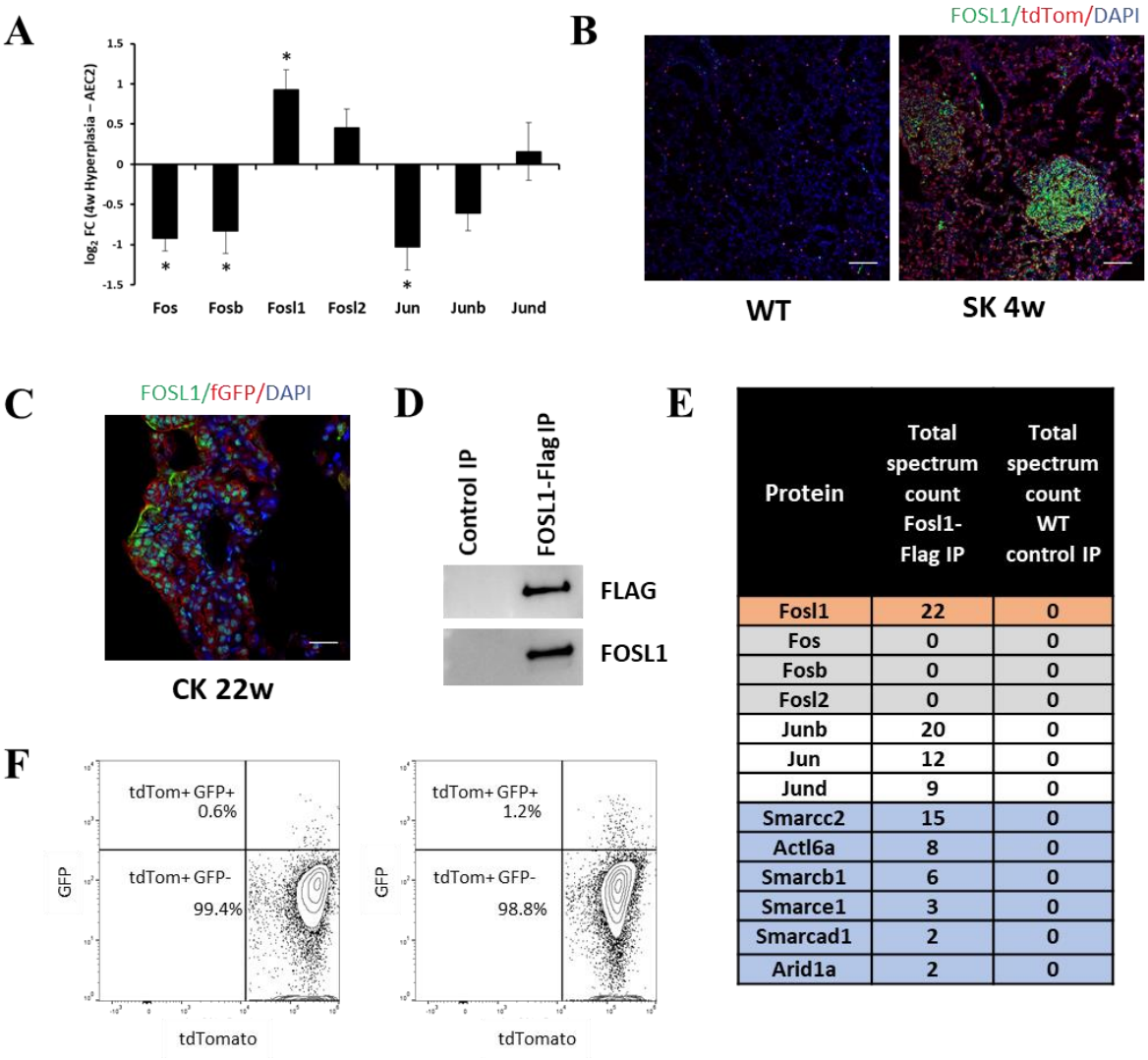
(C-D) Normal AEC2 cells are isolated from SK mice by FACS and plated in Matrigel containing 4-OHT to induce *Kras*^{G12D} expression for 24 hours. They were also treated with AP-1 inhibitor SR 11302 or DMSO (vehicle control) and colony (spheroid) formation ability of the cells was assayed after 8 days.

C. Representative immunofluorescent images showing the sphere formation ability. Scale bar represents 100 µm.
D. Quantification of spheroid size and number from 3 independent experiments.

(E-F) Normal AEC2 cells are isolated from SC mice by FACS and plated in Matrigel. They were treated with AP-1 inhibitor SR 11302 or DMSO (vehicle control) and colony (spheroid) formation ability of the cells was assayed after 8 days.

E. Representative immunofluorescent images showing the sphere formation ability. Scale bar represents 100 µm.
F. Quantification of spheroid size and number from 3 independent experiments.
G. SK mice were injected with two doses of tamoxifen and simultaneously treated with AP-1 inhibitor SR 11302 or DMSO (control). The drug treatment was continued for 3 weeks before the animals were euthanised and lungs harvested. Sections of the lungs were stained for Ki67 and tdTomato (lineage label). The

Figure 5



Protein	Total spectrum count Fosl1-Flag IP	Total spectrum count WT control IP
Fosl1	22	0
Fos	0	0
Fosb	0	0
Fosl2	0	0
Junb	20	0
Jun	12	0
Jund	9	0
Smarcc2	15	0
Actl6a	8	0
Smarcb1	6	0
Smarce1	3	0
Smarcad1	2	0
Arid1a	2	0

bar plot shows the percent of Ki67+ cells among the lineage labelled population in SR 11302 and DMSO treated mice. n=5 mice/condition.

Figure 5. Nucleosome remodelling SWI/SNF complex binds to AP-1 in *KRAS* mutant cells

A. Bar plot summarising the changes to expression levels of AP-1 genes in the SK model as detected by RNA-seq. * indicates statistical significance (q value < 0.05 and |FC| > 1.5).
B. Representative immunostaining images showing the expression of *Fosl1* (green) in SC (WT) and 4-week SK lungs. *Sftpc*-CreER lineage labelled cells are identified by tdTomato (red) expression. DAPI (blue) shows nuclei and the scale bar represents 100 μm.
C. Representative immunostaining image showing the expression of *Fosl1* (green) in 22-week CK lungs. *Scgb1a1*-CreER lineage labelled cells are identified by fGFP (red) expression. DAPI (blue) shows nuclei and the scale bar represents 25 μm.

(D-E) FOSL1 was immunoprecipitated from A549-Fosl1-Flag cell-line. A549 WT was used as control.

D. Representative Western Blot image showing the detection of Flag tag.
E. AP-1 and SWI/SNF complex proteins co-immunoprecipitated with FOSL1 as identified by Mass Spectrometry.
F. FACS plots showing the expression of GFP in lineage labelled (tdTomato+) cells from two SFK mice at 4 weeks post Cre induction.

Figure 6

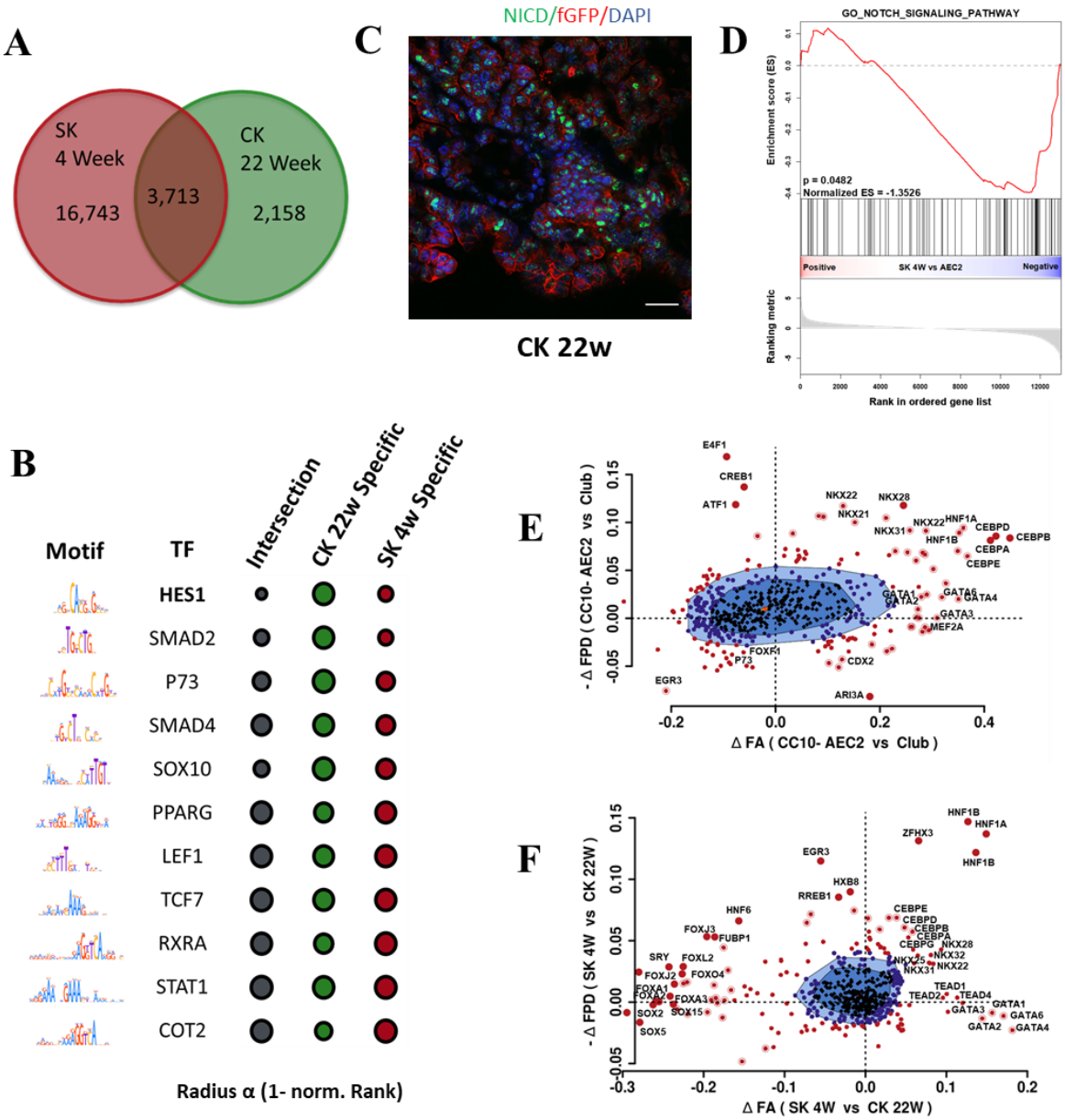

A
SK
4 Week
16,743
3,713
CK
22 Week
2,158
C
NICD/fGFP/DAPI
CK 22w
D
GO_NOTCH_SIGNALING_PATHWAY
Enrichment score (ES)
p = 0.0482
Normalized ES = -1.3526
Positive
SK 4W vs AEC2
Negative
Ranking metric
Rank in ordered gene list
B
Motif
TF
Intersection
CK 22w Specific
SK 4w Specific
HES1
SMAD2
P73
SMAD4
SOX10
PPARG
LEF1
TCF7
RXRA
STAT1
COT2
Radius α (1- norm. Rank)
E
- Δ FPD (CC10- AEC2 vs Club)
Δ FA (CC10- AEC2 vs Club)
F
- Δ FPD (SK 4W vs CK 22W)
Δ FA (SK 4W vs CK 22W)

Figure 6. Signature of the cell of origin in neoplastic cells

A. Venn diagram showing the number of overlapping newly open regions that harbour an AP-1 motif in CK 22w and SK 4w neoplastic cells.
B. Centrimo package was used to find TF motif enrichment in the neighbourhood of putative AP-1 binding sites in the newly open regions that are specific to CK 22w or SK 4w neoplastic cells and the regions that are common as shown in (A). Enriched motifs were rank ordered in each case and few of the top hits are depicted in a bubble plot.
C. Representative immunofluorescence image showing the level of activated Notch1 or NICD (green) in 22w CK lungs. Lineage label fGFP is shown in red and DAI (blue) represents nuclei. Scale bar represents 25 μm.
D. Gene set enrichment analysis of RNA-seq data shows that genes upregulated in Kras2LA lung cancer model (Sweet-Cordero et al., 2005) are enriched in SK 4w cells when compared to AEC2.
E. BaGFoot analysis of ATAC-seq data from CC10- AEC2 and Club cells.
F. BaGFoot analysis of ATAC-seq data from 4w SK and 22w CK neoplastic cells.

Figure 7

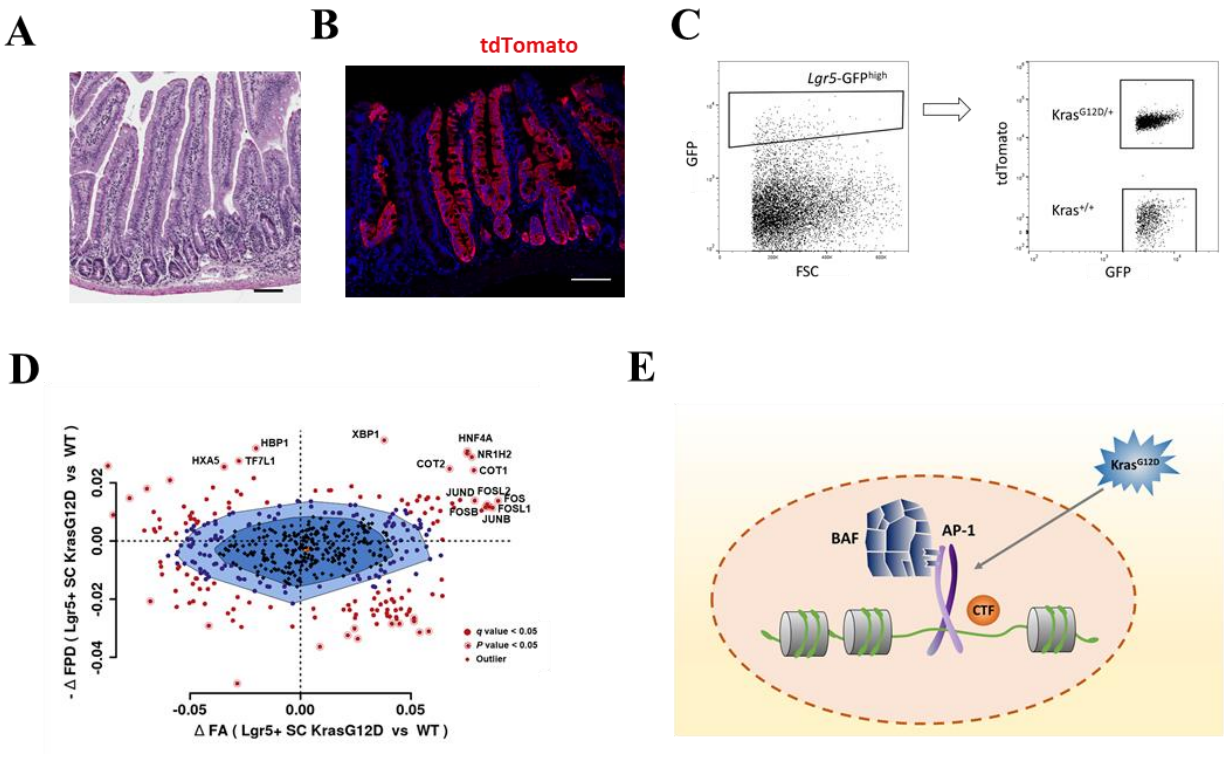

A
B
tdTomato
C
Lgr5-GFPhigh
GFP
FSC
tdTomato
KrasG12D/+
Kras+/+
GFP
D
HBP1
XBP1
HNF4A
HXA5
TF7L1
NR1H2
COT2
COT1
JUND
FOSL2
FOS
FOSB
FOSL1
JUNB
q value < 0.05
P value < 0.05
Outlier
- Δ FPD (Lgr5+ SC KrasG12D vs WT)
Δ FA (Lgr5+ SC KrasG12D vs WT)
0.02
0.00
-0.02
-0.04
-0.05
0.00
0.05
E
KrasG12D
BAF
AP-1
CTF

Figure 7. AP-1 signature in *Kras*^{G12D}-expressing *Lgr5*+ stem cells of the intestine

(A-D) *Lgr5*-EGFP-CreER; LSL-*Kras*^{G12D}; *R26R*-tdTomato mice were injected 4 doses of tamoxifen and their intestines were analysed after 5 weeks.

A. Representative H&E stained image of the intestine. Scale bar represents 100 µm.
B. Representative immunofluorescence image showing ribbons of tdTomato+ (red) cells. DAPI (blue) represents nuclei. Scale bar represents 100 µm.
C. Representative FACS plot showing the gates used to sort *Kras* mutant and WT *Lgr5*+ stem cells.
D. BaGFoot analysis of ATAC-seq data from *Kras* mutant and WT *Lgr5*+ stem cells.
E. Summary. AP-1 complex containing of FOSL1, in cooperation with a collaborating transcription factor (CTF), directly binds to the SWI/SNF complex and directs to regions in the genome to displace nucleosomes and increase accessibility.

Figure S1

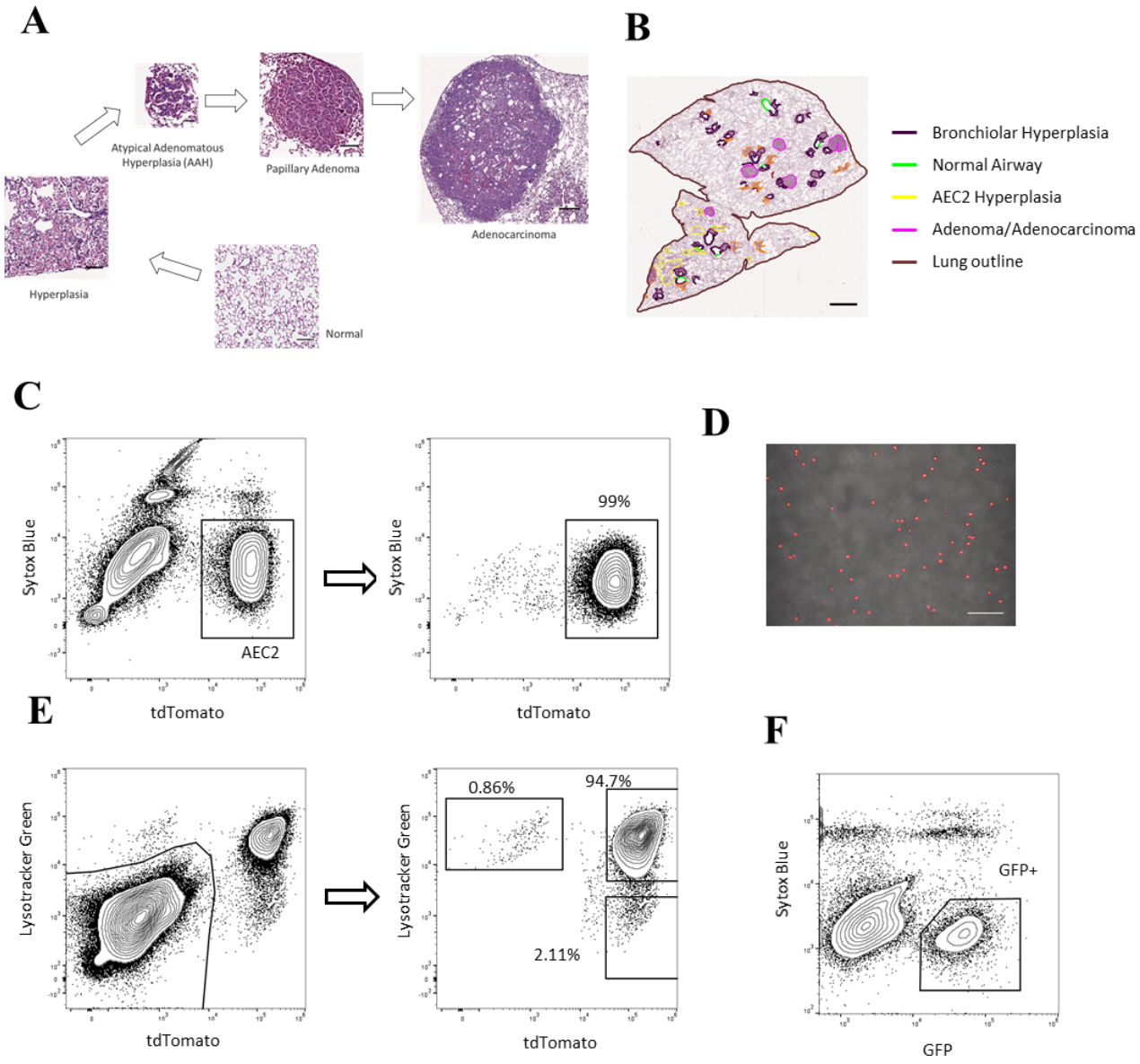

A
Atypical Adenomatous Hyperplasia (AAH)
Papillary Adenoma
Adenocarcinoma
Hyperplasia
Normal
B
Bronchiolar Hyperplasia
Normal Airway
AEC2 Hyperplasia
Adenoma/Adenocarcinoma
Lung outline
C
Sytox Blue
AEC2
tdTomato
99%
D
E
Lysotracker Green
tdTomato
0.86%
94.7%
2.11%
F
Sytox Blue
GFP+
GFP

Figure S1. Neoplastic progression in SK and CK lung adenocarcinoma models

A. Representative H&E images showing the various stages of lung tumorigenesis in the SK model. Normal lung is from an SC mouse (scale bar: 100 µm). Hyperplastic region (scale bar: 100 µm), AAH (scale bar: 50 µm), papillary adenoma (scale bar: 100 µm) and adenocarcinoma (scale bar: 300 µm) are from SK mice at 4 weeks post tamoxifen doses.
B. Representative H&E image of a lung section from a CK mouse at 16 weeks post tamoxifen doses with various histopathological features annotated. Scale bar represents 1 mm.
C. Representative FACS plot showing the purity of sorted AEC2 cells from SC mice.
D. Image showing the sorted tdTomato+ AEC2 cells. Scale bar represents 100 µm.
E. Representative FACS plots of Lysotracker stained lung cells from SC mice. In the plot on the left, 'Not-gate' is used to eliminate GFP- Lysotracker- population and the rest are visualised on the right. About 95% of the cells are both Lysotracker+ and tdTomato+ indicating that Lysotracker can be used to isolate AEC2 cells.
F. Representative FACS plot showing the gate used to sort *Scgb1a1*-CreER lineage labelled (GFP+) cells from CK mice.

Figure S2

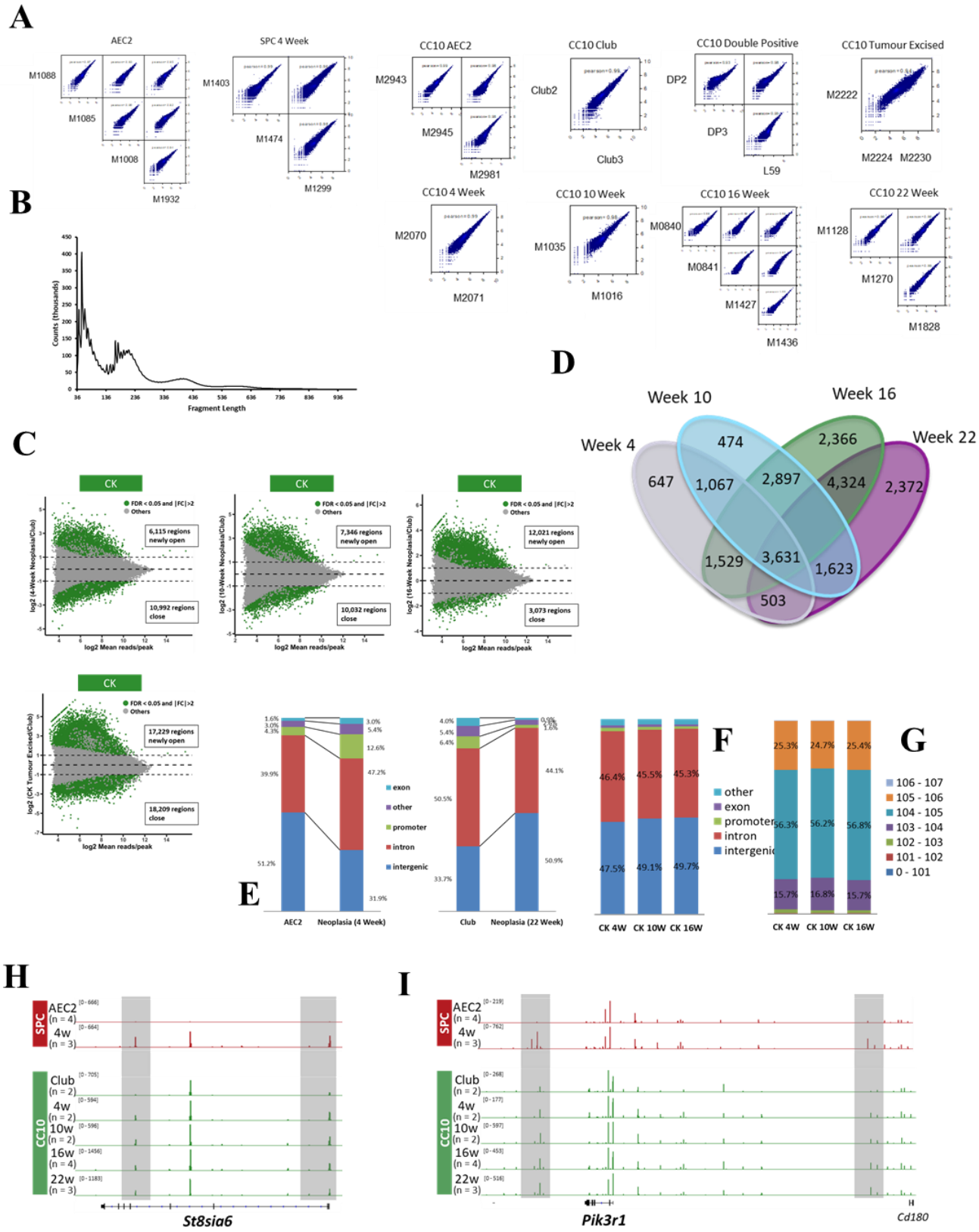

A
AEC2
M1088
M1085
M1008
M1932
SPC 4 Week
M1403
M1474
M1299
CC10 AEC2
M2943
M2945
M2981
CC10 Club
Club2
Club3
CC10 Double Positive
DP2
DP3
L59
CC10 Tumour Excised
M2222
M2224 M2230
B
Counts (thousands)
Fragment Length
CC10 4 Week
M2070
M2071
CC10 10 Week
M1035
M1016
CC10 16 Week
M0840
M0841
M1427
M1436
CC10 22 Week
M1128
M1270
M1828
C
CK
FDR < 0.05 and |FC|>2
Others
6,115 regions newly open
10,992 regions close
log2 (4-Week Neoplasia/Club)
log2 Mean reads/peak
7,346 regions newly open
10,032 regions close
log2 (10-Week Neoplasia/Club)
12,021 regions newly open
3,073 regions close
log2 (16-Week Neoplasia/Club)
17,229 regions newly open
18,209 regions close
log2 (CK Tumour Excised/Club)
D
Week 10
Week 16
Week 4
Week 22
647
474
1,067
2,897
2,366
4,324
2,372
1,529
3,631
1,623
503
E
AEC2
Neoplasia (4 Week)
exon
other
promoter
intron
intergenic
Club
Neoplasia (22 Week)
CK 4W
CK 10W
CK 16W
F
other
exon
promoter
intron
intergenic
G
106 - 107
105 - 106
104 - 105
103 - 104
102 - 103
101 - 102
0 - 101
H
SPC
AEC2 (n = 4)
4w (n = 3)
CC10
Club (n = 2)
4w (n = 2)
10w (n = 2)
16w (n = 4)
22w (n = 3)
St8sia6
I
Pik3r1
Cd180

Figure S2. Mutant KRAS remodels the open chromatin landscape

A. Plots showing the correlation between the biological replicates in the ATAC-seq data from AEC2, Neoplasia SK (4w), Club, CC10+ AEC2 and Neoplasia CK (4w,10w,16w and 22w).
B. Representative plot showing the fragment length (or insert size) distribution in ATAC-seq data.
C. MA-plot showing the differential expression of accessible chromatin regions in CK 4w, 10w, 16w and tumours excised from 22w samples when compared to Club cells.
D. Venn diagram showing the overlap of newly open regions in CK neoplastic cells.
E. Plot showing the distribution of differentially accessible regions in SK (4w vs AEC2) and CK (22w vs Club) models based on their annotation.
F. Plot showing the distribution of differentially accessible regions in CK (4w, 10w and 16w) neoplastic cells based on their annotation.
G. Plot showing the distribution of newly open regions in CK (4w, 10w and 16w) neoplastic cells based on their distance from the nearest TSS.
H. ATAC-seq signal tracks at the *St8sia6* locus.
I. ATAC-seq signal tracks at the *Pik3r1* locus.

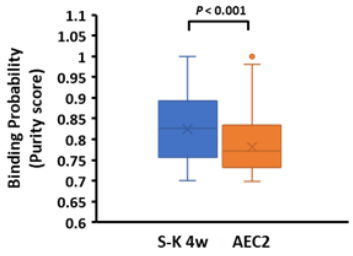
Figure S3

A

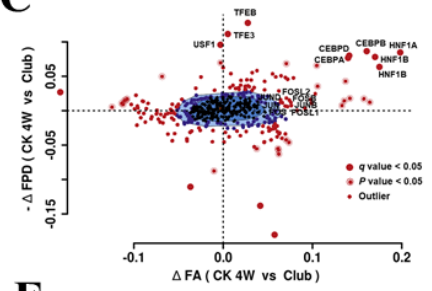
Newly closed regions in SK (4w)

Motif	Transcription Factor	-log10(P-value)	% Target Seq with motif	% Background Seq with motif
	ERG	221	45.35%	18.25%
	Etv2	190	31.48%	10.37%
	ETS1	187	33.00%	11.42%
	Foxo3	191	19.36%	6.07%

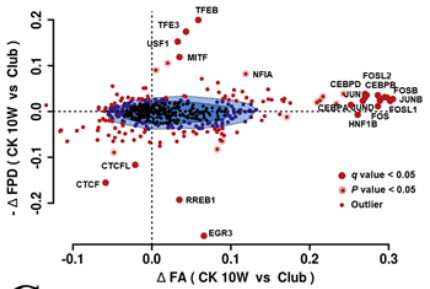
B



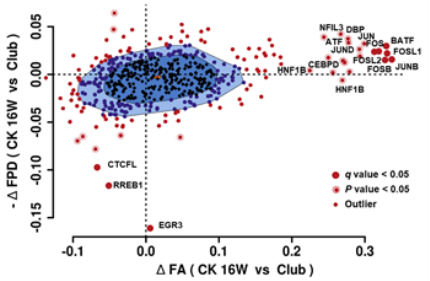
C



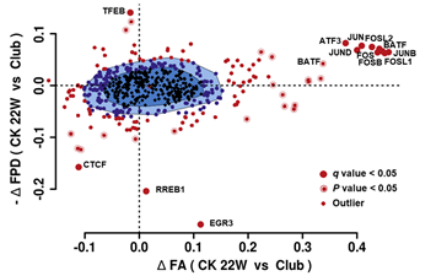
D



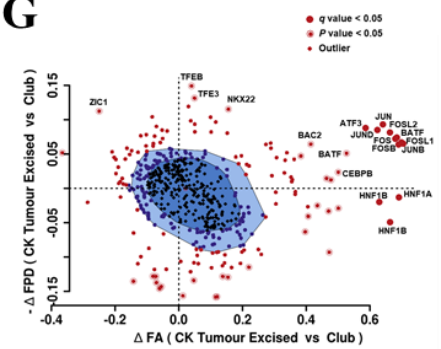
E



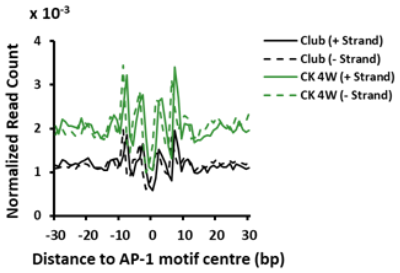
F



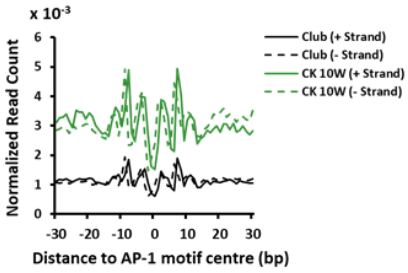
G



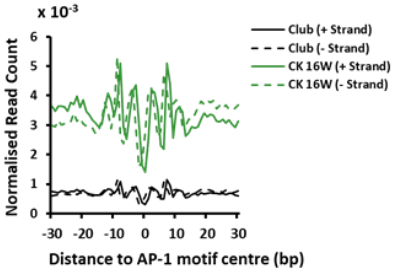
H



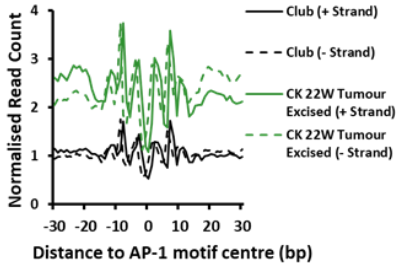
I



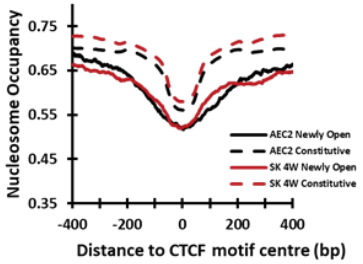
J



K



L



M

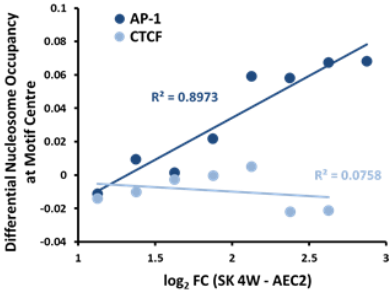


Figure S3. Newly open regions in neoplastic cells are linked to AP-1 activity

A. TF motifs enriched in the newly closed regions (with $\log_2$ FC < -2) of SK neoplastic cells.
B. Box and Whisker plot showing the difference in binding probabilities of AP-1 complex (identified by FOSL1 motif) in AEC2 and 4w SK neoplastic cells as predicted by PIQ algorithm using ATAC-seq data.
C. BaGFoot analysis of ATAC-seq data from Club and 4w CK neoplastic cells.
D. BaGFoot analysis of ATAC-seq data from Club and 10w CK neoplastic cells.
E. BaGFoot analysis of ATAC-seq data from Club and 16w CK neoplastic cells.
F. BaGFoot analysis of ATAC-seq data from Club and 22w CK neoplastic cells.
G. BaGFoot analysis of ATAC-seq data from Club and tumours excised from 22w CK mouse lungs.
H. Plot showing the normalised ATAC-seq read count of CK 4w and Club cells at the AP-1 motif sites present in differentially accessible regions.
I. Plot showing the normalised ATAC-seq read count of CK 10w and Club cells at the AP-1 motif sites present in differentially accessible regions.
J. Plot showing the normalised ATAC-seq read count of CK 16w and Club cells at the AP-1 motif sites present in differentially accessible regions.
K. Plot showing the normalised ATAC-seq read count of CK 22w (excised tumours) and Club cells at the AP-1 motif sites present in differentially accessible regions.

L. Plot showing nucleosome occupancy computed from ATAC-seq of AEC2 and SK 4w neoplastic cells at CTCF motif sites in constitutive and newly open peaks (with $\log_2$ FC > 2).

M. Differential nucleosome occupancy at transcription factor motif centre (AP-1 and CTCF) plotted against the log fold change of newly open regions in SK 4w neoplastic cells when compared to AEC2 cells.

Figure S4

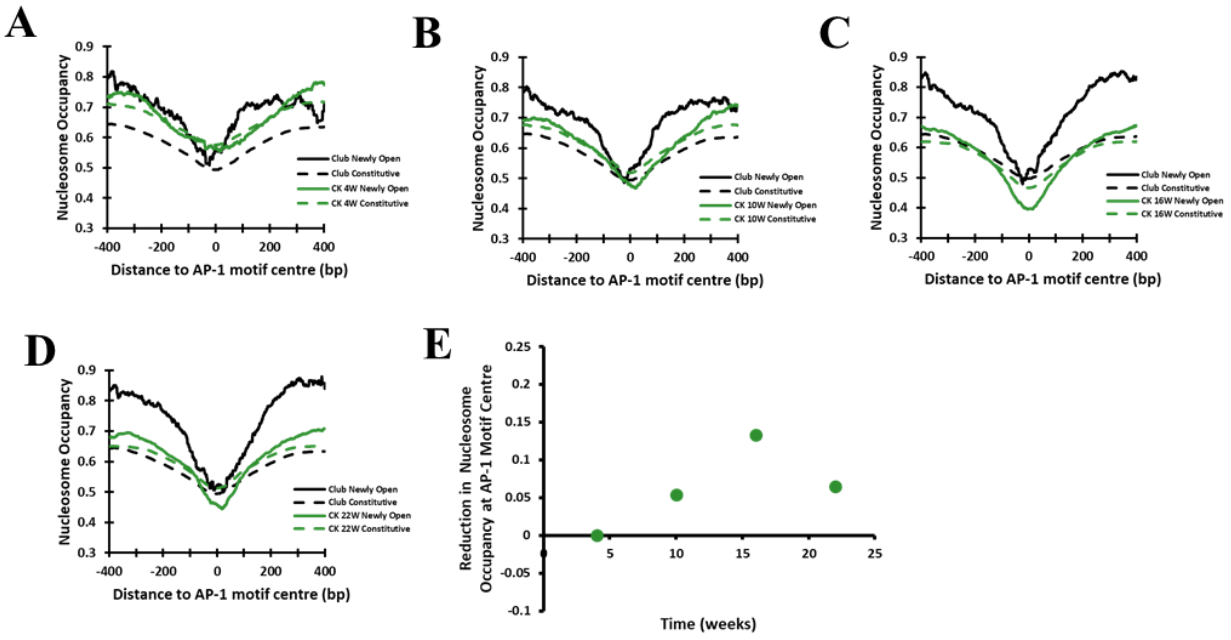

A
B
C
D
E
Nucleosome Occupancy
Distance to AP-1 motif centre (bp)
Club Newly Open
Club Constitutive
CK 4W Newly Open
CK 4W Constitutive
CK 10W Newly Open
CK 10W Constitutive
CK 16W Newly Open
CK 16W Constitutive
CK 22W Newly Open
CK 22W Constitutive
Reduction in Nucleosome Occupancy at AP-1 Motif Centre
Time (weeks)

Figure S4. Newly open regions in neoplastic cells are linked to AP-1 activity

A. Plot showing nucleosome occupancy computed from ATAC-seq of Club and CK 4w neoplastic cells at AP-1 motif sites in constitutive and newly open peaks (with $\log_2$ FC > 2).

B. Plot showing nucleosome occupancy computed from ATAC-seq of Club and CK 10w neoplastic cells at AP-1 motif sites in constitutive and newly open peaks (with $\log_2$ FC > 2).

C. Plot showing nucleosome occupancy computed from ATAC-seq of Club and CK 16w neoplastic cells at AP-1 motif sites in constitutive and newly open peaks (with $\log_2$ FC > 2).

D. Plot showing nucleosome occupancy computed from ATAC-seq of Club and CK 22w neoplastic cells at AP-1 motif sites in constitutive and newly open peaks (with $\log_2$ FC > 2).

E. Average reduction in nucleosome occupancy in the newly open peaks (with $\log_2$ FC > 2) at the AP-1 motif centre plotted against the time at which tumour cells were harvested from the CK model.

Figure S5

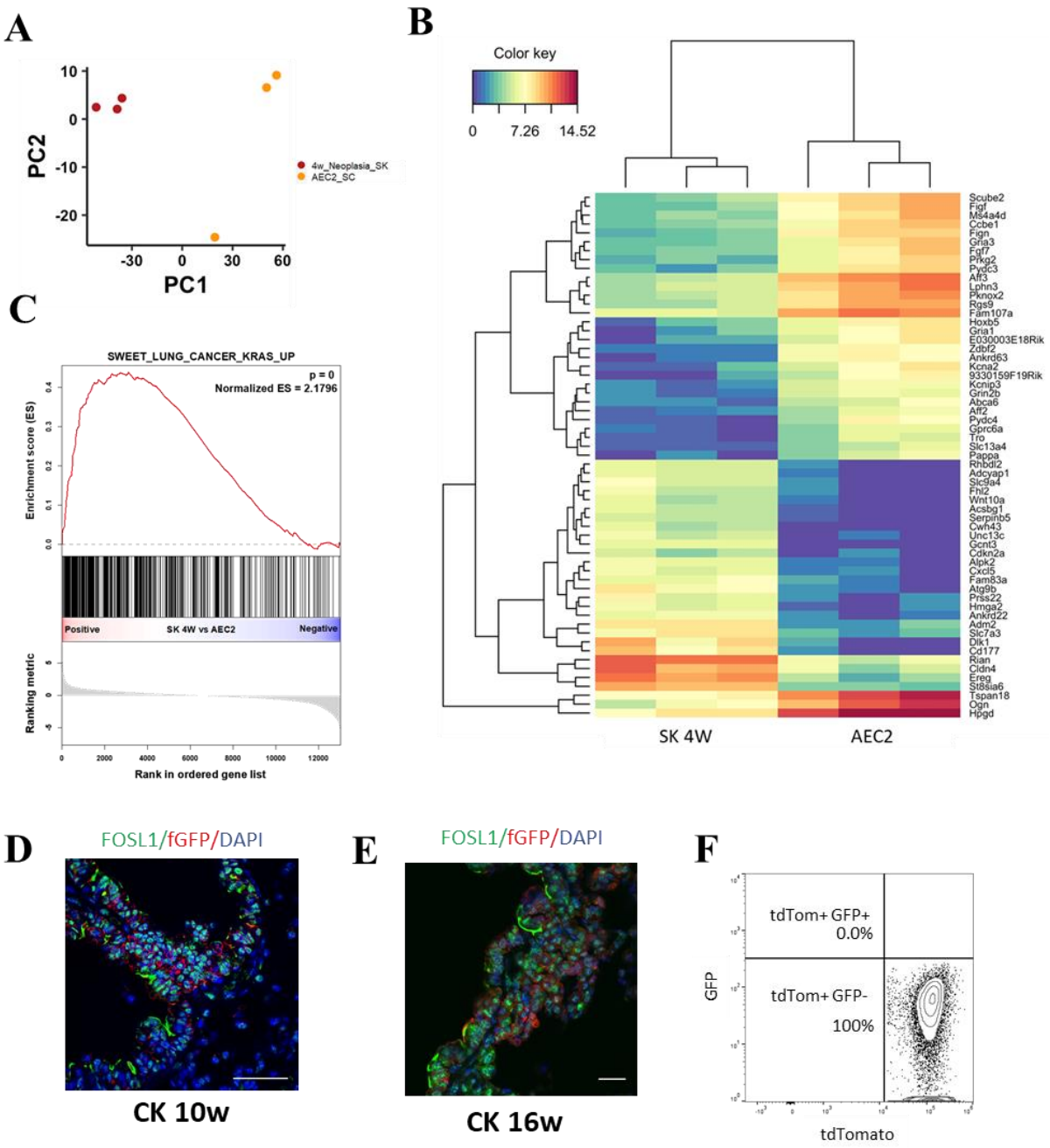

A
PC2
PC1
4w_Neoplasia_SK
AEC2_SC
B
Color key
0
7.26
14.52
SK 4W
AEC2
C
SWEET_LUNG_CANCER_KRAS_UP
p = 0
Normalized ES = 2.1796
Enrichment score (ES)
Positive
SK 4W vs AEC2
Negative
Rankling metric
Rank in ordered gene list
D
FOSL1/fGFP/DAPI
CK 10w
E
FOSL1/fGFP/DAPI
CK 16w
F
tdTom+ GFP+
0.0%
tdTom+ GFP-
100%
GFP
tdTomato

Figure S5. Nucleosome remodelling SWI/SNF complex binds to AP-1 in *KRAS* mutant cells

A. Principal Component Analysis of RNA-seq data from AEC2 and 4-week SK neoplastic cells.
B. Heatmap showing the differentially expressed genes in the RNA-seq data from (A).
C. Gene set enrichment analysis of RNA-seq data shows that Notch activity has reduced in SK 4w neoplastic cells when compared to AEC2.
D. Representative immunostaining image showing the expression of *Fosl1* (green) in 10-week CK lungs. *Scgb1a1*-CreER lineage labelled cells are identified by fGFP (red) expression. DAPI (blue) shows nuclei and the scale bar represents 50 µm.
E. Representative immunostaining image showing the expression of *Fosl1* (green) in 16-week CK lungs. *Scgb1a1*-CreER lineage labelled cells are identified by fGFP (red) expression. DAPI (blue) shows nuclei and the scale bar represents 25 µm.
F. FACS plot showing lineage labelled (tdTomato+) cells from an SK mouse at 4 weeks post Cre induction. These cells are GFP- and are used as a control to set the gates shown in Fig. 5F.

Figure S6

A

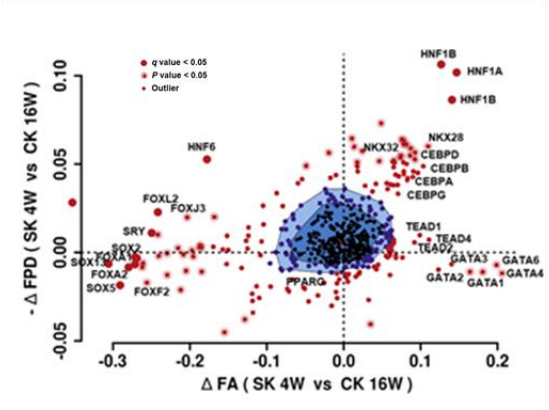


Figure S6. Signature of the cell of origin in neoplastic cells

A. BaGFoot analysis of ATAC-seq data from 4w SK and 16w CK neoplastic cells.

Supplementary Table 1

Sample ID	Mouse ID	Condition	Sequencer	Read length	# of cells used for library prep	Fraction of Reads in Peak	Total Read Count from Sequencer	Final Read Count after filtering	NRF	PBC1	PBC2	Raw peaks	TSS Enrichment score
KLP-L42	M1403	SK Week 4	HiSeq 2500	50 bp PE	75000	0.368954	146613352	34210610	0.60928	0.58541	2.2176	103484	5.130963432
KLP-L48	M1474	SK Week 4	HiSeq 2500	50 bp PE	75000	0.394383	138151542	39640606	0.6586	0.63811	2.53915	119078	9.801743886
KLP-L54	M1299	SK Week 4	HiSeq 2500	50 bp PE	75000	0.434968	169843498	67757514	0.77241	0.76729	4.10525	130859	13.57868544
KLP-L49	M1088	SC AEC2	HiSeq 2500	50 bp PE	150000	0.154434	137414646	36505964	0.65092	0.63031	2.49307	73902	4.441256434
KLP-L51	M1085	SC AEC2	HiSeq 2500	50 bp PE	100000	0.22148	152081604	38537158	0.59242	0.5636	2.08537	96882	5.811432324
KLP-L55	M1008	SC AEC2	HiSeq 2500	50 bp PE	100000	0.156119	141856934	23894248	0.4929	0.49075	2.10034	69986	4.583292869
KLP-L57	M1932	SC AEC2	HiSeq 2500	50 bp PE	75000	0.124499	127462490	37433118	0.61729	0.60918	2.55816	61254	6.349941756
M2070	M2070	CK Week 4	HiSeq 2500	50 bp PE	70000	0.402799	133362302	62031938	0.7649	0.75867	3.96642	118366	16.82462577
M2071	M2071	CK Week 4	HiSeq 2500	50 bp PE	70000	0.380704	111463818	59871442	0.72808	0.71749	3.38321	139128	16.97962636
KLP-L50	M1035	CK Week 10	HiSeq 2500	50 bp PE	75000	0.356775	110365612	37175066	0.84129	0.84156	6.2235	112147	9.241826508
KLP-L52	M1016	CK Week 10	HiSeq 2500	50 bp PE	75000	0.424036	218922992	88178538	0.76703	0.76247	4.0419	128294	15.21026543
KLP-L56	M0840	CK Week 16	HiSeq 2500	50 bp PE	100000	0.388175	118783754	53379402	0.82256	0.8222	5.54085	122765	15.54407307
KLP-L58	M0841	CK Week 16	HiSeq 2500	50 bp PE	75000	0.365713	241339734	62059598	0.47529	0.441	1.67113	112271	15.9125616
KLP-L60	M1427	CK Week 16	HiSeq 2500	50 bp PE	75000	0.482454	203645590	69559950	0.58047	0.557	2.11355	145830	17.667035
KLP-L61	M1436	CK Week 16	HiSeq 2500	50 bp PE	75000	0.493048	191251348	69900552	0.57941	0.55413	2.08647	140536	19.54331736
M1128	M1128	CK Week 22	HiSeq 2500	50 bp PE	75000	0.300112	163530290	60320840	0.65128	0.63267	2.53992	134554	13.42337295
M1270	M1270	CK Week 22	HiSeq 2500	50 bp PE	75000	0.336502	163913418	78488464	0.72522	0.71412	3.31368	136966	17.02392704
M1828	M1828	CK Week 22	HiSeq 2500	50 bp PE	70000	0.38165	173566528	93875356	0.79248	0.78855	4.56707	150127	18.76467227
KLP-L59	Pooled	CC CC10+ AEC2	HiSeq 2500	50 bp PE		0.305946	257468046	36207806	0.29383	0.2869	1.71286	100448	9.560721802
Double_Positive_Mar1	Pooled	CC CC10+ AEC2	HiSeq 2500	50 bp PE	70000	0.435564	189048192	60706390	0.61989	0.59881	2.31166	124668	14.36444234
Double_Positive_May16	Pooled	CC CC10+ AEC2	HiSeq 2500	50 bp PE	75000	0.165323	132637472	43000228	0.58935	0.58519	2.4543	87871	9.393878171
Clara_Mar1	Pooled	CC Club	HiSeq 2500	50 bp PE	57000	0.480134	157243870	78682396	0.77374	0.76802	4.1473	143994	18.75150531
Clara_May16	Pooled	CC Club	HiSeq 2500	50 bp PE	75000	0.396618	128456874	60934224	0.77204	0.76803	4.22304	121083	18.30225684
M2943	M2943	CC CC10- AEC2	HiSeq 2500	50 bp PE	80000	0.238764	140072394	50578398	0.73865	0.73427	3.69149	99827	11.60791645
M2945	M2945	CC CC10- AEC2	HiSeq 2500	50 bp PE	80000	0.239471	136139378	38973054	0.69248	0.68805	3.16391	98452	10.43773046
M2981	M2981	CC CC10- AEC2	HiSeq 2500	50 bp PE	80000	0.334553	124356564	33965820	0.63185	0.62189	2.58649	102336	12.60990641
50B2-L2	M2222	CK Tumour Excised Week 22	NextSeq 500	75 bp PE	80000	0.477069	170295708	76016504	0.77737	0.77724	4.4086	139444	19.52335616
50B2-L3	M2224 and M2230	CK Tumour Excised Week 22	NextSeq 500	75 bp PE	50000	0.301116	218764668	78676994	0.65815	0.64352	2.65808	106505	15.20772029
M3719_Ctrl	M3719	Intestine Lgr5 Ctrl	NextSeq 500	75 bp PE	7000	0.245231	247701674	61742288	0.36051	0.31163	1.34684	128776	10.24676811
M3719_Kras	M3719	Intestine Lgr5 Kras	NextSeq 500	75 bp PE	7000	0.329769	219106580	61113392	0.38274	0.32851	1.35209	143892	13.18083592
M5771_Ctrl	M5771	Intestine Lgr5 Ctrl	NextSeq 500	75 bp PE	20000	0.179635	152299022	41005612	0.74216	0.73508	3.64196	111332	3.24458772
M5771_Kras	M5771	Intestine Lgr5 Kras	NextSeq 500	75 bp PE	20000	0.137442	219376966	63662004	0.64824	0.63522	2.63925	102343	3.747555776

REFERENCES

Aivado, M., Gynes, M., Gorelov, V., Schmidt, W., Röher, H., and Goretzki, P. (2000). " Field cancerization"--an additional phenomenon in development of colon tumors? K-ras codon 12 mutations in normal colonic mucosa of patients with colorectal neoplasms. Der Chirurg; Zeitschrift fur alle Gebiete der operativen Medizen *71*, 1230-1234; discussion 1234-1235.

Baek, S., Goldstein, I., and Hager, G.L. (2017). Bivariate genomic footprinting detects changes in transcription factor activity. Cell reports *19*, 1710-1722.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic acids research *37*, W202-W208.

Barkauskas, C.E., Cronce, M.J., Rackley, C.R., Bowie, E.J., Keene, D.R., Stripp, B.R., Randell, S.H., Noble, P.W., and Hogan, B. (2013). Type 2 alveolar cells are stem cells in adult lung. Journal of Clinical Investigation *123*, 3025-3036.

Barker, N., Ridgway, R.A., van Es, J.H., van de Wetering, M., Begthel, H., van den Born, M., Danenberg, E., Clarke, A.R., Sansom, O.J., and Clevers, H. (2009). Crypt stem cells as the cells-of-origin of intestinal cancer. Nature *457*, 608.

Bell, O., Tiwari, V.K., Thomä, N.H., and Schübeler, D. (2011). Determinants and dynamics of genome accessibility. Nature Reviews Genetics *12*, 554.

Buenrostro, J.D., Corces, M.R., Lareau, C.A., Wu, B., Schep, A.N., Aryee, M.J., Majeti, R., Chang, H.Y., and Greenleaf, W.J. (2018). Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. Cell *173*, 1535-1548.e1516.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature methods *10*, 1213.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E.H., Weissman, J.S., and Qi, L.S. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell *155*, 1479-1491.

Cox, A.D., Fesik, S.W., Kimmelman, A.C., Luo, J., and Der, C.J. (2014). Drugging the undruggable RAS: mission possible? Nature reviews Drug discovery *13*, 828.

Denny, S.K., Yang, D., Chuang, C.-H., Brady, J.J., Lim, J.S., Grüner, B.M., Chiou, S.-H., Schep, A.N., Baral, J., and Hamard, C. (2016). Nfib promotes metastasis through a widespread increase in chromatin accessibility. Cell *166*, 328-342.

der Velden, J.L., Bertoncello, I., and McQualter, J.L. (2013). LysoTracker is a marker of differentiated alveolar type II cells. Respiratory Research *14*, 1-7.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Eferl, R., Hoebertz, A., Schilling, A.F., Rath, M., Karreth, F., Kenner, L., Amling, M., and Wagner, E.F. (2004). The Fos-related antigen Fra-1 is an activator of bone matrix formation. The EMBO journal *23*, 2789-2799.

Eferl, R., and Wagner, E.F. (2003). AP-1: a double-edged sword in tumorigenesis. Nature Reviews Cancer *3*, 859.

Elangovan, I.M., Vaz, M., Tamatam, C.R., Potteti, H.R., Reddy, N.M., and Reddy, S.P. (2018). FOSL1 promotes Kras-induced lung cancer through amphiregulin and cell survival gene regulation. American journal of respiratory cell and molecular biology *58*, 625-635.

Erdoğan, Ö., Xie, L., Wang, L., Wu, B., Kong, Q., Wan, Y., and Chen, X. (2016). Proteomic dissection of LPS-inducible, PHF8-dependent secretome reveals novel roles of PHF8 in TLR4-induced acute inflammation and T cell proliferation. Scientific reports *6*, 24833.

Fanjul, A., Dawson, M.I., Hobbs, P.D., Jong, L., Cameron, J.F., Harlev, E., Graupner, G., Lu, X.-P., and Pfahl, M. (1994). A new class of retinoids with selective inhibition of AP-1 inhibits proliferation. Nature *372*, 107.

Ge, Y., Gomez, N.C., Adam, R.C., Nikolova, M., Yang, H., Verma, A., Lu, C.P.-J., Polak, L., Yuan, S., and Elemento, O. (2017). Stem cell lineage infidelity drives wound repair and cancer. Cell *169*, 636-650. e614.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. cell *144*, 646-674.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010a). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010b). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Huang, C., Ma, W.-Y., Dawson, M.I., Rincon, M., Flavell, R.A., and Dong, Z. (1997). Blocking activator protein-1 activity, but not activating retinoic acid response element, is required for the antitumor promotion effect of retinoic acid. Proceedings of the National Academy of Sciences *94*, 5826-5830.

Jain, R., Barkauskas, C.E., Takeda, N., Bowie, E.J., Aghajanian, H., Wang, Q., Padmanabhan, A., Manderfield, L.J., Gupta, M., and Li, D. (2015). Plasticity of Hopx+ type I alveolar cells to regenerate type II cells in the lung. Nature communications *6*, 6727.

Kamijo, T., Zindy, F., Roussel, M.F., Quelle, D.E., Downing, J.R., Ashmun, R.A., Grosveld, G., and Sherr, C.J. (1997). Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19 ARF. Cell *91*, 649-659.

Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R., Bessy, A., Cheneby, J., Kulkarni, S.R., and Tan, G. (2017). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic acids research *46*, D260-D266.

Koh, P.W., Sinha, R., Barkal, A.A., Morganti, R.M., Chen, A., Weissman, I.L., Ang, L.T., Kundaje, A., and Loh, K.M. (2016). An atlas of transcriptional, chromatin accessibility, and surface marker changes in human mesoderm development. Scientific data *3*, 160109.

Kumar, P.A., Hu, Y., Yamamoto, Y., Hoe, N.B., Wei, T.S., Mu, D., Sun, Y., Joo, L.S., Dagher, R., and Zielonka, E.M. (2011). Distal airway stem cells yield alveoli in vitro and during lung regeneration following H1N1 influenza infection. Cell *147*, 525-538.

Kwon, M.-c., and Berns, A. (2013). Mouse models for lung cancer. Molecular Oncology *7*, 165-177.

Lafkas, D., Shelton, A., Chiu, C., de Leon Boenig, G., Chen, Y., Stawicki, S.S., Siltanen, C., Reichelt, M., Zhou, M., and Wu, X. (2015). Therapeutic antibodies reveal Notch control of transdifferentiation in the adult lung. Nature *528*, 127.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357.

Latil, M., Nassar, D., Beck, B., Boumahdi, S., Wang, L., Brisebarre, A., Dubois, C., Nkusi, E., Lenglez, S., and Checinska, A. (2017). Cell-type-specific chromatin states

differentially prime squamous cell carcinoma tumor-initiating cells for epithelial to mesenchymal transition. Cell stem cell *20*, 191-204. e195.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078-2079.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics (Oxford, England) *30*, 923-930.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology *15*, 550.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nature biotechnology *28*, 495.

Morris, S.A. (2019). The evolving concept of cell identity in the single cell era. Development *146*, dev169748.

Murthy, P.K.L., Srinivasan, T., Bochter, M.S., Xi, R., Varanko, A.K., Tung, K.-L., Semerci, F., Xu, K., Maletic-Savatic, M., and Cole, S.E. (2018). Radical and lunatic fringes modulate notch ligands to support mammalian intestinal homeostasis. eLife *7*, e35710.

Muzumdar, M.D., Dorans, K.J., Chung, K.M., Robbins, R., Tammela, T., Gocheva, V., Li, C.M.-C., and Jacks, T. (2016). Clonal dynamics following p53 loss of heterozygosity in Kras-driven cancers. Nature communications *7*, 12685.

Nabhan, A.N., Brownfield, D.G., Harbury, P.B., Krasnow, M.A., and Desai, T.J. (2018). Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. Science *359*, 1118-1123.

Nikitin, A.Y., Alcaraz, A., Anver, M.R., Bronson, R.T., Cardiff, R.D., Dixon, D., Fraire, A.E., Gabrielson, E.W., Gunning, W.T., and Haines, D.C. (2004). Classification of proliferative pulmonary lesions of the mouse: recommendations of the mouse models of human cancers consortium. Cancer research *64*, 2307-2316.

Pacheco-Pinedo, E.C., and Morrisey, E.E. (2011). Wnt and Kras signaling-dark siblings in lung cancer. Oncotarget *2*, 569.

Pylayeva-Gupta, Y., Grabocka, E., and Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web. Nature Reviews Cancer *11*, 761.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics (Oxford, England) *26*, 841-842.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research *42*, W187-W191.

Rawlins, E.L., and Hogan, B.L. (2006). Epithelial stem cells of the lung: privileged few or opportunities for many? Development *133*, 2455-2465.

Rawlins, E.L., Okubo, T., Xue, Y., Brass, D.M., Auten, R.L., Hasegawa, H., Wang, F., and Hogan, B.L. (2009). The role of Scgb1a1+ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. Cell stem cell *4*, 525-534.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England) *26*, 139-140.

Rohart, F., Gautier, B., Singh, A., and Lê Cao, K.-A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS computational biology *13*, e1005752.

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., and Green, A.R. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389.

Sarkisian, C.J., Keister, B.A., Stairs, D.B., Boxer, R.B., Moody, S.E., and Chodosh, L.A. (2007). Dose-dependent oncogene-induced senescence in vivo and its evasion during mammary tumorigenesis. Nature Cell Biology *9*, 493.

Schep, A.N., Buenrostro, J.D., Denny, S.K., Schwartz, K., Sherlock, G., and Greenleaf, W.J. (2015). Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. Genome research *25*, 1757-1770.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., and Schmid, B. (2012). Fiji: an open-source platform for biological-image analysis. Nature methods *9*, 676.

Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D., and Lowe, S.W. (1997). Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell *88*, 593-602.

Shao, D.D., Xue, W., Krall, E.B., Bhutkar, A., Piccioni, F., Wang, X., Schinzel, A.C., Sood, S., Rosenbluh, J., and Kim, J.W. (2014). KRAS and YAP1 converge to regulate EMT and tumor survival. Cell *158*, 171-184.

Shapiro, E., Biezuner, T., and Linnarsson, S. (2013). Single-cell sequencing-based technologies will revolutionize whole-organism science. Nature Reviews Genetics *14*, 618.

Shaulian, E., and Karin, M. (2001). AP-1 in cell proliferation and survival. Oncogene *20*, 2390.

Sherwood, R.I., Hashimoto, T., O'donnell, C.W., Lewis, S., Barkal, A.A., Van Hoff, J.P., Karun, V., Jaakkola, T., and Gifford, D.K. (2014a). Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. Nature biotechnology *32*, 171.

Sherwood, R.I., Hashimoto, T., O'Donnell, C.W., Lewis, S., Barkal, A.A., van Hoff, J.P., Karun, V., Jaakkola, T., and Gifford, D.K. (2014b). Discovery of directional and

nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. Nature biotechnology *32*, 171-178.

Snippert, H.J., Schepers, A.G., van Es, J.H., Simons, B.D., and Clevers, H. (2014). Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion. EMBO reports *15*, 62-69.

Stark, R., and Brown, G. (2011). DiffBind: differential binding analysis of ChIP-Seq peak data. R package version *100*, 4-3.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences *102*, 15545-15550.

Sutherland, K.D., Song, J.-Y., Kwon, M.C., Proost, N., Zevenhoven, J., and Berns, A. (2014). Multiple cells-of-origin of mutant K-Ras–induced mouse lung adenocarcinoma. Proceedings of the National Academy of Sciences *111*, 4952-4957.

Sweet-Cordero, A., Mukherjee, S., Subramanian, A., You, H., Roix, J.J., Ladd-Acosta, C., Mesirov, J., Golub, T.R., and Jacks, T. (2005). An oncogenic KRAS2 expression signature identified by cross-species gene-expression analysis. Nature genetics *37*, 48.

Tammela, T., Sanchez-Rivera, F.J., Cetinbas, N.M., Wu, K., Joshi, N.S., Helenius, K., Park, Y., Azimi, R., Kerper, N.R., and Wesselhoeft, R.A. (2017). A Wnt-producing

niche drives proliferative potential and progression in lung adenocarcinoma. Nature *545*, 355.

Tata, P.R., and Rajagopal, J. (2017). Plasticity in the lung: making and breaking cell identity. Development *144*, 755-766.

Vallejo, A., Perurena, N., Guruceaga, E., Mazur, P.K., Martinez-Canarias, S., Zandueta, C., Valencia, K., Arricibita, A., Gwinn, D., and Sayles, L.C. (2017). An integrative approach unveils FOSL1 as an oncogene vulnerability in KRAS-driven lung and pancreatic cancer. Nature communications *8*, 14294.

Vierbuchen, T., Ling, E., Cowley, C.J., Couch, C.H., Wang, X., Harmin, D.A., Roberts, C.W., and Greenberg, M.E. (2017). AP-1 Transcription Factors and the BAF Complex Mediate Signal-Dependent Enhancer Selection. Molecular cell *68*, 1067-1082. e1012.

Winslow, M.M., Dayton, T.L., Verhaak, R.G., Kim-Kiselak, C., Snyder, E.L., Feldser, D.M., Hubbard, D.D., DuPage, M.J., Whittaker, C.A., and Hoersch, S. (2011). Suppression of lung adenocarcinoma progression by Nkx2-1. Nature *473*, 101.

Wright, J.R. (1990). Clearance and recycling of pulmonary surfactant. American Journal of Physiology-Lung Cellular and Molecular Physiology *259*, L1-L12.

Wu, J., Huang, B., Chen, H., Yin, Q., Liu, Y., Xiang, Y., Zhang, B., Liu, B., Wang, Q., and Xia, W. (2016). The landscape of accessible chromatin in mammalian preimplantation embryos. Nature *534*, 652.

Wu, M., and Gu, L. (2018). TCseq: Time course sequencing data analysis. R package version *1*.

Xi, R., Murthy, P.K.L., Tung, K.-L., Guy, C.D., Wan, J., Li, F., Wang, Z., Li, X., Varanko, A., and Rakhilin, N. (2019). SENP3-mediated host defense response contains HBV replication and restores protein synthesis. PloS one *14*, e0209179.

Xu, X., Huang, L., Futtner, C., Schwab, B., Rampersad, R.R., Lu, Y., Sporn, T.A., Hogan, B.L., and Onaitis, M.W. (2014). The cell of origin and subtype of K-Ras-induced lung tumors are modified by Notch and Sox2. Genes & development *28*, 1929-1939.

Xu, X., Rock, J.R., Lu, Y., Futtner, C., Schwab, B., Guinney, J., Hogan, B.L., and Onaitis, M.W. (2012). Evidence for type II cells as cells of origin of K-Ras–induced distal lung adenocarcinoma. Proceedings of the National Academy of Sciences, 201112499.

Zacharias, W.J., Frank, D.B., Zepp, J.A., Morley, M.P., Alkhaleel, F.A., Kong, J., Zhou, S., Cantu, E., and Morrisey, E.E. (2018). Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. Nature *555*, 251.

Zhang, X., Choi, P.S., Francis, J.M., Imielinski, M., Watanabe, H., Cherniack, A.D., and Meyerson, M. (2016). Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. Nature genetics *48*, 176.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., and Li, W. (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology *9*, R137.

Zhu, D., Keohavong, P., Finkelstein, S.D., Swalsky, P., Bakker, A., Weissfeld, J., Srivastava, S., and Whiteside, T.L. (1997). K-ras gene mutations in normal colorectal tissues from K-ras mutation-positive colorectal cancer patients. Cancer research *57*, 2485-2492.

CHAPTER 4

CONCLUSION

In this study we have tried to understand the epithelial tissue, which maintains the barrier between the external environment and the body, under homeostatic and oncogenic states. Both the discussions have focused on the importance of stem and progenitor cells which have the ability to proliferate and give rise to fully functional differentiated cell types.

In the first part, we observe the rapidly proliferating intestinal epithelium in homeostasis. The intestinal epithelium consists of crypts and villi with the proliferative *Lgr5*+ stem cells found at the base of the crypts (Barker et al., 2007). Paneth cells which provide a supportive niche to the stem cells are found interspersed between them (Sato et al., 2010). In the upper crypt we find proliferating progenitor cells that differentiate into either secretory cells or absorptive enterocytes (van der Flier and Clevers, 2009). Notch signalling controls both the self-renewal of stem cells at the crypt base and differentiation of progenitors in the upper crypt (van der Flier and Clevers, 2009). Fringe proteins are known to modulate Notch signalling by glycosylating the Notch receptors and thereby affecting their ability to be signaled by DLL or JAG ligands (Haines and Irvine, 2003).

We find that Radical fringe is found in the DLL1 and DLL4 ligand expressing Paneth cells. Loss of RFNG results in reduced ligand expression at the Paneth cell surface resulting in reduced Notch signalling in the *Lgr5*+ stem cells. This affects the ability of the stem cells to self-renew and results in a reduction in their number. This

reveals the importance of tissue organization and niche factors in epithelial homeostasis. In fact, it has recently been shown that overexpression of activated Notch1 in Paneth cells turns them into *Lgr5*+ stem cells (Jones et al., 2019; Yu et al., 2018).

Lunatic Fringe is found in the post-mitotic secretory cells – goblet, enteroendocrine and tuft – scattered along the villus and in the upper crypt. Interestingly, the fourth major secretory cell type, Paneth cell, does not express LFNG. Goblet cells are known to express the ligands DLL1 and DLL4. We find that upon loss of LFNG, the ligand expression on the cells surface of goblet cells reduces. This affects Notch signalling in the enterocyte progenitors found in the upper crypt. However, the role of LFNG, or that of DLL1/4, in the goblet cells of the villi is unknown. It would be interesting to look at the potential Notch based interaction between goblet cells and the mesenchyme. We do not find any phenotype upon deleting Manic fringe in the intestinal epithelial cells. This discovery provides us with the potential to target Notch signalling separately in the upper crypt or at the crypt bottom by targeting Lunatic or Radical fringe respectively.

The mechanism by which the Fringe proteins affect the trafficking of Notch ligands is unknown and would be a potential subject for a future study. Although the three fringes are known to modulate the EGF like repeats on the Notch receptor or ligands, the exact sites at which they glycosylate each protein needs to be characterized. The effect of multiple fringes in the same cell also needs to be understood. It would be interesting to see if one fringe protein can compensate the loss of another fringe. The interaction of Fringes with the Notch ligands and receptors in

the presence of other glycosyltransferases like EOGT needs to be studied. Although, Notch pathway has been shown to be important in a host of developmental, regenerative, homeostatic and pathologic scenarios, the effect of Fringe in those contexts is yet to be unraveled.

In the second part of this dissertation, we have focused on the deviation of epithelia from their homeostatic state into a neoplastic one. Here, we have studied the epigenetic basis of mutant-*Kras* driven neoplastic transformation of pulmonary epithelial cells. The alveoli of the lung are composed of two major cell types: alveolar type I (AEC1) and type II (AEC2) cells. AEC2 are largely considered to act as stem cells and give rise to AEC1 (Barkauskas et al., 2013). The distal bronchiole which opens into alveoli shows the presence of Club and Ciliated cells (Tata and Rajagopal, 2017). Pulmonary epithelial cells are known for their display of plasticity after injury: AEC1 can give rise to AEC2 cells, and Club cells can make basal cells of the airway (Jain et al., 2015; Tata and Rajagopal, 2017).

RAS signalling has been shown to be important in the renewal of AEC2 cells (Desai et al., 2014; Logan and Desai, 2015). Expression of mutant *Kras* in AEC2 or Club cells can lead to the formation of adenocarcinoma (Sutherland et al., 2014). We observed that the chromatin architecture changes significantly during this process. New regions on the chromatin that become accessible in both Club and AEC2 driven tumours are enriched for AP-1 transcription factor motif occurrences. We also observed an increase in the AP-1 TF footprint on the chromatin in neoplastic cells indicating an increased AP-1 activity. AP-1 dimers are composed of FOS (FOS, FOSB, FOSL1 and FOSL2) and JUN (JUN, JUNB and JUND) family proteins. We

find that among all the AP-1 genes, only the expression of *Fosl1* is high in AEC2 derived neoplasm suggesting that the AP-1 complex of interest in these tumours is made up of FOSL1 containing dimers. By immunoprecipitation experiments, we identify that FOSL1 binds to all the JUN proteins, JUN, JUNB and JUND, in KRAS mutant cells. We also find that AP-1 directly binds to nucleosome remodeling complex SWI/SNF. We observe that *Kras* mutant AEC2 cells fail to grow when AP-1 activity is blocked using a small molecule inhibitor SR 11302.

Loss of function SWI/SNF mutations have been found in many cancers including those in the lung, and have often been shown to promote tumorigenesis (Lissanu Deribe et al., 2018). But we have observed that SWI/SNF complex is essential to remodel the chromatin to initiate neoplastic transformation. The order in which cancer cells acquire mutations might explain this apparent contradiction. Observing a study on the interplay between oncogenic *Kras* and the BAF complex in pancreatic cancer (Livshits et al., 2018), it is likely that a fully functional BAF complex is necessary for the initiation of tumours but can be dispensable once the neoplastic transformation is complete.

Two of the important aspects of epithelial biology not covered in this dissertation are development and regeneration. Recent studies have found significant parallels in embryonic development and tissue regeneration (Kotton and Morrisey, 2014). It would be important to understand the changes to chromatin accessibility in these contexts. During the development of embryonic lung, several progenitor populations have been observed (Morrisey and Hogan, 2010). However, it is not yet clear how the *cis*-regulatory elements on the genome influence these progenitors. It

would be important to understand the transcription factors that influence the chromatin architecture and thereby control the differentiation of immature progenitors into fully functional lung epithelium.

REFERENCES

Barkauskas, C.E., Cronce, M.J., Rackley, C.R., Bowie, E.J., Keene, D.R., Stripp, B.R., Randell, S.H., Noble, P.W., and Hogan, B. (2013). Type 2 alveolar cells are stem cells in adult lung. Journal of Clinical Investigation *123*, 3025-3036.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., *et al.* (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature *449*, 1003-1007.

Desai, T.J., Brownfield, D.G., and Krasnow, M.A. (2014). Alveolar progenitor and stem cells in lung development, renewal and cancer. Nature *507*, 190.

Haines, N., and Irvine, K.D. (2003). Glycosylation regulates Notch signalling. Nature reviews Molecular cell biology *4*, 786-797.

Jain, R., Barkauskas, C.E., Takeda, N., Bowie, E.J., Aghajanian, H., Wang, Q., Padmanabhan, A., Manderfield, L.J., Gupta, M., and Li, D. (2015). Plasticity of Hopx+ type I alveolar cells to regenerate type II cells in the lung. Nature communications *6*, 6727.

Jones, J.C., Brindley, C.D., Elder, N.H., Myers Jr, M.G., Rajala, M.W., Dekaney, C.M., McNamee, E.N., Frey, M.R., Shroyer, N.F., and Dempsey, P.J. (2019). Cellular plasticity of Defa4Cre-expressing Paneth cells in response to notch activation and intestinal injury. Cellular and molecular gastroenterology and hepatology *7*, 533-554.

Kotton, D.N., and Morrisey, E.E. (2014). Lung regeneration: mechanisms, applications and emerging stem cell populations. Nature medicine *20*, 822.

Lissanu Deribe, Y., Sun, Y., Terranova, C., Khan, F., Martinez-Ledesma, J., Gay, J., Gao, G., Mullinax, R.A., Khor, T., Feng, N.*, et al.* (2018). Mutations in the SWI/SNF complex induce a targetable dependence on oxidative phosphorylation in lung cancer. Nature Medicine *24*, 1047-1057.

Livshits, G., Alonso-Curbelo, D., Morris IV, J.P., Koche, R., Saborowski, M., Wilkinson, J.E., and Lowe, S.W. (2018). Arid1a restrains Kras-dependent changes in acinar cell identity. Elife *7*, e35216.

Logan, C.Y., and Desai, T.J. (2015). Keeping it together: Pulmonary alveoli are maintained by a hierarchy of cellular programs. BioEssays *37*, 1028-1037.

Morrisey, E.E., and Hogan, B.L. (2010). Preparing for the first breath: genetic and cellular mechanisms in lung development. Developmental cell *18*, 8-23.

Sato, T., van Es, J.H., Snippert, H.J., Stange, D.E., Vries, R.G., van den Born, M., Barker, N., Shroyer, N.F., van de Wetering, M., and Clevers, H. (2010). Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. Nature *469*, 415-418.

Sutherland, K.D., Song, J.-Y., Kwon, M.C., Proost, N., Zevenhoven, J., and Berns, A. (2014). Multiple cells-of-origin of mutant K-Ras–induced mouse lung adenocarcinoma. Proceedings of the National Academy of Sciences *111*, 4952-4957.

Tata, P.R., and Rajagopal, J. (2017). Plasticity in the lung: making and breaking cell identity. Development *144*, 755-766.

van der Flier, L.G., and Clevers, H. (2009). Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium. Annual Review of Physiology *71*, 241-260.

Yu, S., Tong, K., Zhao, Y., Balasubramanian, I., Yap, G.S., Ferraris, R.P., Bonder, E.M., Verzi, M.P., and Gao, N. (2018). Paneth cell multipotency induced by notch activation following injury. Cell Stem Cell *23*, 46-59. e45.