The origin and properties of pro-oncogenic fields in the intestinal epithelium

Ann-Sofie Kirkebjerg Thorsen

Clare College

CRUK, Cambridge Institute

School of Clinical Medicine

University of Cambridge

This dissertation is submitted for the degree of *Doctor of Philosophy*

September 2019

I dedicate this to my family and loved ones.

Declaration

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university of institution for any degree, diploma, or other qualification.

In accordance with the School of Clinical Medicine guidelines, this thesis does not exceed 60,000 words, and it contains less than 150 figures.

Signed:

Date:

Ann-Sofie Kirkebjerg Thorsen

Cambridge

Summary

The intestinal epithelium is maintained by intestinal stem cells that replace each other stochastically over time. Stem cells that are excluded from the stem cell pool will differentiate into the absorptive or secretory cell lineages. However, intestinal cell-fate specification is not always restrictive and early secretory progenitors can revert back to the stem cell pool. Human colon cancer develops by acquisition of oncogenic mutations in the colonic epithelium throughout life. As secretory progenitors have unexpectedly been shown to be a significant source of colonic stem cells, they could serve a cell of origin for tumorigenesis. Furthermore, it has been speculated that tumours could arise in fields that encode pro-oncogenic mutations prior to overt tumour development. Mutations in *KRAS* can be found in 10% of healthy human colons in fields of up to 100 crypts that may be the cause of a subset of colorectal tumours. The extent to which such fields are predisposed to colorectal cancer development is unknown.

The aims of this PhD project were to characterise the cell of origin and stem cell behaviour of *Kras*^{G12D} fields and to develop new tools to recapitulate sequential mutations *in vivo*. Cre mediated lineage tracing of Atoh1 expressing early secretory progenitors carrying *Kras*^{G12D} mutations demonstrated that Kras^{G12D} expression does not change cell fate choices in homeostasis but appears to increase Atoh1+ stem cell contribution in the small intestine after Lgr5 depletion. In addition, Atoh1 derived *Kras*^{G12D} stem cells have a competitive bias and Atoh1 derived *Kras*^{G12D} crypts can multiply over time to create fields and polyps in the colonic epithelium. In addition, a mouse model that utilises Flp and Cre to temporally separate intestinal *Kras*^{G12D} recombination from lineage-tracing was developed to study stem cell behaviour in *Kras*^{G12D} fields. Lineage tracing in this model shows that crypts in *Kras*^{G12D} epithelium have a markedly higher monoclonal conversion rate and accumulate an increased mutation load over time compared to WT crypts. Furthermore, the faster monoclonal conversion rate is shown to be dependent on Mek signalling downstream of Kras. Thus, *Kras*^{G12D} fields may fix secondary mutations at an accelerated rate and so represent pro-oncogenic areas. Lastly, in this thesis a *Rosa26* Dre^{Pr} mouse model, that allows for recombination independently of other recombinases, is developed and used

to initiate lineage tracing in a sequential model of Cre activated intestinal tumorigenesis.

Collectively, the data presented in this thesis allows in-dept investigation of the cell of origin and stem cell behaviour in pro-oncogenic *Kras*^{G12D} fields and contributes to the understanding of how such fields might lead to colon cancer.

Acknowledgements

First and foremost, I would like to thank Professor Douglas Winton for the opportunity to be part of his research group and for providing an excellent scientific working environment, in-depth discussions, encouragement and guidance throughout my doctoral studies.

I would further like to extend my sincere gratitude to Dr. Richard Kemp for always being helpful and offering brilliant insights and direction when I needed it. I would also like to thank Dr. Filipe Lourenco and Mathilde Colombé for their outstanding technical assistance throughout this project – it would not have been possible without your help. I am also thankful to our collaborators. Many thanks to Dr. Ed Morrissey for making the mathematical models for this project. Without his insights and skills this thesis would not have been the same. Thank you to Professor Owen Sansom and Dr. Andrew Campbell for providing the AZD6244 used in this project. And a sincere thank you to the CRUK CI core facilities for their scientific support and to the Cancer Research UK for funding this project.

Additionally, I would like to offer my sincere thankfulness to the Winton lab team members, old as new. You always provided a great and cheerful working environment. Thanks for all the lunches, tea-brakes, Christmas dinners, golf outings, Crisis' and pub-evenings. I have truly loved getting to know each and every one of you. I would also like to extend my deepest appreciation to my close friends: Elisabetta, Mathilde, Darcie and Tatjana. You have been my family away from home these last four years. You have kept me sane even when everything seemed difficult. I will forever be grateful for our time together and I will cherish the memories from all our experiences.

Lastly, I would like to say a big heartfelt thank you to my partner Rune. Thank you for always being loving, supportive and for cooking fantastic dinners even when I came home late from the lab - I love you. To my Mom, Dad and sister: you have always encouraged me and been very proud, thank you.

Contents

List of Figures

List of Abbreviations and Acronyms

ACVR2A = Activin A Receptor Type 2A

ADF = Advanced DMEM/F12

Ah = Aryl hydrocarbon receptor

Alpi = Alkaline Phosphatase Intestinal

APC = Adenomatous Polyposis Coli

Ascl2 = Achaete Scoute Like 2

Atoh1 = Atoh homologue transcription factor 1

BGH = Bovine growth hormone

Bp = Base-pair

BrdU = Bromodeoxyuridine, 5-bromo-2'-deoxyuridine

BSA = Bovine serum albumin

cDNA = complimentary DNA

cm = centimeter

CIMP = CpG island methylation phenotype

CIN = Chromosome instability

Colorectal = Colon and rectal

CYP1A1 = Cytochrome P450 family member 1, subfamily a, polypeptide 1

Dclk = Doublecortin like kinase 1

Dkk1 = Dickkopf-1

Dll = Delta like ligand

DNA = Deoxyribonucleic acid

DSS = Dextran sodium sulfate

DT = Diphtheria toxin

Dtr = Diphtheria toxin receptor

EGF = Epidermal growth factor

EGFR = Epidermal growth factor receptor

EpCam = Epithelial Cell Adhesion Molecule

Erk = Extracellular regulated kinase

ESC = Embryonic stem cells

Et al. = et alia

FAP = Familial Adenomatous Polyposis

FBS = Foetal bovine serum

FSF = Frt STOP Frt

Gap = GTPase activating protein

Gef = Guanine exchange factor

GFP = Green fluorescent protein

H2B = Histone-2B

HBSS = Hank's Balanced Salt Solution

Hopx = Homeobox-only protein x

IHC = immunohistochemistry

IkBa = NF-kappa-B inhibitor alpha

KRAS = Kirsten rat sarcoma viral oncogene

NRAS = Neuroblastoma RAS

HRAS = Harvey RAS

kg = Kilograms

LacZ = Lactose operon Z

λ = Stem cell replacement rate / stem cell / day

Lgr5 = Leucine rich G-protein coupled receptor 5

LRP5/6 = Low density lipoprotein receptor 5/6

LSL = LoxP STOP LosP

M = Molar

mM = milli Molar

M-cells = Microfold cells

Mek = Mitogen activated kinase

Meki = Mek inhibitor

mg = milligrams

MLH1 = MutL homolog 1

mPas = Mild reductive periodic acid Schiff

mRNA = Messenger ribonucleic acid

MSH2 = MutS protein homolog 2

MSI = Microsatellite instability

mtDNA = Mitochondria DNA

mTert = Mouse telomerase reverse transcriptase

mTor = Mammalian target of rapamycin

N = Stem cell number per crypt

NBF = Natural buffered formalin

NES = Nuclear export signal

NF-kB = Nuclear Factor kappa-light-chain-enhancer of activated B cells

Ngn3 = Neurogenin 3

NLS = Nuclear localisation signal

PBS = Phosphate-buffered saline

PBS-T = PBS Tween-20

PCR = Polymerase chain reaction

PFA = Paraformaldehyde

PGK = Phosphoglycerate Kinase

PI3K = Phosphoinositide 3-kinase

pMek = Phosphorylated Mek

PPC = Partially populated crypt

PR = Probability of replacement

Pr = Progesterone receptor

R26 = Rosa26

RPM = Rotations per minute

rsr = rox STOP rox

SD = Standard deviation

SEM = Standard error of mean

SI = Small Intestine

SNP = Single nucleotide polymorphism

SNV = Single nucleotide variant

STAG2 = Stromal Antigen 2

TA = Transit amplifying

TCF = T cell factor

tdTom = tdTomato

TE = Tris-EDTA

TGFBR2 = Transforming Growth Factor, Beta Receptor II

TP53 = Tumour protein 53

WPC = Wholy populated crypt

WT = Wild type

Xbp1 = X-Box Binding Protein 1

List of Appendices

1 Introduction

1.1 The intestinal epithelium

1.1.1 Intestinal tissue organisation

The intestinal tract is lined by a single layer of epithelial cells and consists of the small intestine and the large intestine (Figure 1.1A). The small intestine includes the duodenum, jejunum and the ileum. The large intestine includes the caecum, colon and the rectum. The intestinal tract has two main functions: to absorb nutrients and to protect against insults from the environment. The single layer of cells in the intestinal epithelium invaginates into so-called intestinal crypts of Lieberkühn. The small intestine also has have finger like protrusions or villi which contain cells that mediate nutrient uptake (Gehart and Clevers, 2019). Multiple crypts can supply cells to the villus at the base. The large intestine does not contain villi and the main function of the colon is water reabsorption and stool compaction.

The intestinal epithelium is protected against insults by a thick mucosal barrier, immune-epithelial interactions and by a high cell turn-over. To this end, the entire intestinal epithelium is renewed every 3-4 days in mice and rats (Leblond and Stevens, 1948; Darwich *et al.*, 2014). Tissue renewal is maintained by highly proliferative crypt base columnar stem cells (Cheng and Leblond, 1974; Escobar *et al.*, 2011) (Figure 1.1C). Daughter cells migrate up the crypt to the transit amplifying (TA) zone where they undergo 4-5 rounds of proliferation before they leave to terminally differentiate (Figure 1.1C). As differentiated cells reach the villus top they undergo apoptosis which provides a balancing mechanism for the high levels of proliferation in the crypt base (Cheng and Leblond, 1974; Hall *et al.*, 1994) (Figure 1.1C).

1.1.2 Crypt propagation

During development, as the digestive tube elongates, crypts undergo replication and division to line the entire tube; a process called fission (Clarke, 1972). Fissioning transpires through bifurcation of a single crypt into two, and this starts at the crypt-base and moves up in a zipper-like fashion (Gehart and Clevers, 2019). Fission occurs at slow rates in the mature intestine meaning that crypt multiplication is a dynamic process even in adult intestinal tissue (Snippert *et al.*, 2014). Recently, a study showed that an opposing mechanism called crypt-fusion, the process of two crypts fusing back to one, is also occurring in adult murine intestines (Bruens *et al.*, 2017). Fission and fusion acting in balance could maintain the intestine at a certain length through adult life.

Mature intestinal cell types

The intestinal epithelium is comprised by six different mature cell types that can be separated into two major lineages: absorptive cells and secretory cells. The absorptive lineage consists of enterocytes and microfold (M) cells. The secretory lineage consists of Paneth, goblet, enteroendocrine and tuft cells (Gehart and Clevers, 2019).

1.1.3 Enterocytes and M-cells

Enterocytes are the most numerous cell type in the epithelium and their primary function is to absorb nutrients at their brush-border apical surface and export them basally (Cheng and Leblond, 1974; Noah, *et al.*, 2011). The other absorptive cell type, the M-cell, overlay lymphoid follicles and is specialized to sample and transfer luminal contents to underlying immune cells (Lorenz and Newberry, 2004).

1.1.4 Goblet cells

The goblet cell is the most abundant secretory cell type and accounts for approximately 60% of all secretory cells in the small intestine (Cheng and Leblond, 1974b). The relative proportion of goblet cells increase from the proximal to the distal small intestine with the highest levels found in the colon (Kemper and Specian, 1991). Goblet cells produce and secrete mucus and thus are characterised by their globules of mucin that stain positive for Alcian Blue and negative for Pas-haematoxylin (Spicer, 1960).

1.1.5 Paneth cells

Unlike any other differentiated cells in the crypt-villus axis, Paneth cells, found only in small intestine, migrate downward and reside interspaced between crypt-base stem cells upon differentiation (Cheng, 1974c; Wehkamp and Stange, 2010). Paneth cells are characterised by their granules with high lysozyme activity and their main function is thought to be secretion of antibiotic defensins in response to bacterial stimuli (Deckx, *et al.*, 1967; Wehkamp and Stange, 2010). However, it has been shown that Paneth cells could be a niche providing cell since they express factors required for stem cell maintenance (Sato *et al.*, 2011).

1.1.6 Enteroendocrine cells

Enteroendocrine cells makes up 0.5% of the intestinal epithelium (Cheng and Leblond, 1974a) and their main function is to secrete hormones into the bloodstream to regulate responses such as appetite sensation and insulin production. Thus, enteroendocrine cells can stimulate production of insulin by releasing hormones such as glucagon-like peptide 1 and gastric inhibitory polypeptide upon food ingestion (Gribble and Reimann, 2019).

1.1.7 Tuft cells

Tuft cells are rare and only constitute 0.4% of the small intestinal epithelium during homeostasis and they are characterised by their apical located microvilli (Banerjee *et al.*, 2018). Their function remains unknown, however, recently it was reported that tuft cells are important players during intestinal parasitic worm infections. In response to parasitic infections tuft cells secrete Interleukin-25 which leads to tuft cell expansion and mounting of a type 2 immune reaction to clear the infection (Gerbe *et al.*, 2016; Howitt *et al.*, 2016; von Moltke *et al.*, 2016).

1.2 Lineage specification

Stem cell maintenance and lineage specification is regulated by opposing signalling cascades and transcription factor programmes along the crypt-villus axis. In the crypt base active Wnt and Notch signalling maintains stem cells in an undifferentiated state. As daughter cells move out of the stem cell niche, changes in Wnt and Notch signalling orchestrate differentiation of cells into either the absorptive or secretory cell lineages.

1.2.1 Maintenance of stemness

Intestinal stem cells expresses the Leucine rich G-protein coupled receptor 5 (Lgr5) on their cell surface (Barker *et al.*, 2007) and are thought to be kept in their undifferentiated state by the transcription factor Achaete Scute-Like 2 (Ascl2) (van der Flier *et al.*, 2009). The two main signalling cascades that maintain this phenotype are canonical Wnt signalling and active Notch signalling.

During canonical Wnt signalling the Wnt ligands bind to Frizzled-LRP5-LRP6 receptor complex which in turn inhibits the destruction of ß-catenin. Stabilised ß-catenin will translocate to the nucleus where it regulates gene expression by binding with T cell factors (TCFs) (Gehart and Clevers, 2019). The importance of Wnt-signalling on stem cell maintenance was underlined in four studies where Wnt-signalling was knocked out or inhibited in the murine intestine. Knock out of

the ß-catenin interacting transcription factor TCF-4, either conditionally or in the germline, showed that TCF-4 is essential for maintaining crypt structures and proliferation in adult and neonatal mice, and the lack thereof lead to premature death (Korinek *et al.*, 1998; van Es *et al.*, 2012a). In addition, overexpression of the Wnt antagonist Dickkopf-1 (Dkk1), also lead to decreased proliferation and loss of crypt and villus structures, however mice survived long-term. This was probably due to mosaic Dkk1 expression from constructs and transient Dkk1 expression vectors (Pinto *et al.*, 2003; Kuhnert *et al.*, 2004).

Stem cells are also dependent on active Notch signalling for maintenance of their undifferentiated state (VanDussen *et al.*, 2012). Active Notch signalling requires direct interaction between two adjacent cells, one expressing the Notch ligand in the form of a family member of the Delta-like ligand (Dll) (non-stem cell) and one expressing the Notch receptor (stem cell). Upon interaction, γ-secretase cleaves the intracellular Notch receptor domain which translocate to the nucleus and activates expression of the basic loop helix transcription factor Hes1. Hes1 in turn inhibits expression of the secretory lineage commitment factor Atoh homologue transcription factor 1 (Atoh1), also known as Math1 (Philpott and Winton, 2014; Gehart and Clevers, 2019). The importance of active Notch signalling in stem cells was emphasised in a report showing that intestinal knockout of the Notch ligand Delta-like1 (Dll1) resulted in permanent stem cell differentiation and stem cell loss in the murine intestine (Pellegrinet *et al.*, 2011).

Wnt and Notch ligands are thought to be supplied by niche providing crypt-base cells. Paneth cells have been shown to produce Wnt-ligands *in vivo* (Sato *et al.*, 2011; Farin *et al.*, 2016) thus classifying these as niche cells. However, the finding that Paneth cell loss does not affect Lgr5+ stem cell function, suggest that Paneth cells are not exclusive in serving this function (Kim *et al.* 2012). Recently, sub-epithelial telocytes have been discovered to be an additional niche providing cell-type by secreting Wnt ligands at the bottom of the crypt and Wnt-antagonists at the top (Shoshkes-Carmel *et al.*, 2018).

1.2.2 Secretory and absorptive lineage specification

As daughter cells move out of the crypt-base niche, they lose active Wnt signalling. Wnt signalling is thought to decrease up the crypt-villus axis because of the poor solubility of Wnt ligands and secretion of Wnt inhibitory molecules by mesenchymal cells (Farin *et al.*, 2016; Gehart and Clevers, 2019). This triggers the first binary cell-fate decision: to become a secretory or an absorptive cell.

This decision is dependent on Notch signalling. If cells lack cell-cell contact with Dll expressing neighbour cells as they move up the crypt, Notch signalling, and thus Hes1 expression, will be lost which leads to Atoh1 expression (Philpott and Winton, 2014). In turn, Atoh1 expression will start the secretory lineage commitment programme (Figure 1.2). Yang *et al*. and Shroyer *et al.* showed that germline or intestinal specific *Atoh1* ablation lead to loss of secretory goblet, Paneth and enteroendocrine cells (Yang *et al.*, 2001; Shroyer *et al.*, 2007). Likewise, VanDussen *et al.* showed that Atoh1 overexpression lead to increased levels of secretory cells and loss of absorptive cells (VanDussen and Samuelson, 2010). These findings confirm that secretory cells arise from a common progenitor which is dependent on Atoh1 expression.

Atoh1+ secretory progenitors are poised to become Paneth cells if they migrate down the crypt and encounter active Wnt signalling (van Es, *et al.*, 2005b), or goblet cells if they do not (Figure 1.2). Differentiation of enteroendocrine cells requires subsequent expression of Neurogenin-3 (Ngn3) and NeuroD1 (Jenny *et al.*, 2002; Li *et al.*, 2012) (Figure 1.2).

Thus, Atoh1 expression constitutes the initial “on” switch for the secretory lineage and Atoh1 expressing cells include the earliest secretory progenitors.

On the other hand, daughter cells that leave the stem cell niche but do maintain active Notch signalling become absorptive progenitor TA cells due to the lack of Atoh1 expression. Thus, upon exit from the TA zone they become enterocytes (Gehart and Clevers, 2019) (Figure 1.2). Notch-signalling retardation through γ-secretase inhibition in mice converts proliferative TA-cells into goblets cells,

underlining the crucial effect of Notch-signalling on TA-cell commitment to the absorptive lineage (Milano *et al.*, 2004; van Es *et al.*, 2005a).

Thus, early lineage choices are dependent on Notch-signalling and Atoh1 expression. The relative proportion of secretory versus absorptive cell number relies on Notch-ligand availability and strongly favours the absorptive lineage.

1.3 Tracing cells *in vivo*

Lineage tracing is the process of identifying all offspring of a single cell (Kretzschmar and Watt, 2012). *In vivo* this is mostly achieved through genetic recombination. Here, DNA recombinases are often expressed in a cell-type specific manner. Activation of recombinases in cell(s) of interest activates a conditional reporter gene to permanently label the cell and its descendants. Reporter genes often confer colour labelling to cells for visualisation. The first reporter developed was a Cre inducible LacZ allele (Soriano, 1999). However, since ß-galactosidase activity is not directly visible, most reporters today employ inducible fluorescent proteins (Kretzschmar and Watt, 2012).

Lineage tracing in the intestine is thus carried out by marking a single cell with a colour and following the multiplication of daughter cells or clones over time (Figure 1.3). Therefore, lineage tracing can be used to test stemness *in vivo*. If lineage traced cells have stem cell abilities, they can give rise to daughter cells of multiple lineages that move up the crypt and villus (in the small intestine) to give rise to long-lived clones (ribbons) (Figure 1.3A). In addition, lineage-traced clones can be visualised in the crypt base to study stem cell replacements and clonal growth in pulse-chase experiments conducted over several timepoints (Figure 1.3B-C) (discussed later).

1.4 Plasticity

The intestinal epithelium is a highly regenerative organ. This is displayed by the complete organ renewal every 3-4 days and the impressive capacity of regeneration following tissue damage. Intestinal regeneration is fuelled by rapid cycling stem cells. Remarkedly, when mice are exposed to full body irradiation that results in haematopoietic failure and stem cell depletion, the intestine can still regenerate (Beumer and Clevers, 2016). In addition, if Lgr5+ stem cells are depleted the gut appears to be unaffected (Tian *et al.*, 2011). Together, these results demonstrate the presence of plastic cells that can reacquire stem cell function during tissue regeneration.

1.4.1 Plasticity of progenitors

There are different theories that explains cell plasticity in the intestine, the two main models describe the activation of “sleeping” quiescent stem cells or dedifferentiation of committed progenitors to re-enter the stem cell pool.

Several early studies identified what was believed to be quiescent stem cells in intestinal crypts. Accordingly, by utilizing a *Bmi1*^{CreErt} mouse two groups described the existence of Bmi1+ non-proliferative cells at position +3-4 in the crypt (Sangiorgi and Capecchi, 2008; Yan *et al.*, 2012) (Figure 1.1). Bmi1+ cells gave rise to stem cell derived clones and upon irradiation the number of Bmi1+ cells increased whereas Lgr5+ cells were depleted. Similar results were obtained when using the mouse telomerase reverse transcriptase (mTert)-GFP and the Homeobox-only protein x (*Hopx*)^{CreErt} mouse. Here the authors show that mTert and Hopx marked slowly proliferating cells that gave rise to stem cell clones (Montgomery *et al.*, 2011; Takeda *et al.*, 2011). In addition, mTert+ cells were resistant to irradiation and had a high regenerative potential (Montgomery *et al.*, 2011). Together these publications described the presence of slowly cycling cells that can contribute to the stem cell pool upon reactivation.

The existence of “sleeping” stem cells that are spared proliferative damage have been speculated to be an evolutionary strategy to maintain the

integrity of the intestinal stem cell genome. However, this concept implies that such cells are destined to only serve a purpose if re-activated. This idea has been heavily debated in the community, and therefore, a study was designed to define the exact nature of quiescent stem cells. Instead of defining such cells based on gene-expression Buczacki and colleagues identified quiescent stem cells based on their slowly-cycling properties (Buczacki *et al.*, 2013). In an elegant mouse model, they separated the Cre recombinase into two polypeptides (CreA and CreB). CreA was expressed constitutively in all cells whilst the CreB was inducibly expressed from as a Histone-2B (H2B) fusion construct. Following a pulse of H2B-CreB expression, due to cell division, CreB would be diluted out in proliferating cells, and therefore, the two complimentary halves of Cre would only be present in slowly cycling cells. A dimerizing agent was utilized to activate Cre which resulted in linage tracing of slowly cycling cells. This study showed that slowly cycling cells are secretory progenitor cells that can re-gain stem cell potential during regeneration (Buczacki *et al.*, 2013). Thus, the study exemplified that the quiescent stem cell pool consists of committed progenitors that can serve as stem cells or as functional differentiated cells.

The discovery that progenitors can contribute to the stem cell pool provides a large alternative source of stem cell potential. Several studies have confirmed that stem-cell reactivation of progenitors is occurring *in vivo*. This was emphasised in a study that lineage-traced secretory enteroendocrine cells expressing Ngn3 and showed that such cells produced rare stem cell clones during homeostasis (Schonhoff *et al.*, 2004). An additional study showed that Dll1+ secretory progenitors only contributed to the stem cell pool following damage (van Es *et al.*, 2012b). Furthermore, doublecortin like kinase 1+ (Dclk+) tuft cells have also been shown to produce rare stem cell clones during homeostasis (Westphalen *et al.*, 2014).

Cells of other lineages are also plastic. This was demonstrated in a study that lineage traced the most abundant intestinal cell type: Alkaline Phosphatase Intestinal+ (Alpi+) absorptive enterocytes and progenitors thereof. The results showed that Alpi+ cells only rarely contributed to the stem cell pool during tissue homeostasis (Figure 1.2). In contrast, the contribution of this cell

type increased when Lgr5 cells were depleted, suggesting a role for them during tissue regeneration (Tetteh *et al.*, 2016). Interestingly, Alpi+ stem cell contribution decreased the longer the time interval between Alpi tracing and Lgr5 depletion. This suggested that the further up the crypt Alpi+ cells migrated the smaller their dedifferentiation potential (Tetteh *et al.*, 2016).

Together, these results indicate that committed progenitors contribute to the stem cell pool both during homeostasis, and to a larger extent, during damage. Furthermore, it seems that the degree to which progenitors regain stem cell properties is dictated by where in the lineage hierarchy they reside. The closer to their stem cell origin the greater the stem cell potential.

This is exemplified by studying the earliest secretory progenitors marked by Atoh1 expression. Ishibashi and colleagues related data from Atoh1+, Dll1+ and Dclk1+ lineage tracing to compare their relative stem cell contribution during homeostasis. By doing so the authors showed that Atoh1+ secretory progenitors have a 10-fold higher stem cell contribution potential in homeostasis compared to the further differentiated Dll1+ cells and Dclk1+ cells (Ishibashi *et al.*, 2018). Later, Tomic *et al.* constructed a mouse with impaired Atoh1 phosphorylation potential. Through lineage tracing experiments the authors compared the stem cell potential of Atoh1 and phospho-deficient Atoh1 cells and it was revealed that stem cell contribution is dependent on Atoh1 phosphorylation (Tomic *et al.*, 2018). Furthermore, Atoh1 derived stem cells have an enhanced regenerative potential upon colitis induced injury compared to other secretory progenitors (Castillo-Azofeifa *et al.*, 2019).

Conclusively, these publications demonstrate that Atoh1+ progenitors have the highest homeostatic stem cell contribution potential, compared to other lineage restricted progenitors. In addition, these studies underline that Atoh1+ progenitors contribute significantly to the stem cell pool in homeostasis and during injury (Figure 1.2). This greatly enhances the size of the stem cell pool.

1.4.2 Regulation of plasticity

Thus, it is now clear that intestinal cell specification is not always restrictive, and that cell fate is defined in a background of context dependent continuum. The factors that regulate plasticity are still under investigation. It is clear that tissue injury increases the number of cells that fall back into the stem cell pool. Several studies have indicated that Wnt and Epidermal growth factor receptor (EGFR) signalling are crucial for regeneration following tissue injury and intestinal DNA-damage (Valentin-Vega *et al.*, 2008; Ashton *et al.*, 2010). In addition, Tomic *et al*. showed that post-translational modifications can regulate the availability of cell fate determinants such as Atoh1 (Tomic *et al.*, 2018). In all likelihood, several plasticity regulatory pathways will be uncovered in the future.

1.5 Stem cell dynamics

Plasticity provides a mechanism to protect the integrity of the intestinal epithelium. Another layer of protection is applied during stem cell renewal and competition in the crypt base. As the size of the intestinal stem cell pool remains constant over time, it was believed for many years that stem cells undergo asymmetrical division to produce one cell destined for differentiation and one stem cell. This theory was mainly fuelled by publications showing that intestinal stem cells segregate their DNA asymmetrically during division (Potten *et al.* 2002). However, early studies that utilized a random mutagen or chromosome X-inactivation to study loss of a protein-stain in the murine intestine showed that crypts become monoclonal over time, thus disputing the idea that single stem cells divide asymmetrically (Griffiths *et al.*, 1988; Winton *et al.*, 1988).

1.5.1 Neutral stem cell drift

Two studies set out to decipher the fate of single stem cells to determine the exact nature of stem cell self-renewal in the murine intestine. One study followed clonal growth over time by lineage tracing less than one stem cell per crypt. The results showed that the total number of labelled clones declined over time but the relative size of surviving clones increased, resulting in 80% of crypts being monoclonal

(clonal fixation) after 16 weeks (Lopez-Garcia *et al.*, 2010). The other study labelled multiple Lgr5-GFP+ cells in crypts by utilizing a reporter allele that randomly allocates one of 4 colours per cell (Confetti-allele). To this end, each Lgr5-GFP+ stem cell had a distinct colour at initial labelling which made it possible to follow clonal growth over time (Snippert *et al.*, 2010). The results showed the same pattern of reduced clone numbers but increased clonal size over time as described by Lopez-Garcia and colleagues (Lopez-Garcia *et al.*, 2010). The overall conclusion from these ground-breaking studies was that *"stem cells follow a pattern of neutral drift in which the loss of a stem cell is compensated by the multiplication of a neighbour"* (Lopez-Garcia *et al.*, 2010).

Consequently, during stem cell clonal growth, single cells will divide and replace each other in a stochastic manner, thus, clones can grow and regress, but with time one clone will win and convert the crypt to monoclonality (Figure 1.3B-C). To explain this behaviour in quantitative terms, both papers assumed the number of stem cells per crypt to be approximately 15-20 based on the number Lgr5+ cells defined by Barker *et al.* 2007 and used mathematical modelling to infer that each stem cell is replaced 0.75-1 times per day (Lopez-Garcia *et al.*, 2010; Snippert *et al.*, 2010). These findings indicated that the majority of intestinal stem cell divisions are actually symmetrical and that the decision to exit the stem cell zone and become a TA-cell and committed progenitor is based on competition and extinction from the crypt base.

However, since the two above papers were published other reports used single molecule *in situ* hybridisation to document that the Lgr5+ pool also expresses markers for committed progenitors (such as mTert, Hopx and Bmi1) that are located further up the crypt at location +4 and thus outside the stem cell zone (Figure 1.1) (Itzkovitz *et al.*, 2012; Muñoz *et al.*, 2012). This implied that evaluating stem cell number based on Lgr5-expression may not always be accurate since this population is overlapping with other non-stem cells. Since the publications reviewing neutral drift had based their stem cell numbers (and thus the inferred stem cell replacement rate) on Lgr5-expression, estimates could be inaccurate.

In an elegant report, Kozar and colleagues investigated stem cell number and stem cell replacement patterns by a Lgr5 free method (Kozar, *et al.*, 2013). Here, they utilized a mouse model that encoded a fluorescent reporter that was prevented from being expressed by an unstable CA[30] microsatellite. The CA[30] microsatellite placed the reporter complimentary DNA (cDNA) out of frame and during replication, frameshift mutations in the CA[30] repeat region lead to stable marker expression and clonal labelling. Thus, with mouse ageing an increasing number of crypts were marked, so-called continuous labelling. The results identified a constant level of partially populated crypts (PPCs) over all mouse ages (Kozar, *et al.*, 2013). This fitted the assumption that new marked clones accumulate over time and that some of these will reach monoclonality to become wholly populated crypts (WPCs). As WPCs are derived from PPCs, the size of the PPC population was at a steady size over time. Furthermore, an expected linear increase in WPCs crypts was also observed (Kozar, *et al.*, 2013). Computer simulations estimated that rate of WPC accumulation depended on the mutation rate of CA[30] and the stem cell replacement rate and that the frequency of PPC depended on the CA[30] mutation rate and the number of stem cells per crypt. Thus, by estimating the CA[30] mutation rate and using mathematical modelling Kozar *et al.* concluded that crypts in proximal small intestinal and colon are maintained by 5-7 cells, that are replaced 0.1 – 0.3 times per day, respectively (Kozar, *et al.*, 2013). These numbers were in striking contrast to the numbers reported by Barker *et al.* 2007, Lopez-Garcia *et al*. 2010 and Snippert *et al*. 2010 and showed that only a small pool of “functional” stem cells, or inferred clonogenic fraction, is driving neutral drift in each crypt.

1.5.2 Human stem cell dynamics

Since these studies much effort has been made to understand how stem cells behave in the human colon. For obvious reasons it is not possible to use genetically modified reporter genes in humans and thus the community has searched for other naturally occurring labels that can mark stem cell clones. Taylor and colleagues found that mitochondria DNA (mtDNA) in human colonic stem cells is subject to common spontaneous mutations that cause cytochrome

c deficiency resulting in a negative staining pattern in the crypts affected (Taylor *et al.*, 2003). The group reported that this staining pattern reflected age-accumulated colonic clones (Taylor *et al.*, 2003). These clones represented the first collected continuous-labelling data-set from humans and created an exciting way to lineage trace humane intestinal stem cells. Later, the same research group showed that these mtDNA clones become fixed in the crypt over time and that such mutated crypts spread via fission to create crypt-patches (Greaves *et al.*, 2006). Baker *et al.* concluded that colonic cytochrome c deficient clones follow a pattern of neutral drift and inferred the number of human colonic stem cells to be 5 per crypt (Baker *et al.*, 2014).

Later, Nicholson *et al.* discovered another human clonal mark: detection of loss of O-acetylation of sialomucins using a mild reductive periodic acid Schiff (mPAS) stain. This mark is thought to arise from somatic mutations in the human genome rather than those arising from mtDNA mutations. By screening for mPAS+ clones in healthy human colon tissue from patients of different ages the authors showed that the number of PPCs remained constant with age, but the number of WPCs increased in the same fashion as was described in the murine intestine by Kozar *et al*., 2013. By determining the mutation rate of mPAS staining, and using the relationship between PPCs and WPCs the authors used Bayesian inference to define that stem cell renewal arises from the activity of around 7 stem cells per human colon crypt which replace each other approximately 1.3 times/year/crypt (Nicholson *et al.*, 2018).

Thus, it appears that the number of stem cells per crypt is very similar in humans and mice, however, the rate by which clones drift toward fixation is approximately 100-fold faster in mice compared to humans. This discrepancy could stem from a much slower stem cell proliferation rate in humans or could be because human stem cells divide mostly asymmetrical, creating very few cells for clonal drift (Stamp *et al.*, 2018). However, it seems more likely that the timing of stem cell dynamics in human colonic crypts reflects a resolved behaviour that has not yet been fully explained in the literature.

1.6 Intestinal cancer

Colon and rectal (colorectal) cancer is the third most common cancer worldwide and accounts for approximately 10% of cancer related mortality (Kuipers *et al.*, 2015). There are two major subgroups of colorectal cancer; the hereditary and the non-hereditary forms. Lynch Syndrome, a form of hereditary colorectal cancer, results from germline mutations in DNA mismatch repair genes such as *MutL Homolog 1* (*MLH1), MutS protein Homolog 2 (MSH2), PMS2* and *Epithelial Cell Adhesion Molecule* (*EPCAM).* Impaired DNA mismatch repair during replication causes accumulation of DNA mutations in microsatellite regions (microsatellite instability, MSI) and increased risk of developing cancer at an early age. Familial adenomatous polyposis (FAP), another form of hereditary colorectal cancer, is caused by germline mutations in the *Adenomatous Polyposis Coli* (*APC*) gene that controls the activity of the Wnt signalling pathway (Kuipers *et al.*, 2015). Most patients with this disease develop multiple colorectal adenomas and colorectal cancer at an early age. In the following, the non-hereditary sporadic form of colorectal cancer will be discussed.

1.6.1 Non hereditary colorectal cancer

The common risk factors for sporadic colorectal cancer development are lifestyle and inflammatory diseases such as inflammatory bowel disease, ulcerative colitis or Crohn's Ileocolitis. Sporadic colorectal cancer is initiated with aberrant crypts which can progress into polyps that can develop into early and late adenomas which, with time, can become carcinomas. This process takes 10-15 years and is driven by sequentially acquired somatic mutations and epigenetic alterations that activate oncogenes or silences tumour suppressor genes (Kuipers *et al.*, 2015). In 1990 Fearon and Vogelstein stated that the majority of adenomas that develop into carcinomas follow a sequence of events initiated by mutations and loss of function of *APC*. Next, step-wise acquisition of hypomethylated tumour suppressor genes, oncogenic mutations in *Kirsten Rat Sarcoma* (*KRAS)*, loss of chromosome 18q and *Tumour Protein 53* (*TP53)* mutations drives the process (Fearon and Vogelstein, 1990).

However, since the publication of this paradigm setting paper, it has become evident that sporadic colorectal cancer is a heterogenous disease. For improved clinical intervention it is important to be able allocate clinical sub-types to the disease, thus, many efforts have been made to define such. Most studies divide sporadic colorectal cancer into 3 main subgroups: 1) a subgroup that displays a high degree of chromosome instability (CIN) 2) a subgroup that displays an aberrant CpG island methylation phenotype (CIMP) which results in silencing of tumour suppressor genes and 3) a subgroup that is initiated by somatic mutations in mismatch repair genes causing MSI and hypermutated tumours (Fearon, 2011; Nojadeh *et al.*, 2018). Additionally, a newly recognised subgroup which evolves through sessile serrated polyps and comprises 5-10% of cases has been identified (De Sousa E Melo *et al.*, 2013). The initiating event for such are thought to be oncogenic mutations in *BRAF* and *KRAS* (Kuipers *et al.*, 2015). Furthermore, next generation sequencing of hundreds of human colorectal tumours allocated two major subtypes to sporadic colorectal cancer: the hypermutated tumours wherein the highest mutated genes are *Activin A Receptor Type 2A* (*ACVR2A)* (63%), *APC* (51%), and *Transforming Growth Factor, Beta Receptor II (TGFBR2)* (51%) and non-hypermutated tumours wherein the highest mutated genes are *APC* (80%), *TP53* (60%) *and KRAS* (43%) (The Cancer Genome Atlas Network, 2012). Additionally, several research groups have tried to classify the disease based on clinical outcome (Sadanandam *et al.*, 2013; Phipps *et al.*, 2015).

Recently, a study decided to assign molecular subtypes to classify the disease based on transcriptional behaviour instead of mutation profile. This report combined gene expression data from 18 different studies to define four consensus molecular subtypes: the microsatellite instability immune subtype (CMS1), the hypermutated microsatellite unstable and strong immune activation (CMS2), the epithelial, marked Wnt and Myc signalling activation (CMS3) and epithelial and evident metabolic dysregulation subtype (CMS4) (Guinney *et al.*, 2015). Some subtypes showed a clear enrichment for certain mutations such as *KRAS* mutations in CMS3, however none of them were defined by an individual event and no genetic mutation was limited to a subtype (Guinney *et al.*, 2015).

However, mutations in *APC* were common to all subtypes, which supports the crucial role for this gene in tumorigenesis. Several mouse studies have confirmed this by showing that *Apc* loss can initiate multiple intestinal tumours *in vivo* (Su *et al.*, 1992; Cheung *et al.*, 2009; Robanus-Maandag *et al.*, 2010).

Therefore, it appears that colorectal cancer development cannot be summed up by any specific sequence of events and it seems much is still to be discovered about the initiation of the disease.

1.7 Tumour initiation

1.7.1 Cell of origin

Since the intestine displays a high degree of plasticity it has been heavily debated what cell type constitutes the cell of origin in intestinal tumorigenesis. As the intestinal epithelium is renewed every 3-4 days it is clear that accumulation of cancerous mutations must be present in stem cells, or change the migration abilities of the cell of origin, to be maintained in the tissue over time.

Therefore, it has generally been believed that transformative mutations originate in stem cells. This was disputed by a study showing that early neoplastic lesions contain dysplastic cells and *APC* mutations in the upper crypt only (a non-stem cell zone), with the base of the crypt looking histologically normal (Shih *et al.*, 2001). The authors debated a “Top-Down” origin of intestinal tumours, where cells in intercryptal zones encode the transformative mutational event. These cells could divide and migrate down to occupy the whole crypt and thus could be the cell of origin. This was later refuted in a report studying colonic tissue from patients with FAP to catch the earliest lesions possible (Preston *et al.*, 2003). This study showed that the earliest adenomas consist of discrete monoclonal crypts that stop sharply at the junction of the surface epithelial, and thus do not involve the intercryptal zones (Preston *et al.*, 2003). In addition, adenoma development

was shown to be a result of crypt fission, supporting a “Bottom-Up” origin of colorectal tumours. In this model, the transformative event takes place in stem cells which create clones that expand stochastically making the crypt monoclonal over time. The monoclonal crypt can then spread by fission (so-called lateral spread) to progress tumour development (Preston *et al.*, 2003).

To confirm that stem cells can initiate tumours Barker and colleagues showed that *Apc* deletion in Lgr5+ cells resulted in a high tumour penetrance with rapidly growing adenomas being macroscopic after 3-5 weeks (Barker *et al.*, 2009). This confirmed that intestinal stem cells can in fact initiate tumours if they harbour cancerous mutations. Several other studies have confirmed this by deleting *Apc* or stabilising ß-catenin in stem cells by activating stem cell restricted Cre recombinases: (Sangiorgi and Capecchi, 2008; Zhu *et al.*, 2009; Powell *et al.*, 2014).

In contrast, when researchers have tried to employ non-stem cells as tumour initiating cells *in vivo* it has been obvious that they have much lower tumour initiation capabilities than stem cells. This was emphasised by Barker *et al.* (2009). By using low level intestinal epithelium specific Cre recombinase activity (Ireland *et al.*, 2004), Barker *et al.* deleted *Apc* in TA cells. By doing so, they saw a low penetrance of tumours, although some ß-catenin foci and microadenomas were retained in the upper crypt over time (Barker *et al.*, 2009). In addition, deletion of *Apc* combined with oncogenic activation of Kras in Alpi+ enterocytes did not give rise to tumours *in vivo* (Tetteh *et al.*, 2016). Similarly, deletion of *Apc* in tuft did not provoke any tumour formation. However, dextran sodium sulfate (DSS) exposure (an inflammation inducing agent) lead to multiple colonic tumours arising from tuft cells with *Apc* deletions (Westphalen *et al.*, 2014). In addition, Schwitalla and colleagues constructed a mouse model where Cre is only active in non Lgr5 cells (X-Box Binding Protein 1, *Xbp1*^{CreErt}) (Schwitalla *et al.*, 2013). By using this model they showed that stabilised ß-catenin alone was not sufficient to drive tumour initiation in non-stem cells, however, combining stabilisation of ß-catenin with loss of IκBα (inhibitor of the NF-κB pathway) lead to de-differentiation and tumour initiation of Lgr5 negative cells (Schwitalla *et al.*, 2013).

Together these results suggest that intestinal stem cells have a greater tumour formation capability than non-stem cells. However, differentiated progenitor cells are able to form tumours if they are exposed to intestinal injury/inflammation. This leads to the hypothesis that progenitors will mostly initiate tumours if they fall back into the stem cell pool, thus underlining that stem cells almost certainly are the cell of origin in colorectal cancer. Nevertheless, it is not clear whether cancer drivers promote stem cell contribution from different progenitors or not.

1.7.2 Biased stem cell competition could fuel tumour initiation

As described earlier stem cell drift is determined by a small number of cells and thus it is interesting to speculate that perturbations can have big effects on stem cell homeostasis. An exciting paper set up to explore how different oncogenic alterations affect stem cell behaviour. Vermeulen *et al.* 2013, used low Cre recombinase activity (Kemp *et al.*, 2004) to express oncogenic Kras, p53 or delete of *Apc* in less than one stem cell per crypt. By examining the size of oncogenic stem cell clones over time it was possible, via mathematic modelling, to determine if such clones had any competitive advantage over their WT neighbours. These results showed that heterozygous and homozygous *Apc* deleted stem cells had a 69% and 79% chance of replacing their neighbours, respectively, during clonal drift (Vermeulen *et al.*, 2013). In addition, oncogenic p53 had little effect on clonal growth unless mice were exposed to DSS where after p53 clones had a 58% chance of replacing their neighbours (Vermeulen *et al.*, 2013). Lastly, stem cells expressing oncogenic Kras had a 78% chance of outcompeting their WT neighbours over time (Vermeulen *et al.*, 2013). Together, these observations demonstrated that oncogenic alterations can alter stem cell behaviour and confer competitive bias to such stem cells during clonal drift. This means that stem cells harbouring such oncogenic mutations are more likely to fix in the crypt and thus be present in the intestinal epithelium long-term.

Such biased stem cell behaviour was later confirmed in the human colon. The same paper that studied the neutral mPAS clonal mark in the human colon also studied other biased clonal marks (Nicholson *et al.*, 2018). By screening for loss

of X-chromosome linked proteins the paper identified loss of Stromal Antigen 2 (STAG2) as a clonal mark. Quantification of clones in human colonic tissue allowed them to identify PPCs and WPCs in patients of different ages. If STAG2 clones had a competitive bias it would be expected that the transition from PPC to WPC would be accelerated, compared to that of neutral mPAS clones. Thus, comparison of WPC/PPC ratios showed that STAG2 clones had a strong competitive advantage during monoclonal conversion (Nicholson *et al.*, 2018). This underlines that biased competition also occurs in the human colon and thus this could be the first event in colorectal cancer development.

1.7.3 Tumour origin

In 1987 Fearon and colleagues examined the level of X-inactivation in colorectal tumours from females to determine tumour clonality. This study concluded that all tumours analysed (N=50) were of monoclonal origin (Fearon *et al.,* 1987). This was later confirmed *in vivo*, by studying the staining pattern of Glucose-6-phosphate dehydrogenase (G6PD) in tumours from heterozygotes G6PD deficient mice induced with mutagens or by using Lgr5 driven multicolour lineage tracing system in tumour-prone animals (Griffiths *et al.*, 1989; Schepers *et al.*, 2012). However, several studies have established that colorectal tumours can also be of poly-clonal origin. In 1997, a study used Y-chromosome *in situ* hybridisation in a XO/XY male with FAP to show that 76% of microadenomas in this patient were polyclonal (Novelli *et al.*, 1997). Later, Merritt *et al.* used Apc^{min} mice chimeric for a lineage tracing allele to show that 22/260 tumour analysed were polyclonal (Merritt *et al.,* 1997).

Thus, it seems that most tumours arise from a single crypt, or a single patch of crypts. However, that does not mean that tumours do not encode different mutational clones. Such clones could be acquired through successive mutational clonal sweeps to populate tumours entirely during progression. If so, little to none intra-tumour heterogeneity would be present in tumours. However, intra-tumour heterogeneity in human colorectal tumours has been confirmed by sampling single tumour glands from two sides of the same tumour. The first report to study intra-tumour heterogeneity submitted such samples to genomic

sequencing and methylation pattern analysis to show that there was a high degree of clonal heterogeneity between different tumour glands from the same tumour (Sottoriva *et al.*, 2013). This concept was further explored by the same first author in 2015. By using the same tumour sampling technique and advanced computational modelling the authors could conclude that intra-tumour heterogeneity occurred early in tumour growth and that selective clonal sweeps were rare after the transition to an advanced tumour because of spatial constraints (Sottoriva *et al.*, 2015). From these results a new model of tumour initiation arose: the “Big Bang” model which dictates that colorectal tumours grow as a single expansion of numerous intermixed sub-clones when a critical molecular threshold has been reached (Sottoriva *et al.*, 2015).

Thus, it appears that colorectal tumours originate from tissue that is “primed” with oncogenic clones before overt tumour formation, so-called field characterisation, field cancerisation or field effect.

1.7.4 Field effect primes cancer initiation

Somatic mutations accumulate in normal cells with age and this is thought to be an unavoidable trait of self-renewing tissues. Age accumulated somatic mutations have been debated to mainly arise from intrinsic replication errors thus making ageing the biggest risk factor of cancer development (Tomasetti, *et al.*, 2017). This is highlighted in a study showing that the lifetime risk of developing cancer is highly correlated with the total number of stem cell divisions the tissue has undergone (Tomasetti and Vogelstein, 2015). A field effect can be described in many ways but for the purposes of this thesis they will be defined as: “*the expansion of cells with acquired pro-tumorigenic mutations that does not produce morphological alterations but predisposes to cancer formation*” (Patel *et al.*, 2015).

1.7.5 Field effect in different tissues

A study that examined the total number of mutations in different types of tumours inferred that half or more of all tumour mutations are present in the tissue before apparent tumour formation (Tomasetti, *et al.,* 2013). Thus, the presence of pro-oncogenic fields seems to be a general trait in many different tissues.

That is exemplified in a variety of papers. For instance, ultra-deep sequencing of eye-lid skin, has shown that physiological and histological normal skin carry as many as 0.29 known cancer driver mutations per cell (Martincorena *et al.*, 2015). Analysis of clone sizes showed that some mutant clones were significantly larger than synonymous clones, suggesting increased selection and clonal expansion of some mutations (Martincorena *et al.*, 2015). However, a follow up study later used mathematical modelling to predict that out of 1338 mutant clones identified by Martincorena *et al*., only 6 had clonal sizes that would imply a non-neutral expansion (Simons, 2016). Thus, it appears that most clones, both mutant and synonymous, behave neutrally in the skin epidermis. Together, these studies elegantly showed that fields, in the form oncogenic clones, are present before the onset of tumour formation. Within such fields the pro-oncogenic population is expanded.

Martincorena *et al.* later applied their technology to oesophageal samples from healthy organ donors of different ages. Here they showed that somatic mutations accumulate with age and that some oncogenic mutations are under strong selection in the aging oesophageal tissue (Martincorena *et al.*, 2018). An interesting finding from this study was that a mutation that is only rarely found in oesophageal cancer was found in enormous amounts in the normal tissue (Martincorena *et al.*, 2018). This proves that some non-cancerous mutations can also be subject to positive selection and could be of protective nature.

However, field effects have not only been observed in epithelial tissues. An important study that sequenced normal control blood samples from a variety of cancer patients showed that somatic mutations also accumulate in the haematopoietic system with age (Xie *et al.*, 2014). A subset of these samples also contained oncogenic mutations that are known to cause clonal expansion and blood cancer (Xie *et al.*, 2014).

Together, these studies underline that some oncogenic fields are present before tumorigenic transformation thus priming different tissues to reach the critical threshold for cancer initiation (Hanahan and Weinberg, 2011).

1.7.6 Intestinal field effect

Importantly, field effect is also a known concept in human colonic tissue. Colonic mutational fields are thought to arise through the process of acquiring a mutation in an intestinal stem cell followed by mutated clone fixation and crypt fission (Figure 1.4) (Patel *et al.*, 2015). Many publications have shown that oncogenic mutational fields are present in colonic tissue prior to tumour formation: (Chen *et al.*, 2005; Galandiuk *et al.*, 2012; Nicholson *et al.*, 2018).

One of the most reported field effect causing alterations is the oncogenic *KRAS* mutation (discussed later). It has been shown that *Kras* mutations confer advantaged behaviour to stem cells and that mutated *Kras* crypts have an increased fission rate in the murine intestine (Vermeulen *et al.*, 2013; Snippert *et al.*, 2014). Together, this gives $Kras^{G12D}$ mutations the potential for creating pre-neoplastic fields *in vivo*.

1.8 Ras

The *RAS* genes were first discovered as being involved in the oncogenic Harvey murine sarcoma virus and the Kirsten murine sarcoma virus (Cox and Der, 2010) and they constitute the most frequently mutated oncogenes in human cancer (Cox *et al.*, 2014).

The *RAS* family constitutes: *Neuroblastoma RAS (NRAS), Harvey RAS (HRAS)* and *KRAS*. RAS proteins are membrane bound GTPases and with the help of guanine nucleotide exchange factors (Gefs) and GTPase activating proteins (Gap) they switch between their inactive GDP bound form to the active GTP bound form (Hobbs *et al.,* 2016) (Figure 1.5). Active Ras binds to and activates multiple different targets to induce cellular responses such as proliferation, protein synthesis and survival (Pylayeva-Gupta, *et al.*, 2011). The

two main signalling cascades that Ras proteins activates are: the Phosphoinositide 3-kinae (PI3K) / Akt / mammalian target of rapamycin (mTor) pathway and the Raf / mitogen activated kinase (Map/Mek) / extracellular regulated kinase (Erk) pathway (Pylayeva-Gupta *et al*., 2011) (Figure 1.5).

1.8.1 Oncogenic Ras and tumour initiation

KRAS is the predominant oncogenic *RAS* family member involved in colorectal cancer. There are three main oncogenic hot spots in the *KRAS*: glycine-12 (G12), glycine-13 (G13) and glutamine-61 (Q61). Oncogenic mutations in these hotspots render KRAS constitutively active (Figure 1.5).

The transformative capabilities of RAS proteins have been debated. Early *in vitro* studies showed that RAS alone is not enough to transform fibroblast and kidney cells in culture (Land *et al.*, 1983; Ruley, 1983). However, combining RAS with other oncogenes or loss of tumour suppressor genes lead to transformation of the cell lines (Land *et al.*, 1983; Ruley, 1983). In contrast, the first *in vivo* reports to study *Ras* function overexpressed Hras during mouse development or in adult transgenic mice. These investigations showed that Hras alone could lead to tumour formation in the foetal exocrine pancreas and in the adult mammary and salivary glands and lymphoid tissue, suggesting that Ras expression alone can initiate tumours *in vivo* (Quaife *et al.*, 1987; Sinn *et al.*, 1987).

KRAS is mutated in up to 50% of all colorectal tumours and thus many efforts have been made to understand its tumour initiating potential in the intestine (Pylayeva-Gupta, *et al*., 2011). In an interesting report, Janssen and colleagues over-expressed oncogenic $Kras^{G12V}$ in the entire intestinal epithelium and saw tumours occurring with high penetrance (Janssen *et al.*, 2002). These tumours were negative for nuclear ß-catenin suggesting that tumours arising from Kras activation do not follow the classical *Apc* route. However, with the development of inducible $Kras^{lslG12D/V}$ animals, which encode the G12D or the G12V mutation at the murine *Kras* locus, it was possible to explore what effect endogenous

$Kras^{G12D/V}$ expression has on intestinal tumour development (Jackson *et al.*, 2001; Guerra *et al.*, 2003). This was an important development since *KRAS* is normally mutated but not overexpressed in human colorectal tumours. Two separate studies showed that long term expression of $Kras^{G12D}$ from the endogenous allele leads to hyperproliferating intestinal crypts but not tumour formation (Haigis *et al.*, 2008; Feng *et al.*, 2011). However, intestinal $Kras^{G12V}$ expression appeared to have no immediate morphological effect, although, after long-term exposure (>500 days) mice did develop intestinal tumours (Sansom *et al.*, 2006). Interestingly, expression of oncogenic Kras in combination with loss of tumour suppressor genes greatly increases tumour burden and tumour aggressiveness in the murine intestine (Jackstadt and Sansom, 2016).

Based on these observations, oncogenic *Kras* appears to be a poor initiator of intestinal cancer *in vivo*.

1.8.2 Colonic Kras field effect

Oncogenic *KRAS* mutations in codons 12 and 13 have been discovered in normal human colon mucosa by highly sensitive PCR methods (Dieterle *et al.*, 2004; Kraus *et al.*, 2006; Parsons *et al.*, 2010). Parsons *et al.* predicted the *KRAS* mutant fraction (i.e how many cells had the mutation per sample) in normal mucosa and tumour tissue (Parsons *et al.*, 2010). By doing so they showed that normal mucosa contains a significantly lower *KRAS* mutant fraction than tumours, suggesting that some tumours could arise from *KRAS* fields (Parsons *et al.*, 2010). However, because of the nature of the PCR based experiments is was hard to extrapolate field growth from these results.

The nature of mutant *KRAS* field growth has mostly been studied in patients with Ulcerative Colitis or Chron's Illeocolitis. Such patients are subject to chronic colonic inflammation that causes regeneration and thus increased replication and/or direct DNA damage (Baker *et al.*, 2018). In such conditions clonal expansion is greatly enhanced. In addition, such patients are often subject to colonic resections making it possible to examine mutation burden in a spatial

manner. To this end, one publication showed that a Ulcerative Colitis patient carried oncogenic *KRAS* mutations in 3 independent tumours and normal crypts which were located up to 14 cm apart (Leedham *et al.*, 2009). These results indicate that *KRAS* fields can cover big colonic areas and shows that mutations in *KRAS* could be an early oncogenic event. In addition, an important study that followed the same Chron's Illeocolitis patients over several years could map the spread of mutational fields over time. In one patient a *KRAS* mutation was present in normal and dysplastic crypts 4 years prior to tumour formation (that also carried the *KRAS* mutation), suggesting this mutation was the initial oncogenic event (Galandiuk *et al.*, 2012).

Recently, Nicholson *et al.* showed that mutated *KRAS* fields also increase in size with age in normal non-inflammatory colonic tissue. Here, the authors used samples with a known number of crypts to sequence the oncogenic region of *KRAS*. Based on the mutant allele frequency it was possible to extrapolate the number of mutant crypts per sample. The results showed that mutated *KRAS* fields are present in 10% of healthy human colons and increase in size with age with the largest field recorded as consisting of approximately 100 crypts (Nicholson *et al.*, 2018).

From these studies it appears that colonic *KRAS* mutations are present as fields in some individuals prior to overt tumour formation. Such areas might predispose to cancer development. However, the cell of origin for such fields and the pro-oncogenic mechanism of them is still unknown.

Aim

The Aim(s) of this thesis is to determine the origin and properties of *Kras*^{G12D} pro-oncogenic fields in the intestinal epithelium. Specifically, the project aims to:

1. Determine whether Kras^{G12D} expression affects Atoh1+ secretory progenitor stem cell contribution both during tissue homeostasis and after Lgr5+ cell depletion. This will aid the understanding of how oncogenic *KRAS* fields arise.
2. Determine the stem cell dynamics, defined by number of clonogenic stem cells per crypt and stem cell replacement rate, in *Kras*^{G12D} mutant crypts of *Kras*^{G12D} fields. This will help the to interpret the pro-oncogenic properties of mutated *KRAS* fields.
3. Create a transgenic mouse encoding an inducible Dre recombinase to be used for lineage tracing in sequential mutation models of colorectal cancer *in vivo*.

2 Materials and methods

2.1 Molecular biology and *in vitro* techniques

2.1.1 Genomic DNA (gDNA) and RNA isolation

gDNA was isolated from mouse ear-biopsies or form fractionated intestinal epithelium using the Promega ReliaPrep™ gDNA Tissue Miniprep System (cat. A2052) according to manufacturer's protocol.

2.1.2 Polymerase chain reaction (PCR)

PCR was carried out using the NEB Q5 High-Fidelity DNA Polymerase using manufacturer's instructions. Briefly, 5 μl 5X Q5 Reaction Buffer was mixed with 0.5 μl 10 mM dNTPs, 5 μl primer-mix (forward primer + reverse primer 20 μM final concentration), 13.25 μl H_2O, 0.25 μl Q5 High-Fidelity DNA Polymerase and 1 μl gDNA. For PCR primers see Appendix 1.

2.1.3 Real time Quantitative RT-qPCR

Gene copies were analysed using TaqMan gene expression probes (Thermo Fisher). Briefly, reactions consisted of: 5.4 μL TaqMan Fast Universal PCR Master Mix (Applied Biosystems cat. 4352042), 0.6 μL TaqMan probe and 6 μL gDNA of 5 ng/μL stock. Total reaction volume was 12 μL. Amplification was carried out in a 384 well-plate using a QuantStudio 6 instrument (Thermo Fisher).

2.1.4 Restriction digest, agarose gel purification and ligation

Plasmids and PCR amplified inserts were digested with restriction enzymes following manufacturer's instructions (New England Biolabs). Volume for diagnostic digests were 10 μl and 100 ng of DNA was used. For experimental digests 1 μg of plasmid DNA was used in a reaction volume of 50 μl. All fragments were separated by agarose gel electrophoresis. If fragments were used for subcloning they were cleaned using Zymeclean Gel DNA Recovery kit (Zymo

Research) according to manufacturer's instructions. Ligation reactions were carried out using the T4 DNA ligase with a 1:3 vector to insert molar ratio.

2.1.5 Transformation and *E. coli* expansion

For E. *coli* transformation and expansion XL2-Blue Ultracompetent cells (Stratagene) were used. Briefly, 1 μl of plasmid DNA was added to 30 μL of XL2-Blue cells and kept on ice for 30 minutes. The E. *coli*/DNA suspension was then heat-shocked for 30 seconds at 42°C and 0.9 mL Lysogeny Broth (LB) medium was added. 100 μl of the cell-plasmid suspension was plated on LB agar plates containing appropriate antibiotics and incubated overnight at 37°C. Single colonies were picked and expanded in 5-100 mL of LB medium containing 100 μg/mL Ampicillin at 37°C 2000 RPM overnight. Plasmid DNA isolation was carried out with QIAprep Spin Miniprep kit or QIAGEN Plasmid Maxi Kit according to manufacturer's instructions (Qiagen).

2.1.6 Plasmids and oligonucleotides

Dre coding sequence (Anastassiadis *et al.,* 2009) was synthesized by Invitrogen (construct ID 15ACQONP). The human progesterone receptor (Pr) sequence was amplified from the Cre^{PR2} vector kindly given to us from Günther Schütz laboratory (Kellendonk *et al.,* 1996). The Pr fusion site on Dre has previously been described by Anastassiadis *et al.,* 2009.

2.1.7 Cloning of R_Dre and R_Dre^{Pr} plasmids

The *Rosa26* (*R26)* targeting vector used in this study was kindly given to our lab from Anton Berns and has previously been described in: (Vooijs *et al.*, 2001). The vector contains a total of ~13 kilobases (KB) of the endogenous murine *R26* sequence located at chromosome 6 and is designed for homologous recombination in mouse embryonic stem cells (mESCs). The vector is built so that a 5' ~3 KB homology *R26* sequence precedes a NheI cloning site. This cloning site is followed by a Phosphoglycerate Kinase (PGK) promoter driven Puromycin selection cassette which in turn is followed by ~10 KB 3' *R26* homology sequence. The splice acceptor, Dre and/or Pr coding sequence and a Bovine Growth Hormone (BGH) poly-adenylation signal was merged together by

In-fusion cloning (Takara cat. 121416) in pUC19 (Addgene #5005). The In-fusion reaction contained two or three inserts: the Dre coding sequence (insert 1), the Pr coding sequence (insert 2) and the BGH coding sequence (insert 2 or 3). The consensus splice acceptor sequence was added to the forward primer sequence of insert 1. All inserts were PCR amplified from plasmids and treated with DpnI for 30 minutes at 37°C to remove any residual vector. After cleaning the PCR+DpnI reaction (Zymogene clean and concentrate, cat. D4004) the inserts and the restriction digested vector were mixed in 4 μl (200 ng total) and 1 μl of In-Fusion enzyme and was added. The reaction was incubated for 30 minutes at 50°C and used for E. *coli* transformation. After assembly in pUC19, the constructs were isolated from the vector backbone by restriction digest and agarose gel purification and ligated into a *R26* targeting vector at the NheI cloning site. For in-fusion primer sequences see Appendix 1 and for vector map see Appendix 2.

2.1.8 mESC targeting and colony picking

This procedure was carried out by the Xiangang Zou in the CRUK CI Genome Editing Core. Briefly, C57BL/J mouse embryonic stem cells (mESCs) were electroporated with linearized plasmids. After electroporation the mESCs were plated on feeder cells under puromycin selection (1 ng/μl). Single mESC colonies were picked and transferred to 96-well plates individually. Each plate was grown up and split into two identical plates. One plate was stored at -80°C to be use for later 8-cell stage embryo injections. The duplicate plate was used for gDNA extraction and PCR amplification to identify correctly targeted clones.

2.1.9 mESC gDNA extraction and clone verification

60 μl of DNA extraction buffer (5M NaCl, 1M Tris-HCl ph 7.5, 0.5M EDTA, 20% SDS, 0.2 mg/mL proteinase K) was added to each well of mESCs on 96-well plates and the plates were then incubated at 37°C overnight. After lysis, all wells were transferred into 8-strip PCR tubes. 130 μl 100% cold ethanol was added to each well to precipitate DNA. After precipitation the buffer was removed, and DNA pellets were air-dried. 50 μl of water was added to each pellet to dissolve DNA. To verify 5' and 3' arm insertions PCR amplification was carried out using SequalPrep Long PCR kit (Applied Biosystems, cat. A10498) according to

manufacturer's instructions with F1 and R1 primers (5' arm) or F2 and R2 primers (3' arm) (Appendix 1 for primers).

2.1.10 R_Dre and R_Dre^{Pr} mouse model creation

Figure 2.1 shows the workflow of creating mouse models from targeted mESCs. Briefly, positive C57BL/J mESC clones were injected into fertilized Balb/c oocytes. The oocytes were then placed in pseudo-pregnant Balb/c females. As C57BL/J mESC contributed to the embryo the resulting litters were chimeric for both C57BL/J and Balb/c origin. This manifested as mosaic black/agouti/white coats in resulting litters (Figure 2.1). Pups of which the C57BL/J mESCs contributed to >70% of the coat colour were chosen for further breeding and crossed to C57BL/Js (Figure 2.1). The C57BL/J black coat colour is dominant, therefore pups with white coats resulting from this breeding were considered as negative for C57BL/J mESC origin. ~50% of black coated pups carried either Dre or Dre^{Pr} upon genotyping. These mice are considered to be the origin of the Dre and Dre^{Pr} lines and were used for further sub-breedings (Figure 2.1).

2.1.11 Organoid establishment and maintenance

Organoids were prepared from the colon as previously described (Sato *et al.*, 2009). Briefly, the colon was flushed with Phosphate-buffered saline (PBS) and mesentery fat was carefully removed. The tissue was then opened up longitudinally and cut into 0.5 cm pieces. Tissue was washed in 10 mL cold PBS 10 times where after they were resuspended in 25 mL of PBS/25 mM EDTA and incubated on ice for 30 minutes. PBS/EDTA was removed and 10 mL PBS was added to tissue and samples were hand shaken thoroughly for 10 seconds. This was repeated 4 times to collect crypt different crypt fractions. Crypt enriched fractions (fraction 2-4) were combined and pelleted by centrifugation (1200 RPM, 5 minutes, 4°C). Crypts were then passed through a 70 μM filter and collected by centrifugation (600 RPM, 3 minutes, 4°C) and seeded in a 20 μL drop of ice-cold growth factor reduced Matrigel (Corning, cat. 356231) on pre-warmed 24-well plates. Crypts were maintained in the following medium: Advanced DMEM/F12 (ADF) containing: 0.1% Bovine serum albumin (BSA); 1X L-Glutamine (Gibco, cat. 25030-024); 10 mM HEPES; 1X PenStrep (Gibco, cat. 15070-063);

1X N-2 Supplement (Invitrogen cat. 17502-048), 1X B-27 Supplement (Invitrogen cat. 12587-010); 50 ng/mL human EGF (Prepotech, cat. AF-100-15); 100 ng/mL murine NOGGIN (Prepotech, cat. 250-38); 50% Wnt-conditioned medium (made in-house) and 10% R-spondin-1 conditioned medium (made in house) and kept at 37°C with 5% CO_2.

2.1.12 Library preparation

Organoids were released from Matrigel after 5 days of culture by gently dissociation in ice cold PBS. Single organoids were picked under a light microscope using an EZ-Grip pipettor (RI Life Sciences, ref. 8-72-2802) with 200 μm EZ-TIP (RI Life Sciences, cat. 7-72-2200/1). Single organoids were then transferred directly into 96-well plates containing 20 μL Picopure extraction buffer (cat. KIT0103) and were incubated at 65°C for 3 hours. After lysis, 100 μL 1XTE buffer and 80 μL Ampure XP beads (Beckman Coulter, cat. A63882) were added to each reaction. After 5 minutes of binding at room temperature a magnetic 96-well rack was used for magnetic bead separation. Supernatant was removed and gDNA was washed twice in 75% ethanol and eluted in 20 μl nuclease-free water. Libraries were prepped using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina according to manufacturer's instructions (NEB, cat. E6177). Briefly, the low DNA input protocol was used, fragmentation was carried out for 30 minutes and 11 cycles of PCR was performed using Agilent SureSelect indexing primers. 2 μl of each library was quantified on a 4200 TapeStation System (Agilent) and via Qubit™ dsDNA HS Assay kit (Invitrogen, cat. Q322854). Libraries with similar DNA content from the Tapestation and Qubit analysis were used for exome capture (6 libraries per mouse).

2.1.13 Exome capture and sequencing

1000 ng library DNA was concentrated using a vacuum Concentrator plus (Eppendorf) and exomes were hybridized, captured and amplified using the SureSelect^{XT} Low Input Target Enrichment System for Illumina Paired End Multiplexed Sequencing Library (Agilent Technologies) according to manufacturer's instructions. After exome capture, the libraries were quantified on the 4200 TapeStation System (Agilent Technologies) and by KAPA Library

Quantification Kit according to manufacturer's instructions (KAPA Biosystems Roche, ref. 07960338001). Libraries were pooled (20 nM each) and sequenced on the Illumina HiSeq 4000 (1 lane, paired-end 150 base-pair (bp)).

2.2 *In vivo* and mouse tissue techniques

2.2.1 Mouse models

Rosa26 rox STOP rox tdTomato (R_tdTom^{rsr}): the R_tdTom^{rsr} were generated in house by germline deletion of the LoxP STOP LoxP (lsl) cassette from R_tdTom^{rsrlsl} mouse (Jackson Laboratories, stock no. 021876). This was carried out by Xiangang Zou and Guocheng Lan in the CRUK CI Genome Editing Core. R_tdTom^{rsrlsl} early embryos were cultured with the cell permeable TAT-Cre *in vitro* and embryo transfer was performed as described by (Ryder *et al.*, 2014). PCR screening was used to ensure lsl cassette deletion.

Rosa26 lsl tdTomato (R_tdTom^{lsl}): R_tdTom^{lsl} was bought from the Jackson Laboratory (stock no. 007908) and described here: (Madisen *et al.*, 2010).

Rosa26 Dre (R_Dre) and *Rosa26* Dre^{Pr} (R_Dre^{Pr}): were generated in house at the CRUK CI Genome Editing Core in the Biological Resource Unit by mESC electroporation and homologous recombination of *R26* targeting vectors and subsequent pro-nuclear injections.

Kras^{fsfG12D} and *Kras*^{lslG12D}: was obtained from The Jackson Laboratory (stock no. 008653 and 019104).

Villlin^{Flp}: was kindly provided to us from Dierte Saur Laboratory, Germany.

Atoh1^{CreERT2}: was described and generated by Tomic *et al.,* 2018.

Ctnnb1$^{flox(ex3)}$: was described by (Harada *et al.,* 1999) and was kindly given to us by Richard Gilbertson at CRUK CI.

Ah^{CreErt}: was described and generated by: (Kemp *et al.,* 2004).

Cyp1A1^{FlpO}: was generated in the Winton lab at CRUK CI. In this mouse the FlpO expression cassette is under transcriptional control of the ß-naphthoflavone inducible *Cyp1A1* promoter. This model has not been yet been published.

Villin^{CreErt}: was described by (El Marjou *et al.*, 2004).

2.2.2 Genotyping

Genotyping of animals was performed on ear-biopsies from 3-week old mice. Most genotyping was carried out by Transnetyx. For primers and probes see Appendix 3. Determination of allelic zygosity of $Ah^{CreErt2}$ and $Villin^{Flp}$ was carried out by performing RT-qPCR copy-number assay of ear biopsy gDNA. Determination of allele zygosity of *Atoh1*^{CreErt2} was carried out by PCR amplification of ear biopsy gDNA as previously described (Tomic, *et al.*, 2018).

2.2.3 Crypt isolation of small intestine and colon

After dissection the small intestine and colon were flushed with cold PBS to remove faeces and a 10 cm piece of the small intestine and the entire colon was fed on to a glass pastor pipette and inverted. The inverted intestinal tube was then fed on to a glass spiral and washed in PBS. The spiral was then covered in 50 mL Hank's Balanced Salt Solution (HBSS) with 1 mM Na_2-EDTA and 2 mM NaOH and subjected to 6 cycles of 10 minutes 37°C incubation followed by 10 times 5 seconds pulses of mechanical vibration using a Chemap transformer /vibrator (110/240 Volt). After mechanical vibration the HBSS buffer was collected as crypt-fraction 1-6. Fractions were pelleted at 1200 RPM and the highly crypt enriched fractions (normally 2-6) were combined and subjected to PCR and/or RT-qPCR analysis or dissociated into single cells for flow cytometric analysis or sorting (see below).

2.2.4 Flow cytometry for EpCam and tdTomato

Isolated crypts were dissociated into single cells by resuspending crypt-pellets in 20 mL of 0.05% Trypsin-EDTA (Gibco cat. 25300054). Crypt-Trypsin suspension was incubated at 37°C for 7 minutes and after each minute the suspension was subjected to 10 seconds of vigorous shaking. After 7 minutes, 30 mL of PBS with 2% foetal bovine serum (FBS, Thermo Fisher A38401) was added and the cell suspension was filtered through a 70 μM filter. Single cells were pelleted by centrifugation (1200 RPM, 5 min, 4°C) and washed twice in PBS 2% FBS before being subjected to antibody staining with EpCam Alexa Fluor 647 (BioLegend clone G8.8 0.5 mg/mL, 1:2000) and DAPI (1 mg/mL stock, 1:1000). All flow

cytometry was performed on a BD FACSAria or a BD LSRFortessa Instrument (BD Biosciences).

2.2.5 Paraffin embedding, block sectioning and histochemical staining

All tissue samples were fixed for 24 hours in 4% natural buffered formalin (NBF) or 4% paraformaldehyde (PFA) and transferred into 70% ethanol. All paraffin embedding and block sectioning as well as histochemical staining for Haematoxylin and Eosin (H&E) and Pas/Alcian Blue was carried out at the CRUK CI histology core facility. For H&E staining sections were dewaxed and hydrated, incubated in Haematoxylin for 2 minutes, washed in water, incubated in 2% Acid Alcohol for 20 seconds, washed and incubated in 1% Eosin for 7 minutes, all in a ST5020 Multi-stainer (Leica). For Pas/Alcian Blue staining sections were dewaxed and hydrated in a ST5020 Multi-stainer (Leica), incubated in Alcian Blue for 10 minutes, washed in water, incubated in 0.5% periodic acid for 5 minutes, washed 3 times, incubated in Schiffs reagent for 30 minutes, washed 3 times and counterstained in Mayers Haematoxylin for 30 seconds.

2.2.6 Immunohistochemistry (IHC)

IHC for BrdU, tdTom, GFP, Ki-67, Lysozyme and ß-catenin were carried out by the CRUK CI histology core facility unit, for antibodies and antigen retrieval see Appendix 4.

Staining for phosphorylated Mek (pMek) and Dusp6 was carried out by manual IHC. Antigen retrieval was performed with 10 mM citrate buffer (ph=6) in a pressure cooker. Tissue sections were then incubated in 3% H_2O_2 (in Methanol) for 15 minutes and washed 3 times for 5 minutes in PBS/0.05% Tween 20 (PBS-T). Tissue sections were then blocked in 10% donkey serum (Sigma, D9663) (in PBS-T) for 30 minutes and primary antibodies for pMek or Dusp6 (Cell Signalling, cat. 2338 or Abcam, ab76310, dilution 1:50) were added and incubated on sections over night at 4°C. Sections were then washed 3 times in PBS-T and incubated with biotinylated donkey anti-rabbit secondary antibody (Jackson ImmunoResearch cat. 711-065-152, dilution 1:500) for 40 minutes at room temperature. Sections were washed and incubated with ABC HRP Kit (Vectastain, cat. PK-6100) for 40 minutes at room temperature. Next, sections

were washed and DAB Chromogen substrate (Dako, cat. K3468) was added for dye development for 5 minutes. Counterstain and dehydration were performed on the ST5020 Multi-stainer (Leica).

2.2.7 Crypt-counting: BrdU, lysozyme and Pas/Alcian Blue

All histology sections were counted manually under a light microscope at either 20X or 40X magnification. BrdU+ cells were scored in 25 fully sectioned half crypts. A score of 1 or 0 was given to a positive or negative nucleus, respectively, at all positions up the crypt until the crypt-villus border (small intestine) or the crypt top (colon) was reached. Lysozyme+ cells were scored in 25 fully sectioned crypts. Pas/Alcian Blue+ cells were scored in 15 fully sectioned half-villi. The number of nuclei from the villus bottom to the top and number of Pas/Alcian Blue+ cells clearly associated with a nucleus was quantified.

2.2.8 Subcutaneous insertion of Ru486 slow release pellet(s) in mice

This procedure was carried out by Matthew Clayton at the Biological Resource Unit Core at CRUK CI. Mice receiving pellet insertion surgery were all 8-10 weeks of age. Mice were placed under general anaesthesia and a 1 cm subcutaneous incision was made on the back flank of each mouse in which 1-3 Ru486 pellets (Innovative Research of America, 10 mg/pellet, 90 days release, cat. NX-999) were placed. The wound was closed with surgical glue.

2.2.9 Intra-peritoneal (I.P) and oral gavage administered drugs

Ru486 (Sigma, cat. M8046) was dissolved in 100% ethyl alcohol (Honeywell, cat. E7023) to make a 50 mg/mL solution and further diluted into a working solution of 10 mg/mL in 40% cyclodextrin/H_2O. ß-naphthoflavone (Sigma, cat. N3633) was dissolved to a working solution of 8 mg/mL or 2 mg/mL in corn oil (Sigma, C8267). All 40 mg/kg and 8 mg/kg I.Ps were made from 8 mg/mL or 2 mg/mL stocks, respectively. Tamoxifen (Sigma, cat. T5648) was dissolved in 100% ethyl alcohol (Honeywell, cat. E7023) to make a 200 mg/mL solution and further diluted to a working solution of 20 mg/mL, 5 mg/mL or 1 mg/mL in sunflower oil (Sigma S5007). All 2-4 mg, 0.5 mg and 0.15 mg I.Ps were made with 20 mg/mL, 5 mg/mL or 1 mg/mL stocks respectively. Diphtheria toxin: was dissolved in saline at 50

μg/ml working solution and administered at 50 μg/kg. AZD6244: was dissolved in 0.5% HPMC + 0.1% Tween-80 to a working solution of 6.25 mg/mL and administered at 25 mg/kg.

2.2.10 Whole mounting of mouse small intestine and colon

The small intestine and colon were dissected from mice and flushed with cold PBS. The gut was opened longitudinally, pinned out and fixed for 3 hours at room temperature in 4% PFA. After fixation, whole mounts were washed in PBS and incubated with de-mucifying solution (3 mg/mL DTT, 20% ethanol, 10% glycerol, 0.6% NaCl, 10 mM Tris, pH 8.2) for 20 minutes at room temperature. Mucus was removed by pipetting and wholemounts were stored in PBS at 4°C.

2.2.11 Dapi and EpCam staining of whole-mounted tissue

1-2 cm of intestinal tissue was permeabilised in 0.1% PBS-Tween20 for 24 hours at 4°C. EpCam Alexa Fluor 647 (BioLegend clone G8.8 0.5 mg/mL, 1:100) or DAPI (1 mg/mL stock, 1:100) was then added and the tissue was incubated for 3 days at 4°C. After antibody/DAPI incubation tissue was washed in 0.1% PBS-Tween20 for 24h where after it was subjected to imaging on a TCS SP5 confocal microscope (Leica).

2.2.12 Tumour clearing

Whole-mounted tissue was fixed in 4% PFA overnight at 4°C. Hereafter, tissue was washed in PBS for 8-hours at room temperature. Tissue was then incubated in CUBIC-1A solution (10% Triton, 5% NNNN-tetrakis (2-HP) ethylenediamine (Sigma, 122262), 10% Urea, 25mM NaCl) with DAPI 1:100 (10 mg/mL stock) at 37°C 60 RPM for a total of 5 days. On day 2 and 4 CUBIC-1A + DAPI was refreshed. After CUBIC-1A incubation, tissue was washed in PBS for 2 hours and then placed in Rapi Clear (SunJin Lab., cat. RC152002) and incubated at room temperature until tissue was see-through. Hereafter, tissue was mounted on 1 mm inserts (iSpacer, SinJun Lab.) on glass-slides in Rapi Clear and subjected to imaging on a TCS SP5 confocal microscope (Leica).

2.2.13 Lineage tracing and scoring

For lineage tracing mice were induced with a pulse of ß-naphthoflavone and tamoxifen (40 mg/kg and 0.15 mg, respectively) or Ru486 (50 mg/kg) and tissue was whole mounted at day 4, 7, 10, 14 and 21 after pulse administration. Clone sizes were scored manually under a fluorescent microscope using the 550 nm filter. 2 cm of tissue was mounted muscle side up, on a glass-slide and clone sizes were scored as fractions of 8. All tissue was scored blinded.

2.3 Computational methods

2.3.1 Exome mutation calling

Mutations were called by Lee Hazelwood using a composite sample background of the same type and age as the organoid under detection. The programme Varscan 2.4.3 with quality filter 10 and allele frequency 0.35-0.65 was used to detect somatic mutations.

2.3.2 Clonal collision prediction

Prediction of clonal localisation was carried out by Richard Kemp. A pearl loop-script randomly allocates clones (by a given clonal frequency) on a 400X400 crypt grid. This script assumes that each crypt has 6 neighbouring crypts. Clonal collision and clone patches were quantified manually on print-outs of predicted datasets. For pearl script see Appendix 5.

2.3.3 Determining λ and PR of Atoh1 derived stem cells

This modelling was produced by Edward Morrissey. It relies on the neutral drift model (Lopez-Garcia *et al.* 2010; Vermeulen *et al.* 2013). Previous lineage tracing data-sets were used and by fixing N=7, or λ to 0.25 it was possible to predict how λ or PR, respectively, would affect the WPC/PPC ratio 30 days post lineage trace initiation. In this analysis all 1/8 clone sizes were removed. For R-script see Appendix 6.

3 Determining the effect of Kras^{G12D} expression on the stem cell contribution potential of early secretory progenitors

3.1 Introduction

As discussed in Chapter 1 the intestinal epithelium is maintained by crypt base intestinal stem cells which regenerates the tissue by constantly giving rise to cells of the absorptive and secretory lineages. These different lineages are specified based on differential transcriptional programmes in cells that exit the crypt base. Expression of the transcription factor Atoh1 constitutes the initial “on” switch for the secretory lineage and Atoh1 expressing cells include the earliest secretory progenitors. However, it is now clear that intestinal cell specification is not always restrictive, and that cell fate is a context dependent continuum. In addition, recent studies have shown that Atoh1+ early secretory progenitors contribute to the stem cell pool during injury and homeostasis (Castillo-Azofeifa *et al.,* 2019; Ishibashi *et al*., 2018; Tomic *et al*., 2018), but the effect conferred by pro-oncogenic mutations, such as *Kras*^{G12D}, has not yet been investigated. Additionally, it is known that oncogenic mutations in *KRAS* are present in healthy human colonic tissue (Dieterle *et al.*, 2004; Kraus *et al.*, 2006; Parsons *et al.*, 2010; Nicholson *et al.*, 2018). These mutations exist in fields of up to approximately 100 crypts and such areas could pre-dispose to colorectal cancer in some patients.

Thus, the aim of this chapter is to study the cell of origin of *KRAS* fields by investigating how Kras^{G12D} expression affects cell fate choices of early secretory progenitors.

In this chapter cell fate choices are studied in Kras^{G12D} expressing progenitor cells primed for the secretory path by quantifying their stem cell contribution potential in the small intestine and colon. To develop a suitable *in vivo* model for such studies the *Atoh1*^{CreErt} (Tomic, *et al.*, 2018), *Kras*^{lslG12D} (Jackson *et al.*, 2001) and

R26^{lsltdTom} (Madisen *et al.*, 2010) mouse lines were inter-crossed to generate mice in which both Kras^{G12D} and tdTom expression could be activated in Atoh1+ secretory progenitors. In this model it is possible to trace cell fate choices in early secretory progenitors by activating Cre recombinase in cells that express the Atoh1 transcription factor. Here, stem cell contribution was examined by quantification of stem cell derived clones after saturated pulse-labelling. Such stem cell contribution requires the achievement of two processes acting in series: firstly Atoh1 expressing cells must contribute to the stem cell pool instead of maturing into goblet or Paneth cells (Figure 3.1A) and secondly upon stem cell contribution Atoh1+ stem cells must replace their neighbours to persist as clones in the intestinal crypts over time (Figure 3.1B).

In addition to investigating stem cell contribution this study also aims to investigate the behaviour of Atoh1+ and Atoh1+Kras^{G12D} stem cells by quantifying their monoclonal conversion time (Figure 3.1B). Finally, the Atoh1 derived *Kras*^{G12D} inter-crypt multiplication pattern is also investigated (Figure 3.1C) as well as examination of the tumorigenic potential of long-term Kras^{G12D} activation in Atoh1+ cells.

3.2 Results

The *Atoh1*^{CreErt} allele was crossed to the *Kras*^{lslG12D} allele and the *R26*^{lsltdTom} reporter allele to generate *Atoh1*^{CreErt};*Kras*^{lslG12D}; R26^{lsltdTom} mice referred to as Atoh1;Kras^{G12D};RT. Control animals did not carry the *Kras*^{lslG12D} allele and will be referred to as Atoh1;RT. All animals in this chapter were homozygote for the *Atoh1*^{CreErt} allele and heterozygote for the *Kras*^{lslG12D} and *R26*^{lsltdTom} alleles. Because of differences in clonal frequencies the small intestine and colon will be analysed separately throughout this chapter.

3.2.1 Partial *Kras*^{G12D} and *tdTom* co-recombination is observed in Atoh1+ cells after Cre^{Ert} activation

In the *Atoh1*^{CreErt} mouse the Cre^{Ert} cassette is targeted to the *Atoh1* locus thus restricting Cre^{Ert} expression to Atoh1 expressing cells (Atoh1+) (Tomic, *et al.*, 2018). To be able to analyse the effect of Kras^{G12D} expression in Atoh1+ cells based on reporter expression (tdTom) it is important that the *Kras*^{lslG12D} and *R26*^{lsltdTom} allele co-recombine after Cre^{Ert} activation (Figure 3.2). To examine the level of co-recombination in Atoh1+ cells, Atoh1;RT and Atoh1;Kras^{G12D};RT animals received 3 mg Tamoxifen and 4 days later the small intestine and colon was subjected to flow cytometric cell sorting (Figure 3.3A). To investigate the amount of co-recombination epithelial EpCam+/tdTom- and EpCam+/tdTom+ cells were sorted (Figure 3.3C) and subjected to PCR analysis for *Kras*^{lslG12D} recombination. As seen in Figure 3.3D, *Kras*^{lslG12D} recombination (G12D band) was only present in tdTom+ cells. However, not all tdTom+ cells were recombined at the *Kras*^{lslG12D} allele as un-recombined (LSL) band remained evident in these samples even after Tamoxifen exposure (Figure 3.3D). Thus, partial co-recombination of *Kras*^{G12D} and *tdTom* was observed in Atoh1+ cells after maximum Tamoxifen exposure. This means that a number of tdTom+ cells in Atoh1;Kras^{G12D};RT animals will be *Kras*^{WT} which will negatively skew the data during quantification. Thus, effects observed in Atoh1;Kras^{G12D};RT animals will be underestimated. However, the saturation labelling strategy of these experiments did not allow for further optimisation of *Kras*^{lslG12D} recombination without risking a departure from the steady state due to Tamoxifen toxicity and therefore this was accepted as the working *Kras*^{G12D} resolution.

3.2.2 $Kras^{G12D}$ expression does not alter Atoh1+ cell contribution to the stem cell pool

To investigate stem cell contribution of Atoh1+ and Atoh1+$Kras^{G12D}$ cells during tissue homeostasis, Atoh1;$Kras^{G12D}$;RT and Atoh1;RT animals were administered 3 mg Tamoxifen and analysed 30 days later (Figure 3.4A, Figure 3.5A). Labelled stem cells give rise to differentiated daughter cells that will inherit the label. As such cells proliferate and move up the crypt and villus lineage traced ribbons or clones will appear (Figure 3.4B). Clones that persists with time are stem cell derived. In the small intestine the number of Atoh1+ derived stem cells was relatively low and thus quantification required low magnification microscopy. Therefore, stem cell contribution was quantified by manually counting the total number of tdTom+ crypt/villus ribbons in whole-mounted strips of small intestinal tissue (Figure 3.4B). The total number of tdTom+ ribbons per small intestinal tissue strip was normalised to the length of tissue-strip quantified. Ribbon quantification showed that the amount of Atoh1+ stem cell contribution increased from proximal to distal small intestine but did not change with $Kras^{G12D}$ expression (Figure 3.4C-D).

Epithelial clones expressing tdTom were more common in the colon than the small intestine. Therefore, in the colon stem cell contribution was assessed by quantifying the frequency of tdTom+ crypt clones in approximately 1500 crypts/animal by manual crypt quantification on confocal microscopy images (Figure 3.5B-C). On each tiled image the total number of crypts, the number of crypts wholly populated by tdTom+ cells (WPC), and the number of crypts partially populated by tdTom+ cells (PPC) was quantified (Figure 3.5B, magnified boxes). Single cells were excluded from this analysis as these were considered non-proliferating goblet cells. From this data, the frequency of clones in all crypts was calculated. Quantification of colonic stem cell clone frequency showed no significant difference between Atoh1;$Kras^{G12D}$;RT and Atoh1;RT animals (Figure 3.5C).

Next, examination of stem cell behaviour in the crypt base was carried out. Over time, due to neutral drift, stem cell clones will become fixed or lost. This will be observed as an increase in WPCs and a decrease in PPCs following pulse labelling. If stem cells have a positive bias, they exhibit an increased likelihood generating WPCs. This will present as an accelerated increase in WPCs and decrease in PPCs over time. Such behaviour can be expressed as a ratio of WPC/PPC where an increased ratio suggests accelerated fixation potential of the stem cells analysed. Thus, to examine the stem cell fixation potential of Atoh1+ and Atoh1+Kras^{G12D} derived stem cells WPC/PPC ratios were analysed from the same dataset as was used to obtain clonal frequencies in Figure 3.5B-C. To this end, it was possible to determine stem cell contribution and stem cell behaviour from the same experiment, 30 days post pulse labelling. Interestingly, the WPC/PPC ratio was significantly higher in Atoh1;Kras^{G12D};RT animals compared to Atoh1;RT, suggesting a biased behaviour of Atoh1 derived *Kras*^{G12D} derived stem cells (Figure 3.5D).

As seen from the representative images in Figure 3.5B (magnified boxes) some clones were adjacent (colliding) when analysed. Clonal collisions could arise by chance or by fissioning of one crypt into two. Fissioning events could lead to overestimation of clonal frequencies, as one original clone would be counted as 2 or more. Importantly, *Kras*^{G12D} crypts have a higher fissioning rate than WT crypts which could skew clonal quantification (Snippert *et al.*, 2014).

The data presented in Figure 3.5 counts all clones as individuals and assumes that collisions are due to chance. To check this assumption, it was investigated whether collisions observed in the datasets were due to chance or due to fission. To this end, a mathematical tool that randomly marks clones on a grid of 400 X 400 crypts was created by Richard Kemp (a subset shown in Figure 3.6A). Marking 10% of these crypts randomly (approximately the clonal frequency found in Atoh1;Kras^{G12D};RT and Atoh1;RT at day 30) followed by manual quantification of clonal collisions and patch sizes showed that 26.8($^{+}$/$_{-}$ 6.5)% of all clones would be colliding due to random chance (Figure 3.6A). Next, quantification of clonal collisions in the images used to obtain the data presented in Figure 3.5 was carried out (Figure 3.6B). In this analysis, adjacent partial

clones were also quantified as collisions/patches. This analysis showed that 25.8(+/- 3.2)% and 29.0(+/- 6.9)% of clones were colliding in Atoh1;RT and Atoh1;Kras^{G12D};RT animals, respectively at day 30 (Figure 3.6C). This was not significantly different from that found in the predicted datasets (Figure 3.6C). In addition, analysis of clonal patch sizes showed no significant difference in patch size distribution in the predicted model and experimental mice (Figure 3.6D). This suggest that the clonal collisions recorded 30 days post induction in experimental animals are due to the random nature of crypt marking and not fission, and thus validated the original assumption. In addition, the fission rate does not seem to be increased in Atoh1;Kras^{G12D};RT animals after 30 days.

Taken together these results indicate that Atoh1+ progenitors contribute to the stem cell pool independent of Kras^{G12D} expression but possess a biased behaviour once they have entered the stem cell pool when Kras^{G12D} is present.

3.2.3 Lgr5 depletion increases Atoh1+Kras^{G12D} stem cell contribution in the proximal small intestine but not in the colon

After examining Atoh1+ stem cell contribution during tissue homeostasis it was next investigated in a non-homeostatic setting. Lgr5 has been reported as a marker of rapidly cycling stem cells in the small intestine and colon (Barker *et al.*, 2007). In addition, ablation of Lgr5+ stem cells has been shown to increase the stem cell contribution of both enterocyte and secretory progenitors (Tetteh *et al.*, 2016; Castillo-Azofeifa *et al.*, 2019). Thus, to examine the maximum potential for stem cell contribution of Atoh1+ and Atoh1+Kras^{G12D} cells, Lgr5 stem cell specific deletion in Atoh1;Kras^{G12D};RT and Atoh1;RT animals was desired. To create a model that allowed for this, the Atoh1;Kras^{G12D};RT line was crossed to *Lgr5*$^{Dtr\text{-}GFP}$ animals (Tian *et al.*, 2011) to generate Atoh1;Kras^{G12D};*Lgr5*$^{GFP\text{-}Dtr}$;RT and Atoh1;*Lgr5*$^{GFP\text{-}Dtr}$;RT, referred to as Atoh1;Kras^{G12D};Dtr;RT and Atoh1;Dtr;RT. In this line, Lgr5+ stem cells express the human Diphtheria toxin (DT) receptor (Dtr) and GFP from the *Lgr5* locus (Figure 3.7A). When DT binds to Lgr5+ cells these are ablated (Figure 3.7A). All animals shown in this chapter were heterozygotes for the *Lgr5*$^{GFP\text{-}Dtr}$ allele.

Confirmation of Lgr5+ depletion was carried out by administering *Lgr5*$^{GFP\text{-}Dtr}$ animals with 50 mg/kg DT and assessing GFP expression in the small intestine and colon after 24- and 72 hours by immunohistochemistry (IHC) (Figure 3.7B). Here, GFP served as a marker of Lgr5 expression. As seen from Figure 3.7B, one single dose of DT ablated nearly all Lgr5+ cells after 24 hours in the small intestine and the colon. Lgr5+ cells re-populated the base of the crypts after 72 hours to the same level as the saline control (Figure 3.7B). Importantly, these results were in line with previous reports (Castillo-Azofeifa *et al.*, 2019; Tian *et al.*, 2011; Tomic *et al.*, 2018).

To investigate the effect of Kras^{G12D} expression on Atoh1+ driven contribution after Lgr5+ cell depletion Atoh1;Kras^{G12D};Dtr;RT and Atoh1;Dtr;RT animals were administered 3 mg Tamoxifen and 50 mg/kg DT. Since Tamoxifen is believed to have a slower bioavailability than DT, a 6-hours break between the first Tamoxifen and the subsequent DT injection was given to ensure that Atoh1+ cells were marked at the time of DT ablation. 30 days post induction the small intestine and colon was collected for analysis (Figure 3.8A and 3.10A).

There appeared to be a higher number of tdTom+ cell ribbons present in the proximal small intestine of Atoh1;Kras^{G12D};Dtr;RT animals compared to Atoh1;RT animals when examining the tissue macroscopically (Figure 3.8B). However, the number of ribbons was too great for low magnification quantification. Therefore, clonal frequencies in the proximal small intestine was quantified as described for the colonic data presented in Figure 3.5. The frequency of clones was significantly increased in Atoh1;Kras^{G12D};Dtr;RT animals compared to Atoh1;Dtr;RT animals, with the frequency being approximately 10% and 5%, respectively (Figure 3.8C-D). In addition, examination of the WPC/PPC ratio revealed a significant increase in Atoh1;Kras^{G12D};Dtr;RT animals compared to Atoh1;Dtr;RT (Figure 3.8E). However, clonal collisions also appeared to increase (Figure 3.8C).

To ensure that an increased fission rate in Atoh1;Kras^{G12D};Dtr;RT animals did not falsely increase the clonal frequency and the WPC/PPC ratio, comparison of the expected and experimental clonal-collisions was carried out. The computational tool described in Figure 3.6 was used to predict the amount of random clonal collision and to quantify patch sizes by marking 5% (approximate clonal frequency in Atoh1;Dtr;RT) and 10% (approximate clonal frequency in Atoh1;Kras^{G12D};Dtr;RT) of clones on a 1600 crypt grid (Figure 3.9A+D). This analysis showed that the percentage of clonal collisions and patch sizes observed experimentally was not significantly different to what would be predicted by random clonal marking (Figure 3.9B-C and 3.9E-F). This suggested that the increased clonal frequency found in Atoh1;Kras^{G12D};Dtr;RT was in fact due to increased stem cell contribution of Atoh1+Kras^{G12D} cells in the small intestine after Lgr5 depletion.

Examination of the effect of Lgr5 depletion on Atoh1+ and Atoh1+Kras^{G12D} progenitor behaviour was also carried out in the colon of the same animals (Figure 3.10A). Quantification of clonal frequencies in the colon showed no significant difference between Atoh1;Kras^{G12D};Dtr;RT and Atoh1;Dtr;RT animals, however, as before the WPC/PPC ratio was significantly increased in Atoh;Kras^{G12D};Dtr;RT (Figure 3.10B-D). Interestingly, comparison of the previous homeostatic results (Figure 3.5C) to the results gained after Lgr5 depletion showed no significant difference in clonal frequencies between the two experimental conditions (Figure 3.10E). The same was true for WPC/PPC ratios, Atoh1;Kras^{G12D};Dtr;RT ratios remained significantly higher compared to Atoh1;Dtr;RT, but there was no significant difference in ratios from homeostatic experiments to Lgr5-depletion experiments (Figure 3.10F).

These results suggest that small intestinal Atoh1+Kras^{G12D} cells have an increased contribution potential after Lgr5 depletion, and that colonic Atoh1+ contribution is independent of both Kras^{G12D} expression and Lgr5 ablation.

3.2.4 Atoh1+ derived stem cells have a neutral behaviour whereas Atoh1+ derived *Kras*^{G12D} stem cells have an increased probability of replacement

Stem cells derived from progenitors may have been through partial lineage restriction and it is not known whether such stem cells still carry intrinsic factors that alter their behaviour or tumorigenic abilities. Therefore, to understand whether Atoh1+ derived stem cells act neutrally or if they have an intrinsic bias, mathematical modelling was used to predict stem cell behaviour based on the colonic data from homeostatic experiments (Figure 3.5). Small intestinal data was not included in this analysis as the WPC and PPC populations were not quantified in homeostatic experiments. The modelling presented here was produced by Edward Morrissey and relies on the principle of the neutral drift model previously described by Lopez-Garcia *et al.* 2010 and Vermeulen *et al.* 2013. The neutral drift model is based on *in vivo* pulse-chase experiments where data is collected over several timepoints and clone sizes are quantified in this case as fractions of 8, where 8/8 is a WPC. It utilizes three major factors to predict stem cell behaviour; number of stem cells per crypt (N), stem cell replacement rate/stem cell/day (λ) and probability of replacement (PR). By using data-sets from previous *in vivo* pulse-chase experiments and fixing N=7 (the number of stem cells in the colon estimated by Kozar *et al.* 2013) it was possible to predict how λ and PR would affect the WPC/PPC ratio at day 30 (Figure 3.11A-B).

Firstly, the control data for Atoh1;RT animals was compared to historical control data taking the same assumptions. These were that the number of stem cells per crypt is fixed at N=7 and assuming neutrality in stem cell replacement (by fixing PR to 0.5). Under these conditions this model was used to predict the WPC/PPC ratios after 30 days of pulse-labelling when varying λ (Figure 3.11A). Interestingly, the WPC/PPC found in Atoh1;RT animals predicted that each Atoh1+ derived stem cell had a λ of 0.25 (Figure 3.11A). This is in line or slightly higher with λ's found in previous colonic control pulse chase experiments produced in the Winton lab where λ= 0.22 approximately (data not shown).

The observation that the WPC/PPC ratios was significantly increased in Atoh1;Kras^{G12D};RT suggests that Atoh1 derived *Kras*^{G12D} stem cells have a positive bias. To estimate the bias of Atoh1 derived *Kras*^{G12D} cells the

model in Figure 3.11A was changed to predict WPC/PPC ratios at day 30 by fixing N=7 and λ=0.25 (as found in 11A) and varying the PR (Figure 3.11B). This analysis showed that the WPC/PPC ratio observed in Atoh1;Kras^{G12D};RT animals predicted Atoh1derived *Kras*^{G12D} stem cells to have a PR of 0.65 (Figure 3.9B).

In summary, these result show that Atoh1+ derived stem cells behave neutrally whereas *Kras*^{G12D} stem cells derived from Atoh1+ cells have a competitive advantage.

3.2.5 Atoh1 derived *Kras*^{G12D} crypts have increased fission rate in the proximal small intestine and colon

Next, the nature of inter-crypt colonization over time was examination. The previous experiments have shown that random clonal collisions do create crypt-patches following saturation pulse labelling in the *Atoh1*^{CreErt} model (Figure 3.6). However the size-distribution of these were not different in Atoh1;RT and Atoh1;Kras^{G12D};RT after 30 days (Figure 3.6). Thus, the only route of patch-size growth over time in animals with the same initial clonal induction and patch size distribution is by crypt multiplication. Crypt multiplication happens through the process of crypt-fission. Snippert *et al*. 2014 reported that *Kras*^{G12D} crypts have an increased fission rate compared to WT crypts in the small intestine (Snippert *et al.*, 2014). Thus, to investigate if this behaviour was retained in Atoh1 derived *Kras*^{G12D} crypts, Atoh1;Kras^{G12D};RT and Atoh1;RT animals received 3 mg Tamoxifen on D0 and the small intestine and colon was collected 12- and 24 weeks later for analysis (Figure 3.12A and 3.13A).

Figure 3.12B-E shows patch size analysis in the proximal and distal small intestine (SI-1 and SI-4, respectively) after 12 and 24 weeks. Here, crypt patch sizes were manually quantified under a fluorescent microscope. The 12-week analysis included ~200-350 events and the 24-week analysis included 400-700 events in N=2-3 mice per condition, with more events found in the distal part of the small intestine. Over time, clonal patches grew significantly larger in

Atoh1;Kras^{G12D};RT proximal small intestine (SI-1) compared to Atoh1;RT, with ~75% vs >90% single crypts at 24 weeks, respectively (Figure 3.12C). In addition, Atoh1;Kras^{G12D};RT animals had patch sizes of up to 6 crypts whereas the biggest patch size recorded in Atoh1;RT animals consisted of 3 crypts (Figure 3.12C). Interestingly, no significant difference was observed between Atoh1;Kras^{G12D};RT and Atoh1;RT in the distal small intestine and only few fission events were recorded here (2.5-5% of all events) (Figure 3.12D-E).

Figure 3.13B-C shows patch size distributions in Atoh1;Kras^{G12D};RT and Atoh1;RT colons after 12 and 24 weeks of pulse-labelling. The 12-week analysis included ~550-700 events and the 24-week analysis included 600-1000 events in N=2-3 mice per condition. Quantification of patch sizes showed that Atoh1 derived *Kras*^{G12D} crypts had a significantly increased fission rate compared to Atoh1+ crypts over time (Figure 3.13C).

After 12 weeks, Atoh1;RT animals had 90% single crypts and 10% double crypt patches, whereas Atoh;Kras^{G12D};RT animals contained 85% single crypts and patch sizes of up to 4 crypts (Figure 3.13C). At the 24-week timepoint, Atoh1;RT animals had 85% single crypts and patch sizes of up to 3 crypts. In contrast, Atoh1;Kras^{G12D};RT animals had only 70% single crypt patches at this timepoint, and patches of up to >7 crypts present (Figure 3.13C).

As a note, there were fewer single crypts (patch size = 1) in the 30-day data-sets than in the 12- and 24-week datasets (Figure 3.6 vs. Figure 3.13). This is due to the fact that the 30-day analysis of patch sizes included partial clones as patches (as the mathematical tool did not distinguish full and partial clones during clonal collision prediction). Thus, two colliding partial clones were quantified as a patch of two. At 12- and 24-week post induction, clonal drift was mostly completed hence many partial clones were lost leading to the difference observed.

Together, these results show that Atoh1 derived *Kras*^{G12D} crypts spread via an increased fission rate in the proximal small intestine and colon.

3.2.6 Atoh1;Kras^{G12D};RT animals develop colonic lesions over time

Long-term widespread intestinal activation of Kras^{G12D} has shown to cause hyperproliferative crypts but to not lead to tumour formation (Haigis *et al.*, 2008; Feng *et al.*, 2011). However, surprisingly, when examining the colons of Atoh1;Kras^{G12D};RT at 12 and 24 months post induction, tdTom+ multicrypt-patches that resembled early adenomas when subjected to confocal microscopy were found (Figure 3.14A), something that was never observed in the small intestine or in Atoh1;RT animals (Data not shown). Examination of the total number of tdTom+ lesions 12- and 24 weeks post induction in Atoh1;Kras^{G12D};RT showed the same number of lesions over time (Figure 3.14B). This suggested that the lesions arise early and only progress in size and not number over time.

To further explore the nature of these lesions the colons were subjected to histological examinations. Different types of lesions where observed upon sectioning from 3 different levels in the colonic paraffin blocks (Figure 3.15). Firstly, enlarged crypts with high Ki-67+ hyperproliferation and increased Dusp6 expression were found, here referred to as hyperproliferative crypts (Figure 3.15). Dusp6 is an Erk specific phosphatase that has been reported to be a marker of high MAPK signalling (Buffet *et al.*, 2017) – here it was used as a surrogate marker for *Kras*^{G12D} origin. These crypts seemed to be confined to the basement membrane and did not protrude into the colonic lumen (Figure 3.15). However, crypts that had “umbrella-shaped” protrusions into the lumen were also found (Figure 3.15). These were also recorded as Ki-67^{high} and Dusp6 positive. A few of these lesions seemed to involve several crypts and have a serrated phenotype, these were referred to as “serrated polyp lesion” (Figure 3.15). In all lesions, Ki-67 activity were found in the base of the crypt and never in the distal part (Figure 3.15). None of these lesions found were positive for stabilised nuclear ß-catenin (Figure 3.15). In addition, most of these were found in the upper colon.

In conclusion, these results show that Kras^{G12D} expression in Atoh1+ cells can induce hyperproliferative crypts and serrated colonic polyps over time.

3.3 Discussion

Taken together the results in this chapter show that Kras^{G12D} expression does not change cell fate choices of Atoh1+ cells during tissue homeostasis but might increase the small intestinal Atoh1+ stem cell potential after Lgr5 depletion. In addition, *Kras*^{G12D} mutations that arise in Atoh1+ progenitors can produce *Kras*^{G12D} stem cells with a competitive bias and *Kras*^{G12D} crypts that spread via fission in the proximal small intestine and colon and develop colonic polyps over time.

It has previously been reported that Kras^{G12D} expression and MAPK signalling affects lineage choices in the intestine towards the secretory goblet cell lineage (Haigis *et al.*, 2008; Heuberger *et al.*, 2014; Gagné-Sansfaçon *et al.*, 2016). This could lead to the hypothesis that Kras^{G12D} expression in secretory progenitors could induce a terminal goblet cell differentiation pattern. In theory, this would decrease the stem cell contribution potential of these cells. In contrast, a study that examined gene expression patterns in *KRAS* mutated tumours showed an increased embryonic stem cell like expression pattern in such tumours (Le Rolle *et al.*, 2016). This suggests that *KRAS* mutations infer de-differentiation of intestinal cells which leads to the hypothesis that *Kras*^{G12D} mutations would give progenitor cells an increased stem cell potential.

However, results in this chapter suggest that early Atoh1 expressing progenitors contribute to the stem cell pool independent of Kras^{G12D} expression during homeostasis. Importantly, even though the potential contribution has not changed with Kras^{G12D} expression, stem cell clones originating from Atoh1+Kras^{G12D} cells are observed. This underlines that *Kras*^{G12D} mutations are not excluded from the stem cell pool if they arise in Atoh1 progenitors. This greatly expands the pool of cells that can give rise to mutated *KRAS* colonic fields.

In addition, these findings propose that Kras^{G12D} expression might be important for cell fate decisions in the stem cell pool but is dispensable in cells that have started to express secretory commitment transcription factors such as Atoh1. In the future, to further investigate the effect of Kras^{G12D} signalling on

Atoh1+ stem contribution it will be interesting to administer Mek inhibitor to the Atoh1;RT animals. In addition, it would be interesting to explore Atoh1+ stem cell contribution in $Kras^{G12D}$ activated tissue to determine if secretory progenitors have an increased/decreased role in monitoring the epithelium in *Kras*^{G12D} pro-oncogenic fields.

Interestingly, the results in this chapter suggest that Lgr5 deletion in the small intestine gives Atoh1+$Kras^{G12D}$ cells a higher chance of becoming stem cells compared to Atoh1+ cells. Thus, $Kras^{G12D}$ expression could expand the de-differentiation potential of Atoh1+ cells in response to injury. This could be crucial for patients undergoing chemotherapy or radiotherapy. Such treatments are known to kill rapidly-cycling intestinal stem cells (Potten, 1977; Keefe *et al.*, 2000). Thus, because *Kras*^{G12D} secretory progenitors appear to have an increased regenerative potential during such responses patients could have a higher chance of gaining *KRAS* mutations in intestinal stem cells in response to treatment. This would potentially increase their chance of developing secondary cancer over time.

In contrast, no significant increase in clone frequencies after colonic Lgr5 deletion was observed. This suggest that Atoh1+ stem cell contribution is independent of Lgr5 depletion and $Kras^{G12D}$ expression in the colon. Interestingly, this is in contrast to a report by Castillo-Azofeifa *et al.* 2019. Here, the authors show a 15-fold increase in Atoh1+ traced colonic clones after Lgr5 ablation. In the paper the authors use a Lgr5 depletion strategy in which DT is administered on two consecutive days prior to Atoh1 tracing in addition to 1 day after tracing starts. Taken together, this could mean that the Lgr5 regenerative response is slower in the colon than in the small intestine and that our Atoh1+ tracing do not catch the window of Lgr5 induced regeneration with our same-day DT administration strategy. On the other hand, it is also a possibility that the DT strategy employed by Catillo-Azofeifa *et al.* induces an increase in Atoh1+ cell numbers. This would mean that they in fact sample more Atoh1+ cells which implies that the stem cell contribution potential of each Atoh1+ cell is not changed in response to Lgr5 ablation. Thus, in the future it will be important to test the Lgr5 depletion strategy

applied by Catillo-Azofeifa *et al.* in our Atoh1;Kras^{G12D};Dtr;RT model. This will emphasise whether stem cell contribution is independent on *Kras*^{G12D} during Lgr5 depletion in the colon. Further, it would be interesting to test other injury strategies to fully determine the regenerative response of Atoh1+Kras^{G12D} cells.

After assessing the stem cell contribution potential of Atoh1+ and Atoh1+Kras^{G12D} cells it was next examined whether stem cells derived from secretory progenitors are impaired in their stem cell function. Here, we are the first to show that Atoh1+ derived colonic stem cells have a neutral behaviour with an λ that is similar to that found in other neutral Cre models. In addition, the results show that Atoh1+ derived *Kras*^{G12D} stem cells have a competitive advantage over WT stem cells with a PR of 0.65 vs 0.5 respectively. Vermeulen *et al.* 2013 has previously reported that small intestinal *Kras*^{G12D} stem cells have a PR of 0.78. The difference between the PR recorded in the two experiments could be due to small intestine and colon discrepancies or lacking co-recombination in the model presented here.

However, taken together this suggest that stem cells derived from Atoh1+ progenitors behave like any other stem cell in the crypt base and do not carry any intrinsic factors that alter their behaviour.

Next, the nature of long term Atoh1+ and Atoh1+Kras^{G12D} lineage tracing was investigated, to expose the inter-crypt spread and the tumorigenic potential of such crypts. The results show that Atoh1 derived *Kras*^{G12D} crypts have an increased fission rate in the proximal small intestine and colon compared to Atoh1 derived crypts. This is in line with data from Snippert *et al*. 2014 also showing increased fission in the small intestine following *Kras*^{G12D} lineage tracing. This indicates that Kras^{G12D} expression accelerates fission independent of cell origin. However, Snippert *et al*. reports 60% single crypts after 16 weeks of lineage tracing in the small intestine of *Kras*^{G12D} animals in contrast to 75% after 24 weeks reported here (proximal small intestine). The higher fission frequency reported by Snippert *et al*. could be due to increased clonal collision in their

Lgr5^{CreErt};*Kras*^{lslG12D};Confetti model or it could be due to differential co-recombination of *Kras*^{G12D} and reporter alleles in the two models.

Interestingly, here the results show a differential fission effect of *Kras*^{G12D} crypts in the proximal and distal small intestine, with no difference in fission events between WT and *Kras*^{G12D} crypts in the distal small intestine. This could be due to a differential co-recombination rate in the proximal and distal small intestine.

Importantly, the results report that *Kras*^{G12D} crypts also spread via increased fission in the colon *in vivo*. This appears to occur at the same rate as in the proximal small intestine. This verifies that *Kras*^{G12D} crypt spread also occurs in the colon which supports the modelling of increased mutated *KRAS* crypt fission rate found in human colons (Nicholson *et al.*, 2018).

Lastly, long term exposure of Atoh1 derived *Kras*^{G12D} lead to production of early lesions in the colonic epithelium. The microscopic features of these lesions leads to the hypothesis that they are different stages of the more severe serrated-lesions. These could start as hyperproliferative crypts that develops into umbrella shaped polyp-crypts and serrated lesions over time. The hyper-plasticity resulting from Kras^{G12D} colonic expression has previously been reported and is also observed in other models with more widespread Kras^{G12D} activation (Haigis *et al.*, 2008; Feng *et al.*, 2011) (and the Winton lab). However, polyp lesions have not been reported upon colonic Kras^{G12D} activation alone. The lesions reported here could arise due to the secretory Atoh1+ origin, indicating that Atoh1+ derived *Kras*^{G12D} stem cells are in fact different to *Kras*^{G12D} stem cells of other origins. These differences do not alter their stem cell behaviour, but it is possible that Atoh1 derived *Kras*^{G12D} crypts and patches select for clones with other amplifying genetic mutations that could cause the lesions observed. However, the sporadic nature of the Kras^{G12D} activation in this model could also be the source of initiation. If so, that indicates that *Kras*^{G12D} lesions might require extrinsic factors from surrounding WT crypts to progress into serrated lesions and polyps. A good control for this would be sporadic colonic Kras^{G12D} activation by a stem cell driven Cre.

The serrated neoplasia pathway is a new classification of colon cancer that is thought to comprise 20% of sporadic cases (Rashtak *et al.*, 2017). The precursor lesion for this type of colon cancer is sessile serrated polyps. However, it has been proposed that hyperplastic polyps might in fact be an even earlier stage of the disease (Jass, 2004). It is believed that the initiating event for such lesions are *BRAF*^{V600E} mutations though a recent study found that 20% of serrated polyps contained *KRAS* mutations (Juárez *et al.*, 2017). If the lesions found in Atoh1+Kras^{G12D} animals are in fact hyperplastic and serrated, they could be initial stages of serrated colon-cancer. It is, however, clear that these lesions are not part of the classical colorectal cancer progression pathway as they lack nuclear ß-catenin staining. In the future, it will be interesting to age and characterise these lesions further.

In conclusion, early secretory progenitors could in fact be the origin of *Kras*^{G12D} mutations in stem cells and may be one of contributors generating colorectal *KRAS* fields. This finding means that the pool of cells that can give rise to *KRAS* fields is larger than previously expected.

In the future, it will be interesting to investigate how Kras^{G12D} activation affects precursors of different cell lineages such as Alpi+ enterocytes and Ngn3+ enteroendocrine progenitors. This would map the relative influence of progenitor stem cell contribution on *Kras*^{G12D} field origin

4 Defining stem cell dynamics in intestinal *Kras*^{G12D} fields

4.1 Introduction

Cancer development is believed to be a step-wise process of mutational acquisition, selection and clonal expansion (Baker *et al.*, 2018). Thus, pre-neoplastic cells clonally expand and acquire the molecular identities essential for tumour development over time. In this interpretation, some pre-neoplastic cells carry oncogenic mutations but are not yet morphological distinct (Baker *et al.*, 2018). Consequently, pre-neoplastic clones can be in a "pre-disposed cancerous" state, which is referred to as field characterisation or field effect. The notion that oncogenic mutations can be discovered in cells before overt tumour formation and that mutational burden increases with age supports the idea that field effect is an unavoidable trait of self-renewing tissues (Tomasetti *et al.*, 2013; Alexandrov *et al.*, 2015). Additionally, field effects have been discovered in many different tissue types, such as the skin (Martincorena *et al.*, 2015), the oesophagus (Martincorena *et al.*, 2018), haematopoietic stem cells (Xie *et al.*, 2014) and the colonic epithelium (Galandiuk *et al.*, 2012; Nicholson *et al.*, 2018).

In 1990 the Vogelstein model predicted that colorectal cancer develops through sequential mutational acquisition in the colonic epithelium with the initiating event being loss of *APC* (Fearon and Vogelstein, 1990). In this model, oncogenic *KRAS* and *TP53* mutations occur after early adenomas have formed (Fearon and Vogelstein, 1990). However, recent studies have shown that oncogenic mutations in *KRAS* can be found in 10% of morphological healthy human colon tissue and can exist in fields of up to 100 crypts (Dieterle *et al.*, 2004; Kraus *et al.*, 2006; Parsons *et al.*, 2010; Nicholson *et al.*, 2018). This leads to the hypothesis that *KRAS* mutated fields arise early in colorectal tissue and might pre-dispose to cancer development.

As intestinal stem cells are self-renewing, they are the only cells that live long enough to maintain tumorigenic mutations in the epithelium. Therefore, a field effect might begin with the process of acquiring mutation(s) in a single stem cell which through neutral or biased drift will create monoclonally converted mutated crypts, so-called clonal fixation (Lopez-Garcia *et al.*, 2010; Snippert *et al.*, 2010). Such mutated crypts can spread by fission to create mutated fields within the intestinal epithelial. In the previous chapter and by Vermeulen *et al.* 2013, it was shown that intestinal stem cells harbouring *Kras*^{G12D} mutations have a competitive bias over WT stem cells. This gives *Kras*^{G12D} stem cells an increased chance of outcompeting WT cells and creating monoclonal *Kras*^{G12D} crypts during clonal drift and thus fixing in the tissue. Furthermore, data from the previous chapter and Snippert *et al.* 2014 has shown that *Kras*^{G12D} crypts have an increased fission rate. Together, this gives *Kras*^{G12D} mutations the potential for creating pre-neoplastic fields.

Studies exploring the long-term effects of *Kras* mutations in the murine intestinal epithelium have found that oncogenic *Kras* alone does not initiate cancer (Haigis *et al.*, 2008; Feng *et al.*, 2011). Thus, it is not clear if mutated *KRAS* fields are directly oncogenic or whether they indeed pre-dispose to cancer development. Therefore, it is important to characterise the cellular behaviour and clonal dynamics of such fields to understand their tumorigenic potential. Such results will aid the general understanding of early colorectal cancer development and might help identify new ways to treat or prevent the disease.

To define clonal dynamics in *Kras*^{G12D} fields, the *Cyp1A1*^{FlpO};*Kras*^{fsfG12D};*Ah*^{CreErt};*R26*^{lsltdTom} (referred to as Flp;Kras^{G12D};Cre^{Ert};RT) mouse model, that allows temporal separation of the creation of a *Kras*^{G12D} field from pulse-chase lineage tracing, was created.

The cytochrome P450 family member 1, subfamily a, polypeptide 1 (*Cyp1A1*) *Cyp1A1*^{FlpO} allele was developed in the Winton lab and has not been published yet. Here, the FlpO expression cassette was targeted to the *Cyp1A1* mouse promoter and enhancer element located on chromosome 9. Cyp1A1 is a

xenobiotic metabolising enzyme that is active in the intestinal epithelium, liver and in the immune system (Ireland *et al.*, 2004; Gutié Rrez-Vá Zquez and Quintana, 2018). The *Cyp1A1* promoter is silent until exposed to lipophilic xenobiotics (such as ß-naphthoflavone) which can bind to the cytoplasmic aryl hydrocarbon receptor (Ah). The Ah receptor will in turn translocate to the nucleus and bind to the xenobiotic response element in the *Cyp1A1* promoter to initiate transcription (Matsushitasq *et al.*, 1993). The Ah^{CreErt} allele was developed by placing the Cre^{Ert} expression cassette under the transcriptional control of the rat *Cyp1A1* promoter and delivering this transgene by oocyte pro-nuclear injection (Kemp *et al.*, 2004). To this end, the Ah^{CreErt} allele is created by random genetic insertion and the $Cyp1A1^{FlpO}$ by targeted insertion, but otherwise FlpO and Cre^{Ert} are transcriptionally regulated by the same inducing agent.

A single dose of ß-naphthoflavone will initiate a pulse of Flp and Cre^{Ert} expression from the *Cyp1A1* and Ah promoter, respectively, in the in intestinal epithelium of Flp;$Kras^{G12D}$;Cre^{Ert};RT animals. Since Cre^{Ert} encodes the inducible oestrogen receptor (Ert) this recombinase is not active until Tamoxifen is administered. However, Flp recombinase is directly functional upon expression. Thus, ß-naphthoflavone administration alone will lead to widespread intestinal $Kras^{G12D}$ expression by Flp mediated removal of the frt flanked STOP cassette from the $Kras^{fsfG12D}$ allele (Figure 4.1A). Subsequently, a dose of ß-naphthoflavone and Tamoxifen will activate Cre^{Ert} to remove the loxP flanked STOP cassette on the $R26^{ls/tdTom}$ allele to initiate tdTom lineage tracing in $Kras^{G12D}$ fields (Figure 4.1B). Therefore, combining $Cyp1A1^{FlpO}$ and Ah^{CreErt} in the same model theoretically makes it possible to separate two recombination events temporally and thus facilitates lineaging tracing in $Kras^{G12D}$ intestinal fields. All animals presented in this chapter were heterozygotes for all alleles.

The aim of this chapter is to characterise the nature of stem cell replacement and clonal drift within murine $Kras^{G12D}$ mutant crypts in $Kras^{G12D}$ fields, to expose any pro-oncogenic features of stem cell behaviour. The second aim is to explore the pathways that govern clonal growth in $Kras^{G12D}$ fields.

4.2 Results

4.2.1 Optimisation of the Flp;Kras^{G12D};Cre^{Ert};RT mouse

To be able to utilize *Cyp1A1*^{FlpO} and Ah^{CreErt} alleles temporally it was important that the initial induction of FlpO did not activate the second recombinase in the sequence: Ah^{CreErt}. *Cyp1A1*^{FlpO} is activated by ß-naphthoflavone and Ah^{CreErt}, in theory, requires both ß-naphthoflavone and Tamoxifen to be activated. However, if in fact the induction of *Cyp1A1*^{FlpO} creates background Ah^{CreErt} activity it could skew lineage tracing as some tdTom+ clones would originate early from background activity and not from the intended pulse chase.

Therefore, it was firstly tested whether ß-naphthoflavone administration alone would lead to background activity of Cre recombinase. This was examined by quantifying the number of tdTom+ small intestinal villus ribbons and colonic crypts present 30 days post one dose of either 40, 20, 8 or 4 mg/kg ß-naphthoflavone in Ah^{CreErt};*R26*^{lsltdTom} (referred to as, Ah^{CreErt};RT) animals (Figure 4.2A). In this experiment, tdTom+ expression served as a surrogate marker for Ah^{CreErt} activity. A positive control animal was administered 40 mg/kg ß-naphthoflavone as well as 4 mg Tamoxifen to induce high levels of Ah^{CreErt} activity (Figure 4.2B). A negative control animal did not receive any drugs (Figure 4.2B). The number of tdTom+ ribbons and crypts were counted manually in whole-mounted small intestine and colon under a fluorescent microscope and normalised to the length of the tissue strip scored (figure 4.2C). Importantly, the positive control contained a vast amount of tdTom+ crypts in the proximal small intestine, demonstrating Cre activity in these animals, whereas no tdTom+ crypts were observed in negative controls (Figure 4.2B). Quantification of Ah^{CreErt} background activity showed greater tdTom+ levels in the proximal small intestine (SI-1 and SI-2) than the distal small intestine (SI-3 and SI-4) and colon (Figure 4.2C). In addition, background activity behaved in a dose-responsive manner, with the greatest number of ribbons (>60 ribbons / cm in SI-1) found in animals administered 40 mg/kg ß-naphthoflavone (Figure 4.2B-C). Importantly, the background activity was negligible in animals treated with 8 and 4 mg/kg ß-naphthoflavone with only 0-2 ribbons / cm found in the proximal small intestine and none in the colon (Figure 4.2B-C).

Therefore, 8 mg/kg ß-naphthoflavone was chosen as the *Cyp1A1*^{FlpO} induction dose for any future experiments. Clonal quantification in Ah^{CreErt};RT lineage tracing experiments normally requires approximately 2 cm of small intestinal tissue to quantify 100-300 clones. Thus, the approximate number of background clones scored during lineage tracing in Flp;Kras^{G12D};Cre^{Ert};RT would be a maximum of 4/100 clones (4%) which was viewed as an acceptable error margin.

It was important that the initial *Kras*^{fsf}^{G12D} recombination was highly efficient in creating tissue wide fields. Therefore, to investigate this Flp;Kras^{G12D};Cre;RT animals were administered 8 mg/kg ß-naphthoflavone and 21 days post induction the intestinal epithelium was isolated from the proximal small intestine and colon. gDNA from these samples was subjected to PCR amplification to examine the level of *Kras*^{fsfG12D} recombination (Figure 4.3A). Here, uninduced Flp;Kras^{G12D};Cre^{Ert};RT animals served as a negative control for background *Kras*^{fsfG12D} recombination (Figure 4.3A). In addition, wild-type (WT) gDNA (from ear-biopsy) served as a negative control for PCR specificity (Figure 4.3A). Examination of PCR amplified DNA by gel-electrophoresis showed high levels of *Kras*^{fsfG12D} recombination in both small intestine and colon as the unrecombined band (FSF) was extremely diminished in samples exposed to ß-naphthoflavone (Figure 4.3A). In addition, it appeared that recombination efficiency was higher in the colon than in the small intestine (Figure 4.3A). Importantly, recombination was absent in un-induced samples underlining that recombination was dependent on *Cyp1A1*^{FlpO} induction (Figure 4.3A).

Conclusively, this result showed that 8 mg/kg ß-naphthoflavone was sufficient to induce high levels of *Kras*^{fsfG12D} recombination, whilst inducing negligible levels of tdTom+ cell marking.

The standard dose for Ah^{CreErt} mediated tdTom+ pulse lineage chasing is 40 mg/kg ß-naphthoflavone + 0.15 mg Tamoxifen. This dose marks <1 cell per crypt on average which is desired in lineage trace experiments. To test whether this dose was still applicable for lineage tracing in the Flp;Kras^{G12D};Cre^{Ert};RT model, animals were administered 8 mg/kg ß-naphthoflavone to induce *Cyp1A1*^{FlpO}

activity and 21 days later a dose of 40 mg/kg ß-naphthoflavone + 0.15 or 0.3 mg Tamoxifen to initiate Ah^{CreErt} mediated lineage tracing (Figure 4.3B). Confocal microscopy of the proximal small intestine (highest Ah^{CreErt} activity) 4 days post lineage trace initiation showed that 40 mg/kg ß-naphthoflavone + 0.15 mg Tamoxifen lead to <1 tdTom+ cells / crypt whereas 0.3 mg Tamoxifen lead to >1 tdTom+ cells / crypt (Figure 4.3B). Thus, 40 mg/kg ß-naphthoflavone + 0.15 mg Tamoxifen was the chosen dose for lineage tracing in the $Flp;Kras^{G12D};Cre^{Ert};RT$ model.

However, induced $Flp;Kras^{G12D};Cre^{Ert};RT$ animals developed oral papillomas over time (Figure 4.3C). This was never observed in WT animals and thus was $Kras^{G12D}$ specific. Because of mouth obstruction this allowed $Flp;Kras^{G12D};Cre^{Ert};RT$ animals to be kept alive for approximately 50 days post $Kras^{G12D}$ induction (Figure 4.3C). Lineage tracing is a 21-day experiment and therefore it was concluded that initial $Kras^{G12D}$ recombination should last for 21 days (Figure 4.3D).

In conclusion, the result show that $Cyp1A1^{FlpO}$ and Ah^{CreErt} can be utilized in the same model to temporally separate $Kras^{fsfG12D}$ recombination from lineage tracing. In addition, it appeared that 8 mg/kg ß-naphthoflavone was the optimal dose to use for initial $Cyp1A1^{FlpO}$ activation as this dose created negligible levels of Ah^{CreErt} background activity and high levels of $Kras^{fsfG12D}$ recombination.

4.2.2 Clonal dynamics are significantly accelerated in $Kras^{G12D}$ fields

After thorough optimisation and characterisation of the $Flp;Kras^{G12D};Cre^{Ert};RT$ model lineage tracing was carried out in $Kras^{G12D}$ fields to define clonal dynamics.

$Flp;Kras^{G12D};Cre^{Ert};RT$ ($Kras^{G12D}$) or $Flp;WT;Cre^{Ert};RT$ (WT) animals received 8 mg/kg ß-naphthoflavone to induce $Kras^{G12D}$ expression or not and 21 days later they received 40 mg/kg ß-naphthoflavone + 0.15 mg Tamoxifen to initiate lineage tracing of tdTom+ clones (Figure 4.4A). Animals were sacrificed 4, 7, 10, 14 and 21 days post pulse-chase initiation (Figure 4.4A). The proximal

small intestine and colon were then subjected to manual clone quantification by fluorescent microscopy to score the relative sizes of tdTom+ clones in intestinal crypts recorded as fractions of 8 (Figure 4.4B). Each timepoint contained 3-4 animals. In WT animals a total of 450-750 clones were quantified per timepoint in the small intestine and colon. In *Kras*^{G12D} animals 450-900 and 150-650 clones were quantified in the small intestine and colon respectively, with fewest clones quantified at later timepoints. As a note, the clonal frequency was lower in *Kras*^{G12D} colons leading to the relative low clone count here.

Analysis of average clone size showed that clones in *Kras*^{G12D} small intestine and colon were significantly larger than clones in WT animals at all timepoints (Figure 4.4C). This implies that the rate by which stem cells replace each other during clonal expansion (replacement rate) is increased in *Kras*^{G12D} fields. To investigate clonal fixation time in these animals, the percentage of wholly populated clones (WPC) (scored as 1, Figure 4.4B) over time was plotted (Figure 4.4D). From this analysis, it was clear that clones in *Kras*^{G12D} epithelium have a significantly decreased monoclonal conversion time, with nearly 60% of all clones being scored as WPC at day 21 in both the small intestine and colon, compared to <20% in WT animals (Figure 4.4D). In addition, there appeared to be a lag in the time it took clones to become WPC in *Kras*^{G12D} colons compared to the small intestine (Figure 4.4D). Heatmap analysis of clone size distribution further confirmed that overall, clones grew larger at an accelerated rate in *Kras*^{G12D} small intestine and colon compared to WT (Figure 4.4E).

In conclusion, the lineage tracing data here showed that dynamics of clonal growth appear to be accelerated in *Kras*^{G12D} fields with significantly larger average clone sizes and percentage of WPCs over time compared to that of WT crypts.

4.2.3 Determining stem cell number and stem replacement rate in *Kras*^{G12D} fields

To be able to quantitively compare stem cell dynamics in the *Kras*^{G12D} and WT fields mathematical modelling was used to predict stem cell number per crypt (N) and stem cell replacement rate per stem cell per day (λ). The mathematical model was designed by Edward Morrissey and has previously been applied to describe stem cell behaviour in the mouse intestine (Kozar *et al.*, 2013; Vermeulen *et al.*, 2013). The model assumes neutral behaviour between stem cells in intestinal crypts and it relies on Bayesian inference to explain stochastic stem cell behaviour to predict average clone sizes over 100 days in *Kras*^{G12D} and WT small intestinal and colonic fields (Figure 4.5A).

From the model it is also clear that stem cell dynamics are accelerated in *Kras*^{G12D} fields. Thus, the model predicts that all crypts in *Kras*^{G12D} fields are monoclonal after approximately 60 days of neutral drift whereas 100 days of drift is insufficient for all crypts to become monoclonal in WT fields (Figure 4.5A).

From the average clone sizes the model inferred N to be 6 and 7 and λ to be 0.17 and 0.21 in the WT small intestine and colon respectively (Figure 4.5B). In contrast, N was inferred to be 5 in both the small intestine and colon of *Kras*^{G12D} animals and λ was estimated to 0.31 and 0.26 in the *Kras*^{G12D} small intestine and colon respectively (Figure 4.5B). As a note, the model suggested to a low probability that N=4 and λ=0.18 in the *Kras*^{G12D} colon. This was probably due to the higher variability in the colonic dataset (Figure 4.5B).

Thus, this analysis shows that *Kras*^{G12D} crypts mediate accelerated clonal growth by employing fewer stem cells with a relative increased replacement rate for clonal drift.

4.2.4 Crypts in *Kras*^{G12D} fields have an increased mutational burden

Theoretically, an accelerated monoclonal conversion rate will lead to faster fixation of neutral and oncogenic mutations. Thus, if the rate of mutations per cell division is the same in WT and *Kras*^{G12D} crypts, *Kras*^{G12D} crypts should fix an increased amount of mutations compared to WT crypts during aging. To test this experimentally, exome sequencing was employed to examine mutational burden in WT and *Kras*^{G12D} crypts.

Previous exome sequencing data has shown that the single nucleotide variant (SNV) mutation rate is 0.37 X 10^{-10} per cell division per base pair in normal crypts (Lugli *et al.*, 2017). Therefore, it was expected that the relative acquisition of SNVs over time would be low. Thus, to make it statistically possible to test if there was a difference between *Kras*^{G12D} and WT crypts it was concluded that aged crypts that had time to acquire some mutational burden would be appropriate for this analysis. As Flp;Kras^{G12D};Cre^{Ert};RT animals could not be maintained long term *Villin*^{CreErt};*Kras*^{lslG12D} and *Villin*^{CreErt};WT animals were used for these experiments. In this model *Villin*^{CreErt} allows for conditional intestinal specific expression of Kras^{G12D} by administration of Tamoxifen. Additionally, this analysis focused on colonic crypts to be as close to the human colorectal pathology as possible.

For widespread intestinal Kras^{G12D} expression (or not) 61 and 126 days old *Villin*^{CreErt};*Kras*^{lslG12D} (*Kras*^{G12D}) and *Villin*^{CreErt};WT (WT) animals received 3 mg Tamoxifen (Figure 4.6A). Single crypts were isolated from the colon 186 days post induction (Figure 4.6A). To expand the genetic material for sequencing, crypts were cultured as organoids for 5 days (Figure 4.6A). Thereafter, genomic DNA was isolated from 6 single organoids per animal and these were subjected to exome capture and next generation sequencing (Figure 4.6A).

Mutation calling was performed by Lee Hazelwood. To filter out germline single nucleotide polymorphisms (SNPs) and call somatic SNVs each sample was compared against a sample background. Background samples consisted of a composite sample in which sequencing data from organoid groups of same genotype and age was collated. Organoid SNVs were called by comparing them back to the background sample minus the organoid under investigation. Only monoclonal somatic SNVs were called in this analysis. Such

were expected to be largely heterozygous with an allele frequency of 0.5 therefore the allele frequency window used to call SNVs were set to 0.35-0.65 at a read-depth of 30 reads per base.

This analysis showed that colonic *Kras*^{G12D} crypts acquired significantly more somatic mutations over time compared to WT crypts (Figure 4.6D).

4.2.5 *Kras*^{G12D} fields accelerate ß-catenin mediated tumorigenesis

As described above the clonal dynamics found in *Kras*^{G12D} fields imply that such areas could speed up cancer initiation time by decreasing monoclonal conversion time of additional oncogenic mutations. The increased fission rate observed in *Kras*^{G12D} crypts (explained in chapter 3) could give new mutated crypts in a *Kras*^{G12D} field the potential to spread and thereby create a new field within the field. Over time these could develop into intestinal adenomas and carcinomas in some instances by acquiring additional oncogenic mutations. In addition, it is clear that increased Wnt signalling is an initiating factor of colorectal cancer, underlined by the fact that 80% of human colon tumours contain mutations in *APC* (The Cancer Genome Atlas Network, 2012). Furthermore, it has been shown that conditional deletion of exon 15 and 3 in *Apc* and *Ctnnb1,* respectively, can cause nuclear ß-catenin stabilisation and adenoma development in the mouse intestine (Harada *et al.*, 1999; Robanus-Maandag *et al.*, 2010).

To functionally examine whether *Kras*^{G12D} fields accelerate cancer initiation caused by stabilised nuclear ß-catenin the *Villin*^{Flp};*Kras*^{fsfG12D}; Ah^{CreErt};*Ctnnb1*^{floxex3} (denoted as Kras^{G12D};ß-cat) mouse was created. The Villin promoter is gut specific and is active during development and in adult mice (Maunoury *et al.*, 1992). Thus, in this model, animals were born with intestinal *Kras*^{G12D} fields and sporadic stabilisation of nuclear ß-catenin was created by deleting exon 3 in *Ctnnb1* by Cre activation. The control for these mice were *Villin*^{Flp};*WT*;Ah^{CreErt};*Ctnnb1*^{floxex3} (denoted as WT;ß-cat). Due to time constraints and the difficulty of this breeding only two Kras^{G12D};ß-cat animals were included in the analysis described below.

To induce clonal levels of stabilised nuclear ß-catenin, 8-weeks old Kras^{G12D};ß-cat and WT;ß-cat animals received 40 mg/kg ß-naphthoflavone and 0.5 mg Tamoxifen and were culled when moribund (Figure 4.7A). The clonal induction of nuclear ß-catenin was chosen so that the experiment would be as close to the human equivalent as possible and to allow the animals to survive long enough for the experiment to reveal true differences between Kras^{G12D};ß-cat and WT;ß-cat.

The results revealed a noticeable difference in survival of the two groups of animals (Figure 4.7B). Kras^{G12D};ß-cat animals survived 49 and 73 days whereas WT;ß-cat animals survived to a median of 161 days (N=3 animals) (Figure 4.7B). To investigate tumour progression side by side, matched controls samples were collected upon harvest of Kras^{G12D};ß-cat animals. Upon tissue inspection in such matched samples it was clear that Kras^{G12D};ß-cat had a more severe phenotype than WT;ß-cat animals (Figure 4.7C). This manifested as 49 and >100 tumours in the proximal small intestine (SI-1) of Kras^{G12D};ß-cat animals compared to just 9 and 4 found in WT;ß-cat animals (Figure 4.7C-D).

Together these preliminary results suggest that *Kras*^{G12D} fields in fact do speed up cancer initiated by stabilisation of nuclear ß-catenin. This further supports the hypothesis that *Kras*^{G12D} fields are pro-oncogenic areas.

4.2.6 Mek1/2 inhibition can rescue accelerated clonal dynamics in *Kras*^{G12D} fields

The Ras family proteins are GTPases that are involved in many different cell-signalling cascades. One of the main substrates of the activated GTP bound Ras proteins are the Raf family of protein kinases (Pylayeva-Gupta *et al.,* 2011). Upon Ras-Raf interaction Raf proteins will be left in an active phosphorylated state. The main substrates of Raf mediated phosphorylation are the protein kinases Mek1/2 (Pylayeva-Gupta *et al.,* 2011). Phosphorylated Mek1/2 will in turn phosphorylate Erk1/2 which can phosphorylate a variety of target proteins to induce cellular responses such as: proliferation, protein synthesis and survival (Pylayeva-Gupta

et al., 2011; Cox *et al.*, 2014). *RAS* genes constitute the most frequently mutated oncogenes in human cancer and thus many efforts have been made to create inhibitors for the Ras proteins and their downstream signalling (Cox *et al.*, 2014). The compounds that have had success in the clinic are inhibitors of the downstream effectors Mek1/2 and Erk1/2 (Cox *et al.*, 2014). AZD6244 is a selective Mek1/2 inhibitor that has been shown to be well tolerated by patients in phase I clinical trials (Adjei *et al.*, 2008).

Based on the foregoing, it was next investigated if the accelerated pro-oncogenic clonal dynamics observed in *Kras*^{G12D} fields could be inhibited by the administration of the clinically available Mek1/2 inhibitor AZD6244 (Meki). To test the effect on clonal growth, Meki was administered during lineage tracing in *Kras*^{G12D} fields. Experimentally, the Flp;Kras^{G12D};Cre^{Ert};RT (*Kras*^{G12D}) and Flp;WT;Cre^{Ert};RT (WT) were administered 8 mg/kg ß-naphthoflavone to induce Kras^{G12D} expression (or not) and 21 days later lineage tracing was begun (Figure 4.8A). To obtain Mek1/2 inhibited *Kras*^{G12D} fields at the beginning of lineage tracing, administration of Meki (25 mg/kg) or vehicle treatment was begun one day prior to lineage trace initiation (Figure 4.8A). Meki or vehicle was then administered twice daily by oral gavage and animals were culled 14 days post lineage trace initiation (Figure 4.8A).

Initially it was clear that Meki treatment regressed oral papilloma growth in *Kras*^{G12D} animals, which confirmed that the drug appeared to target Kras^{G12D} affected cells (Figure 4.8A). To ensure that the treatment had reached the small intestine and colon the levels of phosphorylated Mek1/2 was assessed by IHC in animals treated with vehicle or Meki. This confirmed a clear reduction in active Mek1/2 protein levels post treatment in both WT and *Kras*^{G12D} animals (Figure 8B). Next, the size of lineage traced clones and percentages of WPCs were quantified as described for Figure 4.4. The data obtained from this experiment was then compared back to data obtained at day 14 post lineage tracing in Figure 4.4 (Figure 4.8C-D). This analysis revealed that Meki treatment significantly reduced the average clone size and percentage of WPCs in small intestinal and colonic *Kras*^{G12D} fields (Figure 4.8C-D). However, the average clone size and percentage of fixed clones was still significantly greater in Meki treated *Kras*^{G12D} fields compared to WT fields (Figure 4.8C-D). In addition, this

effect was Meki specific as there was no significant difference in *Kras*^{G12D} samples with or without vehicle treatment (Figure 4.8D). Importantly, Meki treatment did not appear to have any influence on clonal dynamics in WT animals (Figure 4.8C-D).

These results show that accelerated clonal dynamics in *Kras*^{G12D} fields were dependent on active Mek1/2 signalling and could be rescued by Mek1/2 inhibition.

Previous reports have shown that Kras^{G12D} expression in the murine intestine mediates hyperproliferation, increased levels of goblet cells and decreased levels of Paneth cells (Feng *et al.*, 2011). Interestingly, it was also shown that Mek inhibitor treatment lead to a reversion of this phenotype in Kras^{G12D} animals (Feng *et al.*, 2011). Therefore, to investigate if the same could be observed in the experimental samples here and to understand the cellular effect of Meki treatment on *Kras*^{G12D} fields in more detail, controls and samples illustrated in Figure 4.8 were subjected to histology analysis for proliferation, Paneth and goblet cells.

Meki treated animals from Figure 4.8, and Meki untreated controls consisting of Flp;Kras^{G12D};Cre^{Ert};RT (*Kras*^{G12D}, N=3) and Flp;WT;Cre^{Ert};RT (WT, N=4) induced with 8 mg/kg ß-naphthoflavone and culled 21 post induction were administered a pulse of BrdU for 45 minutes before termination. The proportion of BrdU+ cells was quantified in the proximal small intestine and the colon by IHC analysis and manual scoring. 25 half-crypts were measured per animal and each cell position from base to top of crypts was scored either BrdU+ or BrdU- (Figure 4.9A+E). This analysis showed that cells in *Kras*^{G12D} crypts had an increased proportion of BrdU+ cells at most cell positions compared to WT crypts in both the small intestine and colon (Figure 4.9B+F). In addition, Meki treatment appeared to reduce proliferation in *Kras*^{G12D} but not in WT crypts (Figure 4.9B+F).

Neutral drift is a result of stem cell division and replacement and thus is impacted by cellular proliferation. During lineage tracing stem cell clonal growth is measured at the crypt base and therefore the proportion of BrdU+ cells at the crypt base could be indicative of stem cell division. To understand whether

reduced stem cell proliferation could be the cause of the repressed stem cell dynamics observed in Meki treated *Kras*^{G12D} fields the proportion of BrdU+ cells at cell position 1-5 (defined as the stem cell zone) was examined. This showed a significantly increased proportion of BrdU+ cells in *Kras*^{G12D} samples compared to WT (Figure 4.9C+G). Additionally, Meki treatment significantly decreased stem cell zone proliferation in *Kras*^{G12D} samples (Figure 4.9C+G). Interestingly, the proliferation in the stem cell zone of Meki treated *Kras*^{G12D} samples was still increased compared to WT samples in the small intestine but not in the colon (Figure 4.9C+G).

Additionally, Kras^{G12D} activation lead to significantly elongated crypts compared to WT (Figure 4.9D+H). This appeared inhibited in *Kras*^{G12D} Meki treated samples as small intestinal and colonic crypts were significantly smaller than untreated *Kras*^{G12D} crypts, however, still significantly longer than WT crypts (Figure 4.9D+H).

To further characterise the *Kras*^{G12D} mediated tissue re-organisation after Meki treatment the number of Paneth and goblet cells was examined in the small intestine of the animals described in Figure 4.9.

Paneth cells were identified by their lysozyme containing granules (Deckx, *et. al*, 1967). The number of lysozyme+ Paneth cells was quantified in 25 crypts per animal (Figure 4.10A). This analysis showed *Kras*^{G12D} animals had significantly fewer Paneth cells per crypt compared to WT animals (Figure 4.10A-B). However, Meki inhibition partially reversed this phenotype as significantly increased numbers of Paneth cells were found per *Kras*^{G12D} crypt after treatment, though this level was still significantly decreased compared to WT animals (Figure 4.10A-B).

Goblet cells were identified by Alcian blue / Pas positive mucin staining (Figure 4.10A). The number of goblet cells was quantified in 15 half villi per animal from villus bottom to villus tip (Figure 4.10A). The number of goblet cells per villus was normalised against the length (in cell number) of the villus quantified. This analysis showed an increased number of villus associated goblet cells and villus

length in *Kras*^{G12D} animals (Figure 4.10A+C). These were significantly decreased after Meki treatment in *Kras*^{G12D} animals, however, the villus length in Meki treated *Kras*^{G12D} samples was still significantly longer compared to WT (Figure 4.10A+C). Importantly, Meki treatment did not seem to have any major effect on WT crypts (Figure 4.9 + 4.10).

In conclusion, the results presented here suggest that Meki treatment partially reveres the *Kras*^{G12D} field phenotype.

4.2.7 The long-term effect of Mek1/2 inhibition

Meki treatment inhibited the accelerated pro-oncogenic clonal dynamics observed in *Kras*^{G12D} fields and reverted the *Kras*^{G12D} phenotype. Therefore, it is possible that Meki could repress the pro-oncogenic effect of *Kras*^{G12D} fields. In theory, at least, Meki treatment could be used as a chemo-preventative in the clinic. A successful chemo-preventative treatment would have little side-effects and theoretically be given in cycles (and not continuously) during patient lifespan. Meki treatment should therefore ideally have a long-term effect to achieve activity even during the pauses of treatment cycles.

Thus, to test whether Meki treatment had long-term effects on clonal dynamics, *Kras*^{G12D} animals were treated with Meki in a cycle of Meki on and Meki off (Meki on/off) and clonal dynamics were examined. Experimentally, Flp;Kras^{G12D};Cre^{Ert};RT (*Kras*^{G12D}) or Flp;WT;Cre^{Ert};RT (WT) animals were administered 8 mg/kg ß-naphthoflavone to induce Kras^{G12D} expression (or not) on day -51 (Figure 4.11A). 21 days later Meki treatment was administered for 15 days as described in Figure 4.8 (Meki on) (Figure 4.11A). On day -15 Meki treatment was stopped (Meki off) and lineage tracing was initiated on day 0 for a total of 14 days of pulse-chasing labelling (Figure 4.11A).

Examination of pMek by IHC showed similar levels in untreated animals and Meki on/off animals both in the small intestine and colon suggesting a return of normal Mek1/2 levels after termination of Meki treatment (Figure 4.11B). The average

clone size and percentage of WPCs found in this experiment was compared back to data from Figure 4.8. The average clone size and percentage of WPCs were significantly increased in *Kras*^{G12D} Meki on/off animals compared to *Kras*^{G12D} Meki samples (Figure 4.11C-D), suggesting a reversion of clonal dynamics after Meki treatment ended. If *Kras*^{G12D} clonal dynamics were completely reverted after termination of Meki treatment no significant difference would be expected between untreated *Kras*^{G12D} and Meki on/off samples. Interestingly, there appeared to be a trend that the clonal dynamics in *Kras*^{G12D} fields were not fully reverted to *Kras*^{G12D} levels after Meki treatment termination (Figure 4.11C-D). This was underlined by the fact that the average clone size and the percentage of WPCs in the *Kras*^{G12D} Meki on/off small intestinal and colon were significantly decreased compared to *Kras*^{G12D} untreated samples, respectively (Figure 4.11C-D). However, the difference in average clone size and percentage of WPCs between *Kras*^{G12D} Meki on/off and *Kras*^{G12D} failed to reach significance in the colon and small intestine, respectively (Figure 4.11C-D).

Conclusively, these results show that clonal dynamics in *Kras*^{G12D} Meki treated fields partially revert back to normal *Kras*^{G12D} levels after treatment suggesting a limited sustained long-term effect of Meki.

To further examine whether *Kras*^{G12D} mediated tissue changes were also reverted after cessation of Meki treatment, BrdU positivity and number of Paneth and goblet cells were examined in Meki on/off animals from Figure 4.11. Next, this data was compared to the data presented in Figure 4.9+10. As no difference was observed in WT samples this analysis only included *Kras*^{G12D} samples.

Analysis of proportion of BrdU+ cells at all cell positions in the small intestinal and colonic crypts showed that cells in *Kras*^{G12D} Meki on/off crypts had reached (and in some cases exceeded) levels in *Kras*^{G12D} crypts (Figure 4.12A-B+E-F). This suggested a proliferative burst in Meki *Kras*^{G12D} fields after treatment ended. Examination of proportion of BrdU+ cells in the stem cell zone showed a significant increase in S-phase cells in *Kras*^{G12D} Meki on/off samples compared to *Kras*^{G12D} Meki in both the small intestine and colon (Figure

4.12C+G). In addition, the proportion of BrdU+ cells in the stem cell zone in *Kras*^{G12D} Meki on/off was not different to levels in *Kras*^{G12D} untreated samples (Figure 4.12C+G). This suggested that proliferation was restored back to *Kras*^{G12D} levels after Meki treatment ended. However, small intestinal Meki on/off crypts were slightly longer than *Kras*^{G12D} crypts whereas Meki on/off colonic crypts were shorter than *Kras*^{G12D} (Figure 4.12D+H).

Lastly, quantification of number of crypt-base Paneth cells and villus associated goblet cells showed decreased Paneth cell numbers and increased goblet cell numbers and villus length in Meki on/off samples compared to Meki samples (Figure 4.13A-C). Interestingly, Paneth cell numbers and villus length in the Meki on/off samples was not different to levels observed in *Kras*^{G12D} untreated samples (Figure 4.13A-D). However, the number of goblet cells in *Kras*^{G12D} Meki on/off samples was significantly increased compared to untreated *Kras*^{G12D} samples suggesting a spike in goblet cell fate choice after Meki termination (Figure 4.13A+D).

Together, BrdU, Paneth and goblet cell counts confirm that Meki treated *Kras*^{G12D} fields revert back to full *Kras*^{G12D} phenotypic status after treatment ended. From these counts, it appeared that this transition was mostly complete.

4.3 Discussion

Conclusively, the results presented here show that clonal dynamics are accelerated in *Kras*^{G12D} fields of the mouse small intestine and colon. In addition, the results demonstrate that *Kras*^{G12D} fields accelerate the process of fixing new mutations as well as accelerating tumour initiation. These features make *Kras*^{G12D} fields more likely to be pro-oncogenic.

The accelerated clonal dynamics can be partially repressed, short and long term, by inhibiting Mek1/2 signalling downstream of activated Kras. Thus, Mek1/2 inhibition could serve as a chemo preventative for cancer associated with oncogenic *KRAS* mutations.

Stem cell dynamics have not previously been calculated in the new steady state that is achieved in crypts entirely populated by *Kras*^{G12D} mutant cells or in any pro-oncogenic field. Vermeulen *et al.* 2013 showed that 30% of crypts were monoclonal after 21 days when a single *Kras*^{G12D} stem cell competes with WT stem cells for the niche during clonal drift. Thus, the finding that 60% of crypts were monoclonal in *Kras*^{G12D} mutant crypts after 21 days show that stem cell dynamics are additionally accelerated during neutral competition in fields compared to when *Kras*^{G12D} cells are competing with WT cells during biased competition.

In addition, mathematical modelling revealed that clonal dynamics in *Kras*^{G12D} fields are accelerated by employing relatively fewer stem cells with an increased stem cell replacement rate per crypt for neutral drift. Even though the difference between inferred N and λ in *Kras*^{G12D} and WT appear modest, these changes in combination have a dramatic effect on monoclonal conversion times. This is exemplified by the fact that it takes *Kras*^{G12D} crypts and WT crypts 15 and 45 days, respectively, to reach an average clone size of 0.75.

Analysis of clonal dynamics in human colons has revealed that crypts contain 7 stem cells and have a neutral monoclonal conversion rate of approximately 6.3 years (Nicholson *et al.*, 2018). This is approximately 100-fold slower than in mice. The results in this chapter show that monoclonal conversion

rate is at least 2-fold accelerated in *Kras*^{G12D} fields. If the same effect is true in humans, this means that fixation of new mutations (some of which could be pro-oncogenic) would only take approximately 3 years in a human colonic *KRAS* field. Thus, the relative difference in clonal dynamics characterised here in mice confers a paramount difference in the human equivalent case and could have a major impact on tumour progression time in patients with mutant *KRAS* fields. Thus, to verify these findings, it would be interesting to determine the human clonal fixation time within a mutated *KRAS* field.

Neutral drift is a process of cell division and clonal replacement (Lopez-Garcia *et al.*, 2010; Snippert *et al.*, 2010). Inherently this links stem cell division to the rate by which single clones convert a crypt to monoclonality. However, the link between the two is not yet clear as there are many more stem cell divisions than inferred replacement events during neutral drift.

Here, the results show a link between neutral drift behaviour and crypt-base BrdU positivity. This is underlined by the fact that *Kras*^{G12D} crypts have both increased BrdU+ cells at the crypt base and accelerated clonal dynamics. Additionally, Meki1/2 inhibition decreases both *Kras*^{G12D} clonal drift and BrdU incorporation. This could suggest that a reduced stem cell cycle time is the general trait that drives the accelerated clonal dynamics in *Kras*^{G12D} fields. However, BrdU incorporation only indicates which and how many cells are in S-phase over a given period of time and does not specify an actual cell-cycle/division time per cell. Therefore, in the future, it will be important to determine whether the stem cell division time in *Kras*^{G12D} fields is in fact reduced. This could be done by employing the Fucci system or triple nucleotide labelling in *Kras*^{G12D} crypts to visualise and quantify the length of the different phases of the cell cycle (Sakaue-Sawano *et al.*, 2008; Podgorny *et al.*, 2018). Another approach could be to measure the time it takes a single cell ‘clone’ to divide immediately post clone-labelling as done by Andersen *et al.* 2019. Here the authors show that Kras^{G12D} expression in skin sebaceous-gland progenitors has only a small effect on cellular division time although gland expansion is increased in a *Kras*^{G12D} dependent manner (Andersen *et al.*, 2019). This could imply that accelerated *Kras*^{G12D} clonal behaviour is not only a result of cellular proliferation.

Thus, other factors could also play a role for the accelerated clonal growth in *Kras*^{G12D} fields. It is speculated that cell-cell interactions and cell morphology could be important mechanisms of epithelial clonal expansion (Vincent, *et al*., 2013; Levayer and Moreno, 2016). Consequently, an altered morphology and physical interaction of stem cells in *Kras*^{G12D} mutant crypts could be a driver of the accelerated clonal dynamics observed. This could be fuelled by the loss of Paneth cells located between stem cells in *Kras*^{G12D} small intestinal mutant crypts (Sato *et al.*, 2011) which could create enhanced cell-cell interaction and stem cell replacement patterns.

To functionally investigate the pro-oncogenic potential of the accelerated clonal dynamics in *Kras*^{G12D} fields, crypt mutational accumulation was tested.

In a recent study it was shown that *Apc*^{min} pre-cancer crypts acquire 11-fold more somatic mutations in their genome compared to normal crypts during ageing (Lugli, *et al.*, 2017). Here, a 1.3-fold increased mutational burden was reported in *Kras*^{G12D} crypts compared to WT crypts. The increased mutational burden could be a direct result of the accelerated clonal dynamics observed in *Kras*^{G12D} mutant crypts or it could be the effect of an additional increased mutation rate. An increased mutation rate would imply that *Kras*^{G12D} cells inherently acquire more mutations per cell division than WT cells. This would make *Kras*^{G12D} mutant cells hyper-mutagenic and thus could be an additive mechanism for enhanced oncogenesis in *Kras*^{G12D} fields. However, the fold difference in mutational burden reported in *Apc*^{min} crypts by Lugli and colleagues compared to what is observed in *Kras*^{G12D} crypts could explain why spontaneous tumours are only rarely formed in *Kras*^{G12D} activated murine intestines compared to frequently in *Apc*^{min} mice (Moser, *et. al*, 1990). In the future, additional analysis of the exome sequencing data presented here, and mathematical modelling will be utilized to determine the exact mutation rate and mutational signature(s) of *Kras*^{G12D} mutant crypts.

To be able to assess whether tumour initiation time was in fact decreased in *Kras*^{G12D} fields sporadic stabilised nuclear ß-catenin clones were left to progress

into tumours over time in WT or *Kras*^{G12D} fields. The results presented here show that the time it takes for stabilised ß-catenin clones to develop into tumours is greatly decreased in *Kras*^{G12D} fields.

It has been described that Kras signalling interacts with the Wnt-signalling pathway thus supporting a synergistic effect of *Kras*^{G12D} and stabilised nuclear ß-catenin (Janssen *et al.*, 2006; Moon *et al.*, 2013; Lemieux *et al.*, 2015). Therefore, it is not clear whether *Kras*^{G12D} fields simply accelerate tumour initiation time and/or additionally increases the rate of growth of nuclear ß-catenin driven tumours. This could be investigated by examining how many nuclear ß-catenin crypts are converted into adenomas in *Kras*^{G12D} and WT fields over time. In addition, it will be important to understand whether the potential added effect could stem from an additional competitive bias of stabilised ß-catenin+/*Kras*^{G12D}+ stem cells during clonal drift. Accordingly, lineage tracing of stabilised nuclear ß-catenin stem cells in *Kras*^{G12D} fields could determine the probability of replacement of ß-catenin+/*Kras*^{G12D} vs *Kras*^{G12D} cells in detail.

Results in this chapter strongly suggest that the clonal behaviour in *Kras*^{G12D} fields makes them pro-oncogenic. Thus, it is exciting to theorise about the possibility of inhibiting this behaviour to prevent cancer initiation, so-called chemo-prevention (Issa and Noureddine, 2017). In this chapter, a clinically available Mek1/2 inhibitor is employed to do so.

The data presented here show that Mek1/2 inhibition of *Kras*^{G12D} fields could repress the accelerated clonal dynamics. Importantly, Meki1/2 inhibition did not seem to have any influence on clonal growth in WT tissue, suggesting that neutral drift in WT tissue is independent on Mek1/2. Thus, Meki could be used as a chemo-preventative agent in patients diagnosed with oncogenic *KRAS* mutations prior to cancer diagnosis without significantly affecting WT crypts. In theory this treatment could prolong such individuals' tumour free period and/or prevent the disease all together.

Interestingly, Meki1/2 inhibition only partially rescued clonal dynamics and cellular changes in *Kras*^{G12D} fields. This could be because of the length of the treatment (15 days), suggesting that a longer treatment period or single

treatments given periodically, could regress the *Kras*^{G12D} phenotype further. However, this could also be because activated Kras does not only signal via Mek1/2 (Cox *et al.*, 2014). This is supported by the findings that Mek1/2 inhibition only has a limited effect on tumour regression in advanced colorectal cancer (Hochster *et al.*, 2015). However, studies have shown that combination of Mek1/2 inhibitors and inhibitors of the Wnt-pathway can supress tumour progression *in vivo* and in clinical phase-IB trials (Krishnamurthy *et al.*, 2018; Moon *et al.*, 2019). Thus, a combination of Mek1/2 inhibition and other targeted therapies could potentially completely rescue clonal dynamics in *Kras*^{G12D} fields. However, it is not clear what the toxic effect of long-term inhibitor therapy would be in "healthy" individuals.

After termination of Mek1/2 inhibition clonal dynamics revert partially back to *Kras*^{G12D} untreated levels. This partial reversion indicates that Mek1/2 inhibition might have a lingering long-term effect. However, proliferation, Paneth cell loss and villus length fully returned to normal *Kras*^{G12D} levels after termination of Meki treatment. This implies that stem cell dynamics respond slower to termination of Mek1/2 inhibition than cellular changes do. This seems intuitive, proliferation and cell death (Paneth cell loss) are immediate cellular responses whereas clonal growth is a dynamic change.

Together, the results in this chapter explain how oncogenic colonic *KRAS* fields could pre-dispose to cancer development. However, the notion that pro-oncogenic *KRAS* fields are present in patients before mutations in genes such as *APC* challenges the sequence of events predicted by Fearon and Vogelstein in 1990. Thus, it will be important to convince the scientific community that a subset of tumours in fact arise in mutated *KRAS* fields. This could be done by investigating mutated *KRAS* allelic frequencies in human tumours. If tumours have arisen in oncogenic *KRAS* fields they would be expected to have a near clonal *KRAS* allele frequency.

5 Inducible Dre recombinase: a tool for temporal sequential colorectal cancer modelling *in vivo*

5.1 Introduction

Colorectal cancer is the third most common cancer worldwide (Kuipers *et al.*, 2015). The disease develops by acquisition of mutations in the colonic epithelium during patient lifespan which can lead to development of early polyp and adenoma lesions that can further progress into late adenocarcinomas (Fearon and Vogelstein, 1990). Even though, Fearon and Vogelstein determined a mutational sequence of colorectal cancer development in 1990, recent studies have shown that the disease is heterogenous and can be driven many diverse events that can classify different colorectal cancer sub-types (The Cancer Genome Atlas Network, 2012; De Sousa E Melo *et al.*, 2013; Sadanandam *et al.*, 2013; Phipps *et al.*, 2015; Nojadeh *et al.*, 2018). Although tumours of different subtypes are driven by different genetic alterations it is clear that the development of any colorectal tumour is a sequential process.

The availability of different recombination tools makes it possible to recreate the sequential nature of the disease *in vivo*. As an example, following stem cells in pro-oncogenic fields or in tumours *in vivo* requires two events separated in time: firstly, a tumour/oncogenic field causing recombination and secondly a recombination event initiating lineage tracing. Expanding the genetic tool box to include different recombinases will be crucial to fine-tune such modelling.

The most widely used DNA recombinase in *in vivo* models is Cre. Cre was discovered in bacteriophages in 1981 by Stenberg and colleagues (Sternberg and Hamilton, 1981) and was shown to recombine DNA between 34 base-pair (bp) loxP sites in mammalian cells in 1988 (Sauer and Henderson, 1988). The discovery of the Cre/loxP system was the starting point of conditional *in vivo* modelling which has been widely used since the first description of recombination driven retina eye-lens cancer in mice (Lakso *et al.*, 1992). Later, Sadowski *et al.*

described a novel DNA-recombinase named Flp that recognises and recombines DNA between 34 bp Frt sites (Sadowski, 1995) which had no cross-reactivity with Cre. This expanded the recombinases available and allowed sequential modelling.

Historically, most disease associated conditional alleles encode loxP sites for Cre driven recombination, for instance; *Tp53*$^{flox2\text{-}10}$ (Marino *et al.*, 2000), *Pik3ca*^{H1047R} (Yuan *et al.*, 2013), *Pten*^{flox} (Groszer *et al.*, 2001), *Apc*^{flox} (Cheung *et al.*, 2009) and *Ctnnb1*$^{flox(ex3)}$ (Harada *et al.*, 1999). Some frt alleles for Flp driven recombination are also available, for instance; *Kras*^{fsfG12D} (Young *et al.,* 2011) and *Tp53*^{frt} (Lee *et al.*, 2012). In the fourth chapter of this thesis a mouse model that combines Flp and Cre is utilized for separating two individual recombination events in time. There, *Cyp1A1*^{FlpO} drives the expression of *Kras*^{fsfG12D} and Ah^{CreErt} drives lineage tracing in *Kras*^{G12D} fields. However, using Flp (and not Cre) for the initial recombination event greatly decreased this model's versatility as it becomes dependent on the limiting Frt allele catalogue for modelling of different phenotypes. Thus, another tool for lineage tracing that can be combined with Cre and Flp could enhance *in vivo* sequential modelling.

Dre is a novel DNA recombinase that recognizes 32 bp rox sites, and it was discovered in a screen for Cre-like enzymes in P1-like phages in 2004 (Sauer and McDermott, 2004). Later, Anastassiadis *et al.* showed that the Dre/rox system does not cross-react with the Cre/loxP system in mouse embryonic stem cells (mESCs) and in mouse embryos (Anastassiadis *et al.*, 2009). Since these initial studies, Dre, and Cre (Hermann *et al.*, 2014; Madisen *et al.*, 2015; Sajgo *et al.*, 2014) and Dre, Cre and Flp (Plummer *et al.*, 2015) have been combined in the same *in vivo* and *in vitro* models to identify cell populations defined by differential promoter activity. Furthermore, Dre and Flp enzymes have also been utilized for designing loxP targeting alleles by allowing removal of Frt-site and rox-site flanked selection cassettes (Motojima *et al.*, 2016).

Importantly, Anastassiadis *et al.* also showed that Dre could be fused to the human progesterone receptor (Pr) to create an inducible Dre^{Pr} fusion

protein. This fusion protein was found to be activatable in fibroblasts by the synthetic anti-progestin Ru486 (Anastassiadis *et al.*, 2009). Additionally, Dre^{Pr} has been shown to be active in Zebrafish (Park and Leach, 2013), however, it has never been proven active in mammalian tissues *in vivo*.

Sequential modelling requires spatiotemporal control of DNA-recombinase expression and activation. To ensure optimal separation the recombinases used in sequence should be expressed under different promoters and should be activated by different ligands. Creating a murine Dre^{Pr} allele would allow for a Ru486 inducible Dre different to the Tamoxifen inducible Cre^{Ert} and the ß-naphthoflavone inducible *Cyp1A1*^{FlpO}.

The housekeeping locus *Rosa26 (R26)* confers a constitutive and ubiquitous expression of targeted genes in a variety of different tissues (Kisseberth *et al.*, 1999) and has been reported as a promoter leading to widespread expression of Dre in the mouse embryo (Anastassiadis *et al.*, 2009). Thus, in theory Dre^{Pr} targeted to the *R26* allele should be expressed in all cell lineages, including intestinal stem cells. Hence, *R26*^{DrePr} (R_Dre^{Pr}) could be a good candidate for a marker-free inducible lineage tracing tool that can be combined with other recombinases without cross-reactivity.

Thus, the aim of this chapter is to develop R_Dre and R_Dre^{Pr} alleles to conduct lineage tracing in both homeostatic intestinal tissue and in a sequential model of intestinal tumorigenesis.

5.2 Results

5.2.1 Vector construction

The R_Dre and R_Dre^{Pr} animals were created by homologous recombination of targeting vectors (Appendix 2 for full vector maps, Figure 5.1A for schematic) into the *R26* locus in mouse embryonic stem cells (mESCs) as described before (Vooijs, *et al.*, 2001). In the R_Dre^{Pr} targeting vector Pr consisted of the Ru486 responsive mutant hormone binding domain of the human progesterone receptor hPR891 (described by Kellendonk *et al.*, 1996). Pr was fused to the N-terminus of Dre thereby destroying parts of the endogenous nuclear localisation signal (NLS) in Dre and the fusion site was identical to the one reported by Anastassiadis *et al.*, 2009. The finished vectors underwent Sanger sequencing to ensure all sequences were correct before proceeding with mESC targeting (data now shown).

5.2.2 mESC integration site validation

Screening for correct 5' integration was carried out by PCR amplifying genomic DNA (gDNA) from mESC colonies with F1 and R1 primers (Figure 5.1A, Appendix 1 for primers). *R26* targeting by mESC homologous recombination was effective, with 72/96 F1+R1 positive colonies, giving the targeting event a success rate of 75% (Figure 5.1B). Moreover, the targeting was of similar success rate in R_Dre and R_Dre^{Pr} colonies (Figure 5.1B).

To ensure that the 3' end of the construct was also integrated correctly the first ten F1+R1 positive colonies from either R_Dre or R_Dre^{Pr} mESC colonies (Figure 5.1B) were further expanded *in vitro* and subjected to PCR amplification with primers F2 and R2 (Figure 5.1A). Importantly, all ten colonies from R_Dre and R_Dre^{Pr} were also correctly integrated at the 3' end (Figure 5.1C).

This analysis validates Dre and Dre^{Pr} integration into the intended locus of *R26* in the mouse genome upon homologous recombination.

5.2.3 Copy number assessment

mESC clones with a single copy insertion of either R_Dre or R_Dre^{Pr} were desired for animal creation. Clones with more than one copy could arise by: a) a tandem insertion at one *R26* locus or b) an insert on both *R26* alleles or c) an additional insert in a random genomic loci. To identify single copy mESC clones a quantitative real time PCR (qRT-PCR) copy number assay was carried out on gDNA from the ten correctly targeted clones shown in Figure 5.1C. The qRT-PCR was performed with a puromycin probe serving as a surrogate marker for "inserted vector" (as targeting vectors contained a puromycin selection cassette). All samples were normalised against a sample known to contain a single copy of puromycin (from a previous *R26* targeted mouse known to be heterozygous) (Figure 5.2).

Importantly, gDNA from the negative control sample with untargeted mESCs did not amplify, underlining the specificity of the puromycin probe (Figure 5.2). As seen on Figure 5.2, all clones except for two in R_Dre group contained a single copy of puromycin. To avoid clones with tandem or random insertions only single copy R_Dre and R_Dre^{Pr} clones were chosen to take forward to mouse oocyte pro-nuclear injections and mouse model creation.

5.2.4 R_Dre mice have widespread Dre expression

The R_Dre model serves as a positive control that allows for investigation of the expression of Dre. To investigate Dre activity in different tissues, R_Dre animals were crossed to animals encoding a rox STOP rox (rsr) tdTomato (tdTom) reporter allele (*R26*^{rsrtdTom}). The *R26*^{rsrtdTom} was created by germline deletion of the loxP flanked STOP cassette (lsl) from the R26^{rsrlsltdTom} mouse obtained from Jackson Laboratories (stock no. 021876). The resulting line; R_Dre;*R26*^{rsrtdTom} will be referred to as RD;RT^{rox} and all experimental animals were heterozygote

for both alleles. Expression of tdTom visualised by red fluorescence (excited at 550nm) served as a marker for Dre activity in these mice.

To analyse Dre activity, adult 8-10 week old RD;RT^{rox} animals were culled and organs from different germ-layers were collected for immunohistochemical (IHC) analysis. From endoderm: stomach, liver, pancreas, lung and the intestines. From mesoderm: kidney, spleen, bone, striated muscle in tongue and smooth muscle in heart. From surface ectoderm: skin. Figure 5.3 shows IHC analysis of tdTom expression in above-mentioned tissues from a control RT^{rox} only mouse and three individual RD;RT^{rox} animals. All tissues examined expressed tdTom in RD;RT^{rox} animals (Figure 5.3). tdTom expression was Dre specific, as the RT^{rox} control did not contain any tdTom expression (Figure 5.3). These results show that R_Dre targeting leads to ubiquitous expression of Dre recombinase in tissues of different germ layers during development that persist in adult tissues. Importantly, this seems to be true, also for the intestine, which was of special interest for this study.

5.2.5 Dre^{Pr} is activated in a Ru486 dose dependent manner

To investigate Dre^{Pr} activity in adult tissue, R_Dre^{Pr} animals were crossed to *R26*^{rsrtdTom} mice to create R_Dre^{Pr}; *R26*^{rsrtdTom} animals (referred to as RD^{Pr};RT^{rox}). All experimental animals were heterozygous for both alleles. As described above, tdTom expression served as a marker of Dre^{Pr} activity in this line.

The Pr addition should retain Dre recombinase in the cytoplasm of cells until administration Ru486. Upon Ru486 ligand binding to Pr, Dre should translocate to the cell nucleus to mediate DNA recombinase activity. To investigate whether Dre^{Pr} was activatable by Ru486 and to determine the maximal induction efficiency of Dre^{Pr}, RD^{Pr};RT^{rox} animals had 1, 2 or 3, 90-day slow release pellets containing 10 mg/pellet Ru486 implanted sub-cutaneously (Figure 5.4A). These pellets provide a slow but continuous release of Ru486 made possible by integrating the drug in a bio-degradable matrix. At 75 days post implantation animals were sacrificed, and tissues from different germ-layers were collected for tdTom expression analysis.

Figure 5.4B shows IHC analysis of tdTom in RD^{Pr};RT^{rox} animals either un-induced or implanted with Ru486 slow release pellets. Importantly, RD^{Pr};RT^{rox} contained no background recombination, as un-induced controls had no tdTom expression in any tissue (Figure 5.4B). Furthermore, R_Dre^{Pr} seemed to be activated in a Ru486 dose dependent manner in the stomach, liver, kidney, pancreas and spleen, with the number of tdTom positive cells increasing with number of pellets implanted (Figure 5.4B). However, the heart, tongue and bone had no tdTom expression (Figure 5.4B).

As the intestine was of special interest to this study the small intestine and colon of animals implanted with Ru486 slow release pellets were subjected to further quantitative analysis. Notably, slow release of Ru486 did in fact lead to accumulation of tdTom in small intestinal crypts/villi and colonic crypts (Figure 5.5A). Importantly, no tdTom positive cells were observed in RD^{Pr};RT^{rox} with no pellet(s) inserted (Figure 5.5A). To quantify the number of tdTom+ epithelial cells upon pellet administration the first five centimetres of the small intestine was subjected to single cell suspension and analysed by flow cytometry to determine the percentage of tdTom+ cells in EpCam+ epithelial cells (EpCam+/tdTom+) (Figure 5.5B). Flow cytometric analysis showed that proximal small intestines from animals with 1 pellet implanted had ~2% EpCam+/tdTom+, animals with 2 pellets had ~3% EpCam+/tdTom+ and animals with 3 pellets had ~5% EpCam+/tdTom+ thus showing an almost linear accumulation of tdTom+ cells when treated with Ru486 slow release pellets (Figure 5.5B-C).

These results show that R_Dre^{Pr} has little to no background activity and suggests that it is activated in a dose dependent manner by Ru486 in various tissues, including the intestines.

5.2.6 Dre^{Pr} is active in intestinal stem cells and can be used for lineage tracing

As seen in Figure 5.6A, tdTom could be observed in whole crypts and villi after Ru486 exposure both in the small intestine and colon. This underlined that Dre^{Pr} could be activated in intestinal stem cells, and that these stem cells could give rise to differentiated daughter cells migrating up the crypt (Figure 5.6A). To quantify the number of lineage-traced clones present after pellet implantation the proximal small intestine and colon were subjected to manual tdTom+ clone quantification by fluorescent microscopy (Figure 5.6B-C). Clones were defined as tdTom+ crypts (Figure 5.6B). The number of crypt clones both in the small intestine and colon arising after 1, 2 or 3 pellet implantations increased in a linear fashion (Figure 5.6B-C). However, a decrease in R_Dre^{Pr} activity from small intestine to colon was observed, with approximately half the number of clones quantified in colon (~100 clones/cm/pellet) compared to small intestine (~200 clones/cm/pellet) (Figure 5.6C).

Next, R_Dre^{Pr}'s ability to operate as a lineage trace tool in the mouse intestine was examined. Lineage tracing is commonly carried out by pulse chase experiments. To investigate the activity of Dre^{Pr} after a pulse of inducer, Ru486 was administered to RD^{Pr};RT^{rox} mice by the intra peritoneal (I.P) route either by a single dose of 50 or 80 mg/kg or by 80 mg/kg administered on three consecutive days. After 14 days animals were culled, and the small intestine and colon were subjected to manual crypt clone quantification by fluorescent microscopy analysis. As seen on figure 5.7A, intraperitoneal administration of Ru486 gave rise to tdTom+ clones both in the small intestine and colon. Additionally, the number of clones observed 14 days post induction increased in a dose responsive fashion both in the small intestine and colon (Figure 5.7B). The number of clones/cm in the small intestine was ~100 vs ~300 following a single or three dose(s) of 80 mg/kg Ru586, respectively, showing an almost linear accumulation of signal (Figure 5.7B). From this analysis 50 mg/kg Ru486 was used as the pulsing dose in lineage trace experiments as this dose appeared to hit <1 stem cell per crypt in small intestine and colon (Figure 5.7C), was well tolerated by the mice and provided an appropriate number of clones.

To further examine whether Dre^{Pr} could be a successful tool for lineage tracing, results obtained on stem cell dynamics in the widely used and validated Ah^{CreErt};tdTom (Ah^{CreErt};RT) model was compared to the RD^{Pr};RT^{rox} model. Ah^{CreErt};RT^{lox} animals received a single dose of 40 mg/kg ß-naphthoflavone and 0.15 mg Tamoxifen and RD^{Pr};RT^{rox} animals received a single dose of 50mg/kg Ru486 and animals from both strains were sacrificed 4, 7, 10, 14 and 21 days post drug administration (Figure 5.8A). The proximal small intestine and colon from both lines were then subjected to manual clone quantification by fluorescent microscopy to score the relative sizes of tdTom+ clones in intestinal crypts (scored as fractions of 8, Figure 5.8B). 3 mice per timepoint were quantified. In RD^{Pr};RT^{rox} animals an average of 100-250 and 85-170 clones/animal/timepoint were quantified in the small intestine and colon, respectively. In Ah^{CreErt};RT^{lox} animals an average of 200-300 and 175-300 clones/animal/timepoint were quantified in the small intestine and colon, respectively. Highest clone numbers were found at earliest timepoints.

During lineage tracing the average clone size and clone size distribution in small intestine and colon were not significantly different in the two models (Figure 5.8C + D).

Finally, to quantitively compare stem cell dynamics in the two lines mathematical modelling was used to predict stem cell number per crypt (N) and stem cell replacement rate per stem cell per day (λ) in RD^{Pr};RT^{rox} and Ah^{CreErt};RT^{lox} animals (Figure 5.9A-B). The mathematical model utilized encoded the same assumptions as the model described in chapter 4. Here, the model predicted average clone sizes over 100 days in RD^{Pr};RT^{rox} and Ah^{CreErt};RT^{lox} animals (Fig 5.9A). The analysis inferred N to be 6 and 7 in the small intestine and colon, respectively, in both the RD^{Pr};RT^{rox} and Ah^{CreErt};RT^{lox} models (Figure 5.10B). Furthermore, λ was estimated to be 0.147 vs 0.15 in the small intestine and to 0.19 vs 0.20 in the colon of RD^{Pr};RT^{rox} and Ah^{CreErt};RT^{lox} animals, respectively (Figure 5.9B). Importantly the average clone sizes presented here are in line with data from lineage tracing in other mouse models as N and λ were similar to what was inferred in previous reports (Kozar *et al.*, 2013;Vermeulen *et al.*, 2013).

Taken together, the data suggests that R_Dre^{Pr} traced stem cells have a non-biased neutral behaviour similar to that of Ah^{CreErt} traced stem cells. Therefore, the R_Dre^{Pr} allele can successfully be used for lineage tracing in small intestine and colon.

5.2.7 R_Dre^{Pr} can trace single cell derived clones in intestinal tumours

A lineage tracing tool that is not activated by Tamoxifen and which is different from Cre is needed to follow single cell derived clones in Cre recombination facilitated tumour models such as the *Ctnnb1*$^{flox(ex3)}$ model (Harada *et al.*, 1999). Here, a new mouse model to trace the fate of single cells in intestinal tumours, is presented: the Ah^{CreErt};*Ctnnb1*$^{lox(ex3)}$;RD^{Pr};RT^{rox} mouse. In this model, Ah^{CreErt} was utilized to induce tumour initiation and R_Dre^{Pr} was utilized to trace clones derived from single cells within these tumours.

To use the Ah^{CreErt} and R_Dre^{Pr} as sequential tools in the same model it was important to determine that the drugs used to activate Ah^{CreErt} (ß-naphthoflavone and Tamoxifen) did not activate the second recombinase in the sequence: Dre^{Pr}. If ß-naphthoflavone and Tamoxifen would activate Dre^{Pr}, background tdTom clones could greatly skew lineage tracing. In addition, it was important to confirm the previous finding that Cre does not recombine RT^{rox}, as this would also create background tdTom clones.

To this end, RD^{Pr};RT^{rox} and Ah^{CreErt};RT^{rox} animals were administered 40 mg/kg ß-naphthoflavone and 2 mg Tamoxifen on three consecutive days. Ah^{CreErt};RT^{lox} animals + ß-naphthoflavone and Tamoxifen served as a positive control for Cre activity. RD^{Pr};RT^{rox} animals + Ru486 served as a positive control for Dre^{Pr} activity. As shown in Figure 5.10A, the ß-naphthoflavone / Tamoxifen treated RD^{Pr};RT^{rox} and Ah^{CreErt};RT^{rox} animals did not contain any tdTom+ cells in the small intestine or colon seven days following initial I.P when subjected to confocal microscopy. In addition, flow cytometry quantification of recombination in the proximal small intestine showed 0% EpCam+/tdTom+ in both RD^{Pr};RT^{rox} and Ah^{CreErt};RT^{rox} animals after ß-naphthoflavone and Tamoxifen exposure

(Figure 5.10B). Importantly, the positive control Ah^{CreErt};RT^{lox} + ß-naphthoflavone / Tamoxifen and RD^{Pr};RT^{rox} + Ru486 did contain 91% and 1.5% EpCam+/tdTom+ cells respectively, underlining Cre and Dre activity in both models (Figure 5.10A-B). These results demonstrate that Ah^{CreErt} and R_Dre^{Pr} can be used sequentially in one mouse line without background activity.

Next, pilot studies to confirm the ability to trace from single cells in early and late tumours were carried out. Ah^{CreErt};*Ctnnb1*$^{lox(ex3)}$;RD^{Pr};RT^{rox} animals were administered 40 mg/kg ß-naphthoflavone + 2.5 mg Tamoxifen to initiate tumours and after 21 days, lineage tracing was initiated with a pulse of 50 mg/kg Ru486. Animals were culled 5, 10 and 15 days after lineage trace induction (Figure 5.11A). This experimental set-up was enough to stabilize ß-catenin and produce hyperplastic crypts in the proximal small intestine 26 days post *Ctnnb1* recombination (Figure 5.11B). At 36 days post *Ctnnb1* recombination some of these small intestinal crypts had developed into tumours (Figure 5.11B). To visualise clones, whole-mounted tissue was subjected to tissue clearing and fluorescence microscopy. Ru486 induction marked single cells in the bottom of the hyperplastic crypts in early tumour tissue (Figure 5.11C, Day 5). With time, fewer but larger clones were observed (Figure 5.11C, Day 10). In late tumours, fewer than 5 clones/tumour were observed (Figure 5.11C, Day 15). Although few, these clones appeared to originate from a single cell and to be fully clonal (Figure 5.11C, Day 15).

To determine the number of traceable cells per tumour Ah^{CreErt};*Ctnnb1*$^{lox(ex3)}$; RD^{Pr};RT^{rox} animas where subjected to the same induction scheme as described above and aged until maximum tumour burden (38 and 40 days post induction) (Figure 5.12A-B). Hereafter, each whole-mounted tumour was subjected to tissue clearing and fluorescence microscopy to score the number of clones per tumour. As seen in Figure 5.12C tumours of different sizes and with different number of clones were observed. Figure 5.12D shows the distribution of clones in all tumours. Out of a total of 133 tumours analysed (N=2 animals) 68 tumours (51%) did not contain any clones. Furthermore, 29% of tumours contained a single clone and the largest number of clones observed was 8 (observed in a single tumour)

(Figure 5.12D). The mean clone number per tumours was found to be 0.92 (Figure 5.12E).

In conclusion these results show that R_Dre^{Pr} can be used for intra-tumour lineage tracing to follow stem cells in tumours from Ah^{CreErt};*Ctnnb1*$^{lox(ex3)}$; RD^{Pr};RT^{rox} animals.

5.3 Discussion

Different DNA recombinases are needed to drive sequential activating or deletion of disease-causing alleles and reporters to model step-wise intestinal cancer *in vivo*. In this chapter the inducible R_Dre^{Pr} mouse model is introduced. The results show that Dre^{Pr} is activated in a dose-dependent manner and that it can be used for lineage tracing in the small intestine and colon. In addition, R_Dre^{Pr} can be combined with Ah^{CreErt} for lineage tracing in stabilised ß-catenin driven tumours.

Once generated the R_Dre and R_Dre^{Pr} animals were crossed to a *R26*^{rsrtdTom} reporter to produce RD/RT^{rox} and RD^{Pr}/RT^{rox} models. When assessing Dre driven tdTom expression in different tissues in adult RD/RT^{rox} animals a strong and widespread expression pattern was observed. This highlights Dre activity during development. Furthermore, analysis of tdTom expression in various tissues of RD^{Pr}/RT^{rox} animals after Ru486 slow release pellet implantation showed that most tissues expressed nuclear Dre^{Pr} after Ru486 exposure, confirming Dre^{Pr} activity in adult tissue. However, heart, bone, tongue and lung expressed little to no tdTom. The lack of tdTom accumulation could have several explanations; 1) Dre^{Pr} is not expressed in these tissues in adult mice 2) Ru486 cannot reach these tissues to activate Dre^{Pr} or 3) the tdTom locus is less susceptible to recombination in these tissues. Vooijs *et al.,* have described differential recombination efficiency in different adult mouse tissues when using a *R26* targeted Cre^{Ert} mouse (Vooijs *et al*., 2001). In this report the authors find a reduced recombination efficiency of *Brca*^{flox19} and *Trp53*$^{flox2\text{-}10}$ in adult lung tissue compared to other highly recombined tissues such as kidney and spleen. In addition, a study has shown that activation of a lung specific Cre produces 50% less *R26* YFP+ reporter cells in the adult lung compared to the embryonic counterpart (Sinha and Lowell, 2017). Together this suggest that *R26* expression is decreased in some adult tissues such as the lung and/or that general delivery of inducing agents is less efficient in reaching such tissues. To fully determine that this is occurring in the RD^{Pr}/RT^{rox} animals it will be of importance to perform qRT-PCR on adult heart, bone, tongue and lung tissues to examine Dre^{Pr} mRNA expression here.

However, this immediately means, that R_Dre^{Pr} most likely cannot be used for recombination in these tissues by sub-cutaneous Ru486 administration.

In addition to not being activatable in all tissues, it is clear from the results presented here that the induction frequency of R_Dre^{Pr} is low compared to other inducible recombinases frequently used worldwide (such as Ah^{CreErt}). The highest percentage of tdTom seen in the small intestinal epithelial cells is 5% after Ru486 exposure. The low induction rate is observed 75 days after implantation of Ru486 slow release pellets (10mg/pellet). Assuming the pellets have a consistent diffusion of Ru486 from the pellet matrix over 90 days, they release approximately 0.1 mg Ru486/day/pellet. With a relatively low daily dose only a small population of cells in different tissues are expected be activated per day. This would result in only long-lived clones and newly activated cells expressing tdTom after 75 days – possibly explaining the low percentage of tdTom expression observed. To fully test the maximum induction rate of R_Dre^{Pr} it would be interesting to administer one pulse of maximum concentration Ru486 to these animals by other routes or solutions than tried so far.

The low activation rate could also suggest that the Pr fusion to Dre creates a very strong cytoplasmic retention of Dre^{Pr} and a poor nuclear translocation after Ru486 ligand binding to Pr. It has been described that nuclear import efficiency of proteins fused to the Pr receptor relies on the balance of efficiency between encoded NLSs and nuclear export signals (NESs) and the localisation properties of Pr (Kakar *et al.*, 2007). In the future, it would be interesting to model different Dre^{Pr} fusion proteins containing NL and NE signals of different strengths to optimise the translocation properties of Dre^{Pr}.

Here, R_Dre^{Pr} is reported to be a new tool for pulse-chase lineage tracing in the intestine. For such experiments it is of high importance to have minimal reporter background recombination in the absence of inducer as this would skew the clone sizes quantified. Importantly, the results here show that R_Dre^{Pr} have zero background activity. Furthermore, an induction rate of <1 cell per crypt is optimal for pulse chase lineage tracing. This is the hit rate observed 24h after Ru486 I.P

administration, with very few double-hits per crypt observed. In addition, even though the number of clones/mouse/timepoint quantified here was lower than that of Ah^{CreErt} animals the data collected seems robust with little variation. Additionally, the results show that different ways of measuring stem cell dynamics, either with Dre^{Pr} or Ah^{CreErt} concludes the same N and λ. Importantly, the N and λ reported here was in line with previous reported values (Kozar, *et al.*, 2013) thus suggesting that Dre^{Pr} activity and exposure to Ru486 does not change the behaviour of intestinal stem cells. Conclusively, it appears that Dre^{Pr} lineage tracing marks a pool of stem cells that behave in the same manner as stem cells marked by Ah^{CreErt}.

In contrast, models that rely on highly efficient recombination, for example *Villin*^{CreErt}, are often subject to significant background recombination. This makes such models good tools for whole tissue recombination but inappropriate for sporadic induced recombination. However, sporadic induced activation of cancer drivers is ideal to model cancer *in vivo*, as this recapitulates the perceived route to cancer. Thus, it seems that the Dre^{Pr} model is a good tool for low frequency, lack of background, recombination.

Lineage tracing in tumours has previously been carried out by using Cre^{Ert} induced reporter expression in spontaneous tumour models such as *Apc*^{min} (Kozar, *et al.*, 2013; Yanai *et al.*, 2017). Others have used co-expression of *Lgr5*^{CreErt} and reporters to describe the clonal origin of intestinal tumours (Schepers *et al.*, 2012). This is the first description of a recombination driven sequential model for lineage tracing in intestinal tumours. The results show that each tumour in the Ah^{CreErt}; *Ctnnb1*$^{lox(ex3)}$;RD^{Pr};RT^{rox} model contain an average of 0.92 traced clones at the end of the experimental animal life span. This underlines that single cell derived clones do survive throughout the experimental outline thus emphasising the long-term application of this tool. However, 51% of analysed tumours did not contain any clones when analysed. No clones detected could have several explanations 1) these tumours where not induced with Ru486 2) traced clone(s) were lost over the time-course 3) no stem cell activity in the tumour. In 2013 Kozar *et al.* reported that spontaneous *Apc*^{min} adenomas are maintained by 9 stem cells per tumour gland. The fact that only 1-8 clones are

detected per Ah^{CreErt};*Ctnnb1*$^{lox(ex3)}$; RD^{Pr};RT^{rox} tumour may further support the finding that adenomas are in fact maintained by a small number of stem cells. In the future, to investigate the intra-tumour clonal dynamics it will be important to analyse tumour clones at early timepoints.

Together, these results demonstrate that Dre^{Pr} can successfully be used to lineage trace in sequential mutation tumour models. In the future it would be exciting to combine Cre, Flp and Dre^{Pr} in one model to re-create the step-wise process of intestinal cancer. It could be speculated that an efficient model for instance could be: *Cyp1A1*^{FlpO};*Kras*^{fsfG12D};Ah^{CreErt};*Ctnnb1*$^{lox(ex3)}$;*R26*^{DrePr}; *R26*^{rsrtdTom}. Here, a pulse of ß-naphthoflavone would drive Flp induced intestinal *Kras*^{G12D} field characterisation. Later, a pulse of ß-naphthoflavone and Tamoxifen would induce Cre^{Ert} driven ß-catenin stabilisation and tumour growth. Lastly, once tumours had formed a dose of Ru486 would drive Dre^{Pr} induced lineage tracing to inform tumorigenic clonal dynamics.

6 Discussion and perspectives

Pre-neoplastic oncogenic fields might be a major driver of colorectal cancer initiation. Mutant clonal fields accumulate in the colonic epithelium over time and when a certain molecular threshold has been reached the primed tissue encodes the properties needed to initiate and propagate tumours (Hanahan and Weinberg, 2011).

Mutant *KRAS* fields are present in human epithelium before overt tumour formation and thus could pre-dispose to cancer development in a subset of patients (Nicholson *et al.*, 2018). Intestinal stem cell dynamics orchestrate tumour initiation and it is known that *Kras*^{G12D} changes the competitive behaviour of stem cells (Vermeulen *et al.,* 2013). Thus, *KRAS* fields could be over-represented in human colons compared to other mutations due to the stem cell advantage conferred by the oncogenic *KRAS* mutation. This can be quantified, but at the start of this project it was not clear how stem cell dynamics are defined in *KRAS* mutant fields and whether such could be of pro-oncogenic nature. Additionally, since the intestinal epithelium is plastic in the sense that committed progenitors can revert back to stem cell status, it is not clear which cells are the cell of origin for pro-oncogenic mutations that causes fields to arise or whether oncogenic mutations could change cell fate choices of progenitors.

If intestinal mutations only occur in stem cells, the total target size for oncogenic mutations would be relatively small. Such a small target population might not be large enough to explain the relatively high mutational burden found in human colonic tissues and cancers. Thus, mature cells and progenitors, that can revert back to stem cells, are in all likelihood a large source of oncogenic mutations found in intestinal stem cells.

Other studies that have investigated the cell of origin in colorectal cancer have expressed oncogenes or deleted tumour suppressors in progenitors of different lineages to assess whether such could give rise to tumours (Barker *et al.*, 2009; Schwitalla *et al.*, 2013, Tetteh *et al.*, 2016). However, to our knowledge,

the studies included in this thesis are the first to quantitatively compare stem cell contribution from early secretory progenitors with or without oncogenes expressed. Thus, these results offer novel measures of the relative cell fate choices of progenitors encoding *Kras*^{G12D}. Such measurements are made possible by the *Atoh1*^{CreErt};*Kras*^{lslG12D};*R26*^{lsltdTom} model as this mouse confers large numbers of marked clones, even during homeostatic conditions, thus allowing statistical testing of differences.

In the third chapter of this thesis, it was discovered that Kras^{G12D} expression did not change the stem cell contribution potential of early secretory progenitors during tissue homeostasis. Even though no difference in stem cell contribution was observed, it was still clear that *Kras*^{G12D} mutant secretory progenitors make a significant contribution to self-renewal during tissue homeostasis. However, after Lgr5 depletion, Kras^{G12D} expression increased the stem cell contribution potential of Atoh1-expressing cells in the small intestine. Therefore, it appears that in a certain context there is an enrichment or higher influx rate of *Kras*^{G12D} mutated progenitors into the stem cell pool, which indicates a greater role for progenitors as sources of oncogenic mutations in such cases. Furthermore, the results show that *Kras*^{G12D} stem cells derived from Atoh1+ cells have an advantage over their WT neighbours and that *Kras*^{G12D} mutant crypts derived from Atoh1+ cells can spread via fission, both in the small intestine and colon.

Increased stem cell contribution of *Kras*^{G12D} progenitors infers an advantageous bias for such oncogenes to populate the stem cell pool. With the additional *Kras*^{G12D} stem cell competitive bias these two mechanisms acting in sequence gives *Kras*^{G12D} oncogenic mutations extremely beneficial circumstances for clonal fixation and field initiation. Thus, oncogene induced plasticity offers an additional layer of mutation-enabled bias. Additional studies into this phenomenon will be important for understanding colorectal cancer initiation.

Furthermore, this thesis is the first study that determines stem cell bias and mutant crypt spread of oncogenes derived from progenitors in the same mouse model. To further explore Atoh1+ derived *Kras*^{G12D} crypts, it could be envisioned that the *Atoh1*^{CreErt};*Kras*^{lslG12D}; *R26*^{lsltdTom} model could also be used to

asses stem cell dynamics within Atoh1 derived *Kras*^{G12D} fields. To this end, a reporter encoding a different colour would be useful, for instance *R26*^{lslYFP} (Madisen *et al.*, 2010). The *Atoh1*^{CreErt};*Kras*^{lslG12D};*R26*^{lslYFP} model could be crossed to the R_Dre^{Pr};R_tdTom^{rox} developed in chapter five of this thesis. Within this model, Atoh1+ derived YFP+Kras^{G12D} clones could be left to develop into fields over time, and Dre^{Pr} could then be used to lineage trace YFP+/tdTom+ cells within *Kras*^{G12D} fields to determine stem cells dynamics of Atoh1+ derived *Kras*^{G12D} mutant crypts. This would also allow the direct comparison of stem cell dynamics in WT crypts neighbouring such fields.

No studies to date have attempted to define stem cell dynamics within pro-oncogenic fields *in vivo*. This is possibly because the community have been lacking the right model(s) to do so. Thus, the development of the *Cyp1A1*^{FlpO};*Kras*^{fsfG12D};*Ah*^{CreErt}; *R26*^{lsltdTom} model and the findings presented here are inherently novel. Using this model, it is shown that stem cell dynamics are accelerated in *Kras*^{G12D} fields. We find that stem cell fixation is achieved 2-fold faster in *Kras*^{G12D} fields by employing fewer clonogenic stem cells per crypt with a relatively increased replacement rate for neutral drift. Accelerated clonal dynamics are shown to be of pro-oncogenic nature, as new mutations are fixed faster in *Kras*^{G12D} fields compared to WT fields. This could speed up cancer initiation as the crypts in such fields could reach the molecular threshold for cancer initiation faster. In addition, the results show that the accelerated *Kras*^{G12D} clonal dynamics are Mek1/2 dependent and that Mek1/2 inhibitor treatment can slow them down, thereby offering an exciting possibility for chemoprevention.

As discussed in chapter 4, if the increased fixation rate observed in *Kras*^{G12D} fields is translated directly into the human setting, the fixation time within *KRAS* fields would be decreased to approximately 3 years. However, as human stem cell dynamics are much slower than in the mouse, it is interesting to speculate whether the dynamic range for increased fixation rate could be even greater here? If that would be the case, clonal dynamics in *KRAS* fields could potentially be accelerated more than 2-fold in humans.

In chapter 4, stem cell dynamics are defined under *Kras*^{G12D} homeostatic tissue conditions. However, as discussed in chapter 1, *Kras*^{G12D}

fields are readily found in patients exposed to recurrent inflammation (Leedam *et al.*, 2009; Galandiuk *et al.*, 2012). Thus, to further develop the understanding of pro-oncogenic abilities of *Kras*^{G12D} fields, the *Cyp1A1*^{FlpO};*Kras*^{fsfG12D};*Ah*^{CreErt}; *R26*^{lsltdTom} model could be exposed to inflammatory agents such as DSS. In such a setting, stem cell dynamics in *Kras*^{G12D} fields could be defined during inflammation. In addition, lifestyle, such as obesity, offers one of the biggest risk factors for colorectal cancer development. To this end, it has been shown that *Apc*^{min} mice on western diet have increased tumour initiation capacity, compared to mice on normal diet (Baltgalvis *et al.*, 2009). Thus *Cyp1A1*^{FlpO};*Kras*^{fsfG12D}; *Ah*^{CreErt};*R26*^{lsltdTom} mice fed a western diet could be used to study how obesity affects stem cell dynamics in *Kras*^{G12D} fields. Collectively, such studies could aid the understanding of why certain patient groups have an increased tumour incidence.

This study concludes that *Kras*^{G12D} fields speed up cancer initiation time since new mutations are fixed faster here. Additionally, it is clear that *Kras*^{G12D} fields alone cannot initiate cancer formation and that multiple other mutations are needed. It is entirely possible that some oncogenic mutations will be positively biased in *Kras*^{G12D} fields, which would make such combinations additive in their pro-oncogenic properties. To measure such an effect, the *Cyp1A1*^{FlpO};*Kras*^{fsfG12D};*Ah*^{CreErt};*R26*^{lsltdTom} model could be crossed to other alleles such as *Tp53*^{lslR172H} (Goh *et al.*, 2015). In such an enhanced model, it would be possible to define the probability of replacement of *Kras*$^{G12D+}$/*p53*$^{R172H+}$ stem cells over *Kras*^{G12D} stem cells.

In addition, it is also possible that oncogenic fields encoding several mutations are present in the human epithelium. To study the stem cell dynamics in such fields *Cyp1A1*^{FlpO};*Kras*^{fsfG12D}; *Ah*^{CreErt};*R26*^{lsltdTom} mice could be crossed to for instance *Tp53*^{frt} (Lee *et al.*, 2012). Here, Flp induction would create intestinal fields of *Kras*^{G12D} and deleted *Tp53* and lineage tracing could define stem cell dynamics within *Kras*^{G12D} and deleted *Tp53* fields. Such studies could shed light on additive pro-oncogenic effects *in vivo*.

Furthermore, results in this thesis mainly focus on studying the effect of the oncogenic mutation *Kras*^{G12D}. However, several different oncogenic *KRAS*

mutations can be found in the human colonic epithelium, such as the *KRAS*^{G12V} variant (Nicholson *et al.*, 2018). In addition, there are indications that these mutations confer different abilities to the intestinal tissue (Sansom *et al.*, 2006; Haigis *et al.*, 2008; Feng *et al.,* 2011). Additionally, oncogenic mutations in other genes such as *STAG2*, *PIK3CA* and *ERB2/3* have also been shown to be present in human colonic tissue prior to cancer development (Lee-Six *et al.*, 2018). In the future, it will be important to determine whether these and other colorectal cancer drivers such as mutations in *APC, TP53, FBXW7, PTEN* and *SMAD4* are present in pre-neoplastic fields in human colons. Additionally, defining the stem cell dynamics in fields of different mutations *in vivo* could map the relative cancer-priming abilities of each mutation. This could help classify high-risk patients that should be on increased observation in any future screening procedures for colorectal cancer.

Lastly, development of the inducible Dre^{Pr} and FlpO alleles introduced in the fourth and fifth chapter of this thesis greatly enhances novel modelling of colorectal cancer *in vivo*. So-far most colorectal cancer tumour models consist of inducing the loss of tumour suppressor genes or gain of oncogenes at the same time (Jackstadt and Sansom, 2016). The addition of the new recombination tools R_Dre^{Pr} and *Cyp1A1*^{FlpO} makes it possible to create models that can temporally separate up to three individual recombination events *in vivo* (Cre, Flp and Dre). This makes it achievable to create models where, for instance, it can be tested how the sequence of different genetic alterations affects intestinal tumour development.

In conclusion, the results in this thesis present a thorough characterisation of *Kras*^{G12D} field initiation and the pro-oncogenic features within such fields. This study offers a new way to think about pro-oncogenic abilities of different genetic alterations. Here, the additive effect of progenitor stem cell contribution, stem cell bias, crypt spread and stem cell dynamics within fields characterise the overall pro-oncogenic features of *Kras*^{G12D} (Figure 6-1A). Based on the findings presented in this thesis, a model for *Kras*^{G12D} field development and their pro-oncogenic characteristics is presented in Figure 6-1B. Here, *Kras*^{G12D} fields can

arise from mutant secretory progenitors that contribute to a normal extent during homeostasis and in an increased manner after Lgr5 depletion. Clonal victory and crypt fission can then create *Kras*^{G12D} fields and once present, these fields are able to fix subsequent oncogenic mutations at an accelerated rate. Tantalisingly, this increased fixation can be repressed by Mek1/2 inhibition.

7 References

Adjei, A. A. *et al.* (2008) 'Phase I Pharmacokinetic and Pharmacodynamic Study of the Oral, Small-Molecule Mitogen-Activated Protein Kinase Kinase 1/2 Inhibitor AZD6244 (ARRY-142886) in Patients With Advanced Cancers', *J Clin Oncol*, 26(13), pp. 2139–2146. doi: 10.1200/JCO.2007.14.4956.

Alexandrov, L. B. *et al.* (2015) 'Clock-like mutational processes in human somatic cells', *Nature Genetics*. Nature Publishing Group, 47(12), pp. 1402–1407. doi: 10.1038/ng.3441.

Anastassiadis, K. *et al.* (2009) 'Dre recombinase, like Cre, is a highly efficient site-specific recombinase in E. coli, mammalian cells and mice.', *Disease models & mechanisms*, 2(9–10), pp. 508–15. doi: 10.1242/dmm.003087.

Andersen, M. S. *et al.* (2019) 'Tracing the cellular dynamics of sebaceous gland development in normal and perturbed states', *Nature Cell Biology*. Nature Publishing Group, 21(8), pp. 924–932. doi: 10.1038/s41556-019-0362-x.

Ashton, G. H. *et al.* (2010) 'Focal adhesion kinase is required for intestinal regeneration and tumorigenesis downstream of Wnt/c-Myc signaling.', *Developmental cell*. Elsevier, 19(2), pp. 259–69. doi: 10.1016/j.devcel.2010.07.015.

Baker, A.-M. *et al.* (2014) 'Quantification of Crypt and Stem Cell Evolution in the Normal and Neoplastic Human Colon', *Cell Reports*. Cell Press, 8(4), pp. 940–947. doi: 10.1016/J.CELREP.2014.07.019.

Baker, K. T. *et al.* (2018) 'Precancer in ulcerative colitis: the role of the field effect and its clinical implications', *Carcinogenesis*, 39(1), pp. 11–20. doi: 10.1093/carcin/bgx117.

Baltgalvis, K. A. *et al.* (2009) 'The Interaction of a High-Fat Diet and Regular Moderate Intensity Exercise on Intestinal Polyp Development in Apc Min/+ Mice'. doi: 10.1158/1940-6207.CAPR-09-0017.

Banerjee, A. *et al.* (2018) 'Interpreting heterogeneity in intestinal tuft cell structure and function', *The Journal of Clinical Investigation*. American Society for Clinical

Investigation, 128(5), pp. 1711–1719. doi: 10.1172/JCI120330.

Barker, N. *et al.* (2007) 'Identification of stem cells in small intestine and colon by marker gene Lgr5.', *Nature*, 449(7165), pp. 1003–7. doi: 10.1038/nature06196.

Barker, N. *et al.* (2009) 'Crypt stem cells as the cells-of-origin of intestinal cancer', *Nature*. Nature Publishing Group, 457(7229), pp. 608–611. doi: 10.1038/nature07602.

Beumer, J. and Clevers, H. (2016) 'Regulation and plasticity of intestinal stem cells during homeostasis and regeneration'. doi: 10.1242/dev.133132.

Bruens, L. *et al.* (2017) 'In Vivo Imaging Reveals Existence of Crypt Fission and Fusion in Adult Mouse Intestine', *Gastroenterology*. W.B. Saunders, 153(3), pp. 674-677.e3. doi: 10.1053/J.GASTRO.2017.05.019.

Buczacki, S. J. a *et al.* (2013) 'Intestinal label-retaining cells are secretory precursors expressing Lgr5.', *Nature*. Nature Publishing Group, 495(7439), pp. 65–9. doi: 10.1038/nature11965.

Buffet, C. *et al.* (2017) 'DUSP5 and DUSP6, two ERK specific phosphatases, are markers of a higher MAPK signaling activation in BRAF mutated thyroid cancers', *PLOS ONE*. Edited by J.-I. Park. Public Library of Science, 12(9), p. e0184861. doi: 10.1371/journal.pone.0184861.

Castillo-Azofeifa, D. *et al.* (2019) 'Atoh1+ secretory progenitors possess renewal capacity independent of Lgr5+ cells during colonic regeneration', *The EMBO Journal*. EMBO Press, 38(4), p. e99984. doi: 10.15252/embj.201899984.

Chen, R. *et al.* (2005) 'The initiation of colon cancer in a chronic inflammatory setting', *Carcinogenesis*. Narnia, 26(9), pp. 1513–1519. doi: 10.1093/carcin/bgi106.

Cheng, H. (1974c) 'Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. IV. Paneth cells', *Am J Anat*, 141(4), pp. 521–535. doi: 10.1002/aja.1001410406.

Cheng, H. and Leblond, C.P. (1974a) 'Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. III. Entero-endocrine cells', *American Journal of Anatomy*, 141(4), pp. 503–19. doi: 10.1002/aja.1001410405.

Cheng, H. and Leblond, C.P. (1974b) ‘Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine II. Mucous cells’, *American Journal of Anatomy*. Wiley Subscription Services, Inc., A Wiley Company, 141(4), pp. 481–501. doi: 10.1002/aja.1001410404.

Cheng, H. and Leblond, C P (1974) ‘Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine V. Unitarian theory of the origin of the four epithelial cell types’, *American Journal of Anatomy*. Wiley Subscription Services, Inc., A Wiley Company, 141(4), pp. 537–561. doi: 10.1002/aja.1001410407.

Cheung, A. F. *et al.* (2009) ‘Complete deletion of Apc results in severe polyposis in mice’, *Oncogene*. Nature Publishing Group, 29(12), pp. 1857–1864. doi: 10.1038/onc.2009.457.

Clarke, R. M. (1972) ‘The effect of growth and of fasting on the number of villi and crypts in the small intestine of the albino rat’, *J. Anat*, 112(1), pp. 27–33.

Cox, A. D. *et al.* (2014) ‘Drugging the undruggable RAS: Mission Possible?’ doi: 10.1038/nrd4389.

Cox, A. D. and Der, C. J. (2010) ‘Ras history, Small GTPases’, *Taylor and Francis Online*, 1:1, pp. 2–27. doi: 10.4161/sgtp.1.1.12178.

Darwich, A. S. *et al.* (2014) ‘Meta-Analysis of the Turnover of Intestinal Epithelia in Preclinical Animal Species and Humans’, *Drug Metabolism and Disposition*, 42(12), pp. 2016–2022. doi: 10.1124/dmd.114.058404.

Deckx, R. J., Vantrappen, G. R. and Parein, M. M. (1967) ‘Localization of lysozyme activity in a Paneth cell granule fraction’, *BBA - Enzymology*. doi: 10.1016/0005-2744(67)90136-2.

Dieterle, C. P. *et al.* (2004) *Detection of Isolated Tumor Cells by Polymerase Chain Reaction-Restriction Fragment Length Polymorphism for K-ras Mutations in Tissue Samples of 199 Colorectal Cancer Patients*. doi: 10.1158/1078-0432.CCR-1355-02.

van Es, J. H. van *et al.* (2012a) ‘A Critical Role for the Wnt Effector Tcf4 in Adult Intestinal Homeostatic Self-Renewal’, *Molecular and Cellular Biology*. American Society for Microbiology Journals, 32(10), pp. 1918–1927. doi:

10.1128/MCB.06288-11.

van Es, J. H. *et al.* (2012b) 'Dll1+ secretory progenitor cells revert to stem cells upon crypt damage', *Nature Cell Biology*. Nature Publishing Group, 14(10), pp. 1099–1104. doi: 10.1038/ncb2581.

van Es, J. H., van Gijn, M. E., *et al.* (2005a) 'Notch/γ-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells', *Nature*. Nature Publishing Group, 435(7044), pp. 959–963. doi: 10.1038/nature03659.

van Es, J. H., Jay, P., *et al.* (2005b) 'Wnt signalling induces maturation of Paneth cells in intestinal crypts', *NATURE CELL BIOLOGY*, 7(4). doi: 10.1038/ncb1240.

Escobar, M. *et al.* (2011) 'Intestinal epithelial stem cells do not protect their genome by asymmetric chromosome segregation', *Nature Communications*, 2, p. 258. doi: 10.1038/ncomms1260.

Farin, H. F. *et al.* (2016) 'Visualization of a short-range Wnt gradient in the intestinal stem-cell niche', *Nature*. doi: 10.1038/nature16937.

Fearon, E., Hamilton, S. R. and Vogelstein, B. (1987) 'Clonal Analysis of Human Colorectal Tumors', *Science*. American Association for the Advancement of Science, 238(4824), pp. 193–197. doi: 10.1126/science.2889267.

Fearon, E. R. (2011) 'Molecular genetics of colorectal cancer.', *Annual review of pathology*, 6, pp. 479–507. doi: 10.1146/annurev-pathol-011110-130235.

Fearon, E. R. and Vogelstein, B. (1990) 'A genetic model for colorectal tumorigenesis', *Cell*, 61(5), pp. 759–767. doi: 10.1016/0092-8674(90)90186-I.

Feng, Y. *et al.* (2011) 'Mutant KRAS promotes hyperplasia and alters differentiation in the colon epithelium but does not expand the presumptive stem cell pool.', *Gastroenterology*. NIH Public Access, 141(3), pp. 1003-1013.e1–10. doi: 10.1053/j.gastro.2011.05.007.

van der Flier, L. G. *et al.* (2009) 'Transcription Factor Achaete Scute-Like 2 Controls Intestinal Stem Cell Fate', *Cell*. Cell Press, 136(5), pp. 903–912. doi: 10.1016/J.CELL.2009.01.031.

Gagné-Sansfaçon, J. *et al.* (2016) 'SHP-2 phosphatase contributes to KRAS-driven intestinal oncogenesis but prevents colitis-associated cancer development.', *Oncotarget*. Impact Journals, LLC, 7(40), pp. 65676–65695. doi:

10.18632/oncotarget.11601.

Galandiuk, S. *et al.* (2012) 'Field Cancerization in the Intestinal Epithelium of Patients With Crohn's Ileocolitis', *YGAST*, 142, pp. 855-864.e8. doi: 10.1053/j.gastro.2011.12.004.

Gehart, H. and Clevers, H. (2019) 'Tales from the crypt: new insights into intestinal stem cells', *Nature Reviews Gastroenterology & Hepatology*. Nature Publishing Group, 16(1), pp. 19–34. doi: 10.1038/s41575-018-0081-y.

Gerbe, F. *et al.* (2016) 'Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites', *Nature*. Nature Publishing Group, 529(7585), pp. 226–230. doi: 10.1038/nature16527.

Goh, A. M. *et al.* (2015) 'Mutant p53 accumulates in cycling and proliferating cells in the normal tissues of p53 R172H mutant mice', *Oncotarget*, 6(20). doi: 10.18632/oncotarget.4956.

Greaves, L. C. *et al.* (2006) 'Mitochondrial DNA mutations are established in human colonic stem cells, and mutated clones expand by crypt fission', *PNAS*, 103(3), pp. 714–719. doi: 10.1073/pnas.0505903103.

Gribble, F. M. and Reimann, F. (2019) 'Function and mechanisms of enteroendocrine cells and gut hormones in metabolism', *Nature Reviews Endocrinology*, 15, pp. 226–237. doi: 10.1038/s41574-019-0168-8.

Griffiths, D. F. *et al.* (1989) 'The clonal origin of experimental large bowel tumours.', *British journal of cancer*. Nature Publishing Group, 59(3), pp. 385–7. doi: 10.1038/bjc.1989.77.

Griffiths, D. F. R. *et al.* (1988) 'Demonstration of somatic mutation and colonic crypt clonality by X-linked enzyme histochemistry', *Nature*. Nature Publishing Group, 333(6172), pp. 461–463. doi: 10.1038/333461a0.

Groszer, M. *et al.* (2001) 'Negative Regulation of Neural Stem/Progenitor Cell Proliferation by the Pten Tumor Suppressor Gene in Vivo', *Science*. American Association for the Advancement of Science, 294(5549), pp. 2186–2189. doi: 10.1126/science.1553558.

Guerra, C. *et al.* (2003) 'Tumor induction by an endogenous K-ras oncogene is highly dependent on cellular context', *Cancer Cell*. Elsevier, 4(2), pp. 111–120.

doi: 10.1016/S1535-6108(03)00191-0.

Guinney, J. *et al.* (2015) 'The consensus molecular subtypes of colorectal cancer', *Nature Medicine*, 21(11), pp. 1350–1356. doi: 10.1038/nm.3967.

Gutié Rrez-Vá Zquez, C. and Quintana, F. J. (2018) 'Regulation of the Immune Response by the Aryl Hydrocarbon Receptor'. doi: 10.1016/j.immuni.2017.12.012.

Haigis, K. M. *et al.* (2008) 'Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon', *Nature Genetics*, 40(5), pp. 600–608. doi: 10.1038/ng.115.

Hall, P. A. *et al.* (1994) 'Regulation of cell number in the mammalian gastrointestinal tract: the importance of apoptosis.', *Journal of Cell Science*, 107(1), pp. 3569–3577.

Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of cancer: The next generation', *Cell*, pp. 646–674. doi: 10.1016/j.cell.2011.02.013.

Harada, N. *et al.* (1999) *Intestinal polyposis in mice with a dominant stable mutation of the β-catenin gene*, *The EMBO Journal*. doi: 10.1093/emboj/18.21.5931.

Hermann, M. *et al.* (2014) 'Binary recombinase systems for high-resolution conditional mutagenesis', *Nucleic Acids Research*, 42(6), pp. 3894–3907. doi: 10.1093/nar/gkt1361.

Heuberger, J. *et al.* (2014) 'Shp2/MAPK signaling controls goblet/paneth cell fate decisions in the intestine', *PNAS*, 111(9), pp. 3472–3477. doi: 10.1073/pnas.1309342111.

Hobbs, G. A., Der, C. J. and Rossman, K. L. (2016) 'RAS isoforms and mutations in cancer at a glance.', *Journal of cell science*. The Company of Biologists Ltd, 129(7), pp. 1287–92. doi: 10.1242/jcs.182873.

Hochster, H. S. *et al.* (2015) 'Phase II study of selumetinib (AZD6244, ARRY-142886) plus irinotecan as second-line therapy in patients with K-RAS mutated colorectal cancer', *Cancer Chemotherapy and Pharmacology*. Springer Berlin Heidelberg, 75(1), pp. 17–23. doi: 10.1007/s00280-014-2609-3.

Howitt, M. R. *et al.* (2016) 'Tuft cells, taste-chemosensory cells, orchestrate

parasite type 2 immunity in the gut.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 351(6279), pp. 1329–33. doi: 10.1126/science.aaf1648.

Ireland, H. *et al.* (2004) 'Inducible cre-mediated control of gene expression in the murine gastrointestinal tract: effect of loss of $β$-catenin', *Gastroenterology*, 126(5), pp. 1236–1246. doi: 10.1053/j.gastro.2004.03.020.

Ishibashi, F. *et al.* (2018) 'Contribution of ATOH1+ Cells to the Homeostasis, Repair, and Tumorigenesis of the Colonic Epithelium', *Stem Cell Reports*. Cell Press, 10(1), pp. 27–42. doi: 10.1016/J.STEMCR.2017.11.006.

Issa, I. A. and Noureddine, M. (2017) 'Colorectal cancer screening: An updated review of the available options.', *World journal of gastroenterology*. Baishideng Publishing Group Inc, 23(28), pp. 5086–5096. doi: 10.3748/wjg.v23.i28.5086.

Itzkovitz, S. *et al.* (2012) 'Single-molecule transcript counting of stem-cell markers in the mouse intestine', *Nature Cell Biology*, 14(1), pp. 106–114. doi: 10.1038/ncb2384.

Jackson, E. L. *et al.* (2001) 'Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras.', *Genes & development*. Cold Spring Harbor Laboratory Press, 15(24), pp. 3243–8. doi: 10.1101/gad.943001.

Jackstadt, R. and Sansom, O. J. (2016) 'Mouse models of intestinal cancer', *The Journal of Pathology*. John Wiley & Sons, Ltd, 238(2), pp. 141–151. doi: 10.1002/path.4645.

Janssen, K.-P. *et al.* (2006) 'APC and oncogenic KRAS are synergistic in enhancing Wnt signaling in intestinal tumor formation and progression.', *Gastroenterology*, 131(4), pp. 1096–109. doi: 10.1053/j.gastro.2006.08.011.

Janssen, K. *et al.* (2002) 'Targeted expression of oncogenic K-ras in intestinal epithelium causes spontaneous tumorigenesis in mice', *Gastroenterology*. W.B. Saunders, 123(2), pp. 492–504. doi: 10.1053/GAST.2002.34786.

Jass, J. R. (2004) 'Hyperplastic polyps and colorectal cancer: is there a link?', *Clinical Gastroenterology and Hepatology*. W.B. Saunders, 2(1), pp. 1–8. doi: 10.1016/S1542-3565(03)00284-2.

Jenny, M. *et al.* (2002) 'Neurogenin3 is differentially required for endocrine cell

fate specification in the intestinal and gastric epithelium', *EMBO Journal*, 21(23), pp. 6338–6347. doi: 10.1093/emboj/cdf649.

Juárez, M. *et al.* (2017) 'KRAS and BRAF somatic mutations in colonic polyps and the risk of metachronous neoplasia.', *PloS one*. Public Library of Science, 12(9), p. e0184937. doi: 10.1371/journal.pone.0184937.

Kakar, M. *et al.* (2007) 'Optimizing the protein switch: Altering nuclear import and export signals, and ligand binding domain', *Journal of Controlled Release*. Elsevier, 120(3), pp. 220–232. doi: 10.1016/J.JCONREL.2007.04.017.

Keefe, D. M. K. *et al.* (2000) 'Chemotherapy for cancer causes apoptosis that precedes hypoplasia in crypts of the small intestine in humans', *Gut*, 47, pp. 632–637. doi: 10.1136/gut.47.5.632.

Kellendonk, C. *et al.* (1996) 'Regulation of Cre recombinase activity by the synthetic steroid RU 486', *Nucleic Acids Research*, 24(8), pp. 1404–1411. doi: 10.1093/nar/24.8.1404.

Kemp, R. *et al.* (2004) 'Elimination of background recombination: somatic induction of Cre by combined transcriptional regulation and hormone binding affinity.', *Nucleic acids research*, 32(11), p. e92. doi: 10.1093/nar/gnh090.

Kemper, A. C. and Specian, R. D. (1991) 'Rat small intestinal mucins: A quantitative analysis', *The Anatomical Record*. Wiley Subscription Services, Inc., A Wiley Company, 229(2), pp. 219–226. doi: 10.1002/ar.1092290209.

Kim, T.-H., Escudero, S. and Shivdasani, R. A. (2012) 'Intact function of Lgr5 receptor-expressing intestinal stem cells in the absence of Paneth cells', *PNAS*, 108(10), pp. 3932–3937. doi: 10.1073/pnas.1113890109.

Kisseberth, W. C. *et al.* (1999) 'Ubiquitous Expression of Marker Transgenes in Mice and Rats', *Developmental Biology*, 214(1), pp. 128–138. doi: 10.1006/dbio.1999.9417.

Korinek, V. *et al.* (1998) *Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4*. doi: 10.1038/1270.

Kozar, S., Morrissey, E., Nicholson, A. M., van der Heijden, M., *et al.* (2013) 'Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas.', *Cell stem cell*, 13(5), pp. 626–33. doi:

10.1016/j.stem.2013.08.001.

Kozar, S., Morrissey, E., Nicholson, A. M., Van Der Heijden, M., *et al.* (2013) 'Continuous Clonal Labeling Reveals Small Numbers of Functional Stem Cells in Intestinal Crypts and Adenomas', *Cell Stem Cell*, 13, pp. 626–633. doi: 10.1016/j.stem.2013.08.001.

Kraus, M. *et al.* (2006) 'The balanced induction of K-ras codon 12 and 13 mutations in mucosa differs from their ratio in neoplastic tissues', *International Journal of Oncology*. Spandidos Publications, 29(4), pp. 957–964. doi: 10.3892/ijo.29.4.957.

Kretzschmar, K. and Watt, F. M. (2012) 'Lineage tracing', *Cell*. Elsevier Inc., 148(1–2), pp. 33–45. doi: 10.1016/j.cell.2012.01.002.

Krishnamurthy, A. *et al.* (2018) 'Phase Ib Results of the Rational Combination of Selumetinib and Cyclosporin A in Advanced Solid Tumors with an Expansion Cohort in Metastatic Colorectal Cancer', *Cancer Research*. American Association for Cancer Research, 78(18), pp. 5398–5407. doi: 10.1158/0008-5472.CAN-18-0316.

Kuhnert, F. *et al.* (2004) 'Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 101(1), pp. 266–71. doi: 10.1073/pnas.2536800100.

Kuipers, E. J. *et al.* (2015) 'Colorectal cancer'. doi: 10.1038/nrdp.2015.65.

Lakso, M. *et al.* (1992) *Targeted oncogene activation by site-specific recombination in transgenic mice (crellox/lens development/sinian virus 40 large tumor antigen)*, *Proc. Nail. Acad. Sci. USA*. doi: 10.1073/pnas.89.14.6232.

Land, H., Parada, L. F. and Weinberg, R. A. (1983) 'Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes', *Nature*. Nature Publishing Group, 304(5927), pp. 596–602. doi: 10.1038/304596a0.

Leblond, C. P. and Stevens, C. E. (1948) 'The constant renewal of the intestinal epithelium in the albino rat', *The Anatomical Record*. John Wiley & Sons, Ltd, 100(3), pp. 357–377. doi: 10.1002/ar.1091000306.

Lee-Six, H. *et al.* (2018) 'The landscape of somatic mutation in normal colorectal epithelial cells', *bioRxiv*, p. 416800. doi: https://doi.org/10.1101/416800.

Lee, C.-L. *et al.* (2012) 'Generation of primary tumors with Flp recombinase in FRT-flanked p53 mice.', *Disease models & mechanisms*. The Company of Biologists Ltd, 5(3), pp. 397–402. doi: 10.1242/dmm.009084.

Leedham, S. J. *et al.* (2009) 'Clonality, Founder Mutations, and Field Cancerization in Human Ulcerative Colitis–Associated Neoplasia', *Gastroenterology*. W.B. Saunders, 136(2), pp. 542-550.e6. doi: 10.1053/J.GASTRO.2008.10.086.

Lemieux, E. *et al.* (2015) 'Oncogenic KRAS signalling promotes the Wnt/β-catenin pathway through LRP6 in colorectal cancer', *Oncogene*. Nature Publishing Group, 34(38), pp. 4914–4927. doi: 10.1038/onc.2014.416.

Levayer, R. and Moreno, E. (2016) 'Communicative & Integrative Biology How to be in a good shape? The influence of clone morphology on cell competition'. doi: 10.1080/19420889.2015.1102806.

Li, H. J. *et al.* (2012) 'Notch signaling differentially regulates the cell fate of early endocrine precursor cells and their maturing descendants in the mouse pancreas and intestine', *Developmental Biology*. Academic Press, 371(2), pp. 156–169. doi: 10.1016/J.YDBIO.2012.08.023.

Lopez-Garcia, C. *et al.* (2010) 'Intestinal stem cell replacement follows a pattern of neutral drift.', *Science (New York, N.Y.)*, 330(6005), pp. 822–5. doi: 10.1126/science.1196236.

Lorenz, R. G. and Newberry, R. D. (2004) 'Isolated Lymphoid Follicles Can Function as Sites for Induction of Mucosal Immune Responses', *Annals of the New York Academy of Sciences*. John Wiley & Sons, Ltd (10.1111), 1029(1), pp. 44–57. doi: 10.1196/annals.1309.006.

Lugli, N. *et al.* (2017) 'Enhanced Rate of Acquisition of Point Mutations in Mouse Intestinal Adenomas Compared to Normal Tissue', *Cell Reports*, 19. doi: 10.1016/j.celrep.2017.05.051.

Madisen, L. *et al.* (2010) 'A robust and high-throughput Cre reporting and characterization system for the whole mouse brain.', *Nature neuroscience*.

Nature Publishing Group, 13(1), pp. 133–40. doi: 10.1038/nn.2467.

Madisen, L. *et al.* (2015) 'Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance.', *Neuron*. Elsevier, 85(5), pp. 942–58. doi: 10.1016/j.neuron.2015.02.022.

Marino, S. *et al.* (2000) 'Induction of medulloblastomas in p53-null mutant mice by somatic inactivation of Rb in the external granular layer cells of the cerebellum.', *Genes & development*. Cold Spring Harbor Laboratory Press, 14(8), pp. 994–1004. doi: 10.1101/GAD.14.8.994.

El Marjou, F. *et al.* (2004) 'Tissue-specific and inducible Cre-mediated recombination in the gut epithelium', *Genesis*, 39(3), pp. 186–193. doi: 10.1002/gene.20042.

Martincorena, I. *et al.* (2015) 'Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 348(6237), pp. 880–6. doi: 10.1126/science.aaa6806.

Martincorena, I. *et al.* (2018) 'Somatic mutant clones colonize the human esophagus with age.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, p. eaau3879. doi: 10.1126/science.aau3879.

Matsushitasq, N. *et al.* (1993) *A Factor Binding to the Xenobiotic Responsive Element (XRE) of P-4501A1 Gene Consists of at least Two Helix-Loop-Helix Proteins, Ah Receptor and Arnt**.

Maunoury, R. *et al.* (1992) 'Developmental regulation of villin gene expression in the epithelial cell lineages of mouse digestive and urogenital tracts.', *Development (Cambridge, England)*, 7(115), pp. 717–28.

Merritt, A. J., Gould, K. A. and Dove, W. F. (1997) *Polyclonal structure of intestinal adenomas in Apc Min mice with concomitant loss of Apc from all tumor lineages*, *Medical Sciences Communicated by Eric S. Lander*. doi: 10.1073/pnas.94.25.13927.

Milano, J. *et al.* (2004) 'Modulation of Notch Processing by g-Secretase Inhibitors Causes Intestinal Goblet Cell Metaplasia and Induction of Genes Known to Specify Gut Secretory Lineage Differentiation', *TOXICOLOGICAL SCIENCES*,

82, pp. 341–358. doi: 10.1093/toxsci/kfh254.

von Moltke, J. *et al.* (2016) 'Tuft-cell-derived IL-25 regulates an intestinal ILC2–epithelial response circuit', *Nature*. Nature Publishing Group, 529(7585), pp. 221–225. doi: 10.1038/nature16161.

Montgomery, R. K. *et al.* (2011) 'Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells.', *Proceedings of the National Academy of Sciences of the United States of America*, 108(1), pp. 179–84. doi: 10.1073/pnas.1013004108.

Moon, B.-S. *et al.* (2013) 'role of Oncogenic K-ras in cancer Stem cell Activation by Aberrant Wnt/β-catenin Signaling', *JNCI: Journal of the National Cancer Institute*, 106(2). doi: 10.1093/jnci/djt373.

Moon, J.-H. *et al.* (2019) 'Targeting Î²-catenin overcomes MEK inhibition resistance in colon cancer with KRAS and PIK3CA mutations', *Cellular and Molecular Biology British Journal of Cancer*, p. 120. doi: 10.1038/s41416-019-0434-5.

Moser, A. R., Pitot, H. C. and Dove, W. F. (1990) 'A Dominant Mutation That Predisposes to Multiple Intestinal Neoplasia in the Mouse', *Science*. American Association for the Advancement of Science, 247(4940), pp. 322–324. doi: 10.1126/science.2296722.

Motojima, M. *et al.* (2016) 'Conditional knockout of Foxc2 gene in kidney: efficient generation of conditional alleles of single-exon gene by double-selection system', *Mammalian Genome*, 27, pp. 62–69. doi: 10.1007/s00335-015-9610-y.

Muñoz, J. *et al.* (2012) 'The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent "+4" cell markers.', *The EMBO journal*, 31(14), pp. 3079–91. doi: 10.1038/emboj.2012.166.

Nicholson, A. M. *et al.* (2018) 'Fixation and Spread of Somatic Mutations in Adult Human Colonic Epithelium', *Cell Stem Cell*. Elsevier, 0(0). doi: 10.1016/j.stem.2018.04.020.

Noah, T. K., Donahue, B. and Shroyer, N. F. (2011) 'Intestinal development and differentiation', *Experimental Cell Research*. Academic Press, 317(19), pp. 2702–2710. doi: 10.1016/J.YEXCR.2011.09.006.

Nojadeh, J. N., Behrouz Sharif, S. and Sakhinia, E. (2018) 'Microsatellite instability in colorectal cancer.', *EXCLI journal*. Leibniz Research Centre for Working Environment and Human Factors, 17, pp. 159–168. doi: 10.17179/excli2017-948.

Novelli, M. R. *et al.* (1997) 'Polyclonal Origin of Colonic Adenomas in an XO/XY Patient with FAP', *Science*. American Association for the Advancement of Science, 272(5265), pp. 1187–1190. doi: 10.1126/science.272.5265.1187.

Park, J. T. and Leach, S. D. (2013) 'TAILOR: transgene activation and inactivation using lox and rox in zebrafish.', *PloS one*. Public Library of Science, 8(12), p. e85218. doi: 10.1371/journal.pone.0085218.

Parsons, B. L. *et al.* (2010) 'ACB-PCR Quantification of K-RAS Codon 12 GAT and GTT Mutant Fraction in Colon Tumor and Non-Tumor Tissue', *Cancer Investigation*, (28), pp. 364–375. doi: 10.1080/07357901003630975.

Patel, A. *et al.* (2015) 'Field cancerisation in colorectal cancer: A new frontier or pastures past?', *World J Gastroenterol J Gastroenterol*, 21(2113), pp. 3763–3772. doi: 10.3748/wjg.v21.i13.3763.

Pellegrinet, L. *et al.* (2011) 'Dll1- and Dll4-Mediated Notch Signaling Are Required for Homeostasis of Intestinal Stem Cells', *Gastroenterology*. W.B. Saunders, 140(4), pp. 1230-1240.e7. doi: 10.1053/J.GASTRO.2011.01.005.

Philpott, A. and Winton, D. J. (2014) 'Lineage selection and plasticity in the intestinal crypt', *Current Opinion in Cell Biology*. Elsevier Ltd, 31, pp. 39–45. doi: 10.1016/j.ceb.2014.07.002.

Phipps, A. I. *et al.* (2015) 'Association Between Molecular Subtypes of Colorectal Cancer and Patient Survival', *Gastroenterology*. W.B. Saunders, 148(1), pp. 77-87.e2. doi: 10.1053/J.GASTRO.2014.09.038.

Pinto, D. *et al.* (2003) 'Canonical Wnt signals are essential for homeostasis of the intestinal epithelium.', *Genes & development*. Cold Spring Harbor Laboratory Press, 17(14), pp. 1709–13. doi: 10.1101/gad.267103.

Plummer, N. W. *et al.* (2015) 'Expanding the power of recombinase-based labeling to uncover cellular diversity'. doi: 10.1242/dev.129981.

Podgorny, O. *et al.* (2018) 'Triple S-Phase Labeling of Dividing Stem Cells', *Stem*

Cell Reports, 10, pp. 615–626. doi: 10.1016/j.stemcr.2017.12.020.

Potten, C. S. (1977) *Extreme sensitivity of some intestinal crypt cells to X and 'Y irradiation*, *Nature*. doi: 10.1038/269518a0.

Potten, C. S., Owen, G. and Booth, D. (2002) 'Intestinal stem cells protect their genome by selective segregation of template DNA strands', *Journal of Cell Science*, 115(11), pp. 2381–2388.

Powell, A. E. *et al.* (2014) 'Inducible loss of one Apc allele in Lrig1-expressing progenitor cells results in multiple distal colonic tumors with features of familial adenomatous polyposis', *Am J Physiol Gastrointest Liver Physiol*, 307, pp. 16–23. doi: 10.1152/ajpgi.00358.2013.-Individuals.

Preston, S. L. *et al.* (2003) 'Bottom-up histogenesis of colorectal adenomas: origin in the monocryptal adenoma and initial expansion by crypt fission.', *Cancer research*, 63(13), pp. 3819–25. Available at: http://cancerres.aacrjournals.org/content/63/13/3819.abstract (Accessed: 16 October 2015).

Pylayeva-Gupta, Y., Grabocka, E. and Bar-Sagi, D. (2011) 'RAS oncogenes: weaving a tumorigenic web', *Nature Publishing Group*, 11. doi: 10.1038/nrc3106.

Quaife, C. J. *et al.* (1987) 'Pancreatic neoplasia induced by ras expression in acinar cells of transgenic mice', *Cell*. Cell Press, 48(6), pp. 1023–1034. doi: 10.1016/0092-8674(87)90710-0.

Rashtak, S. *et al.* (2017) 'Sessile Serrated Polyps and Colon Cancer Prevention.', *Cancer prevention research (Philadelphia, Pa.)*. American Association for Cancer Research, 10(5), pp. 270–278. doi: 10.1158/1940-6207.CAPR-16-0264.

Robanus-Maandag, E. C. *et al.* (2010) 'A new conditional Apc-mutant mouse model for colorectal cancer', *Carcinogenesis*. Narnia, 31(5), pp. 946–952. doi: 10.1093/carcin/bgq046.

Le Rolle, A.-F. *et al.* (2016) 'Oncogenic KRAS activates an embryonic stem cell-like program in human colon cancer initiation', *Oncotarget*, 7(3). doi: 10.18632/oncotarget.6818.

Ruley, H. E. (1983) 'Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture', *Nature*. Nature

Publishing Group, 304(5927), pp. 602–606. doi: 10.1038/304602a0.

Ryder, E. *et al.* (2014) 'Rapid conversion of EUCOMM/KOMP-CSD alleles in mouse embryos using a cell-permeable Cre recombinase', *Transgenic Research*, 23(1), pp. 177–185. doi: 10.1007/s11248-013-9764-x.

Sadanandam, A. *et al.* (2013) 'A colorectal cancer classification system that associates cellular phenotype and responses to therapy.', *Nature medicine*, 19(5), pp. 619–25. doi: 10.1038/nm.3175.

Sadowski, P. D. (1995) 'The Flp Recombinase of th 2-μm Plasmid of Saccharomyces cerevisiae', *Progress in Nucleic Acid Research and Molecular Biology*. Academic Press, 51, pp. 53–91. doi: 10.1016/S0079-6603(08)60876-4.

Sajgo, S. *et al.* (2014) 'Dre - Cre sequential recombination provides new tools for retinal ganglion cell labeling and manipulation in mice.', *PloS one*. Public Library of Science, 9(3), p. e91435. doi: 10.1371/journal.pone.0091435.

Sakaue-Sawano, A. *et al.* (2008) 'Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression', *Cell*. doi: 10.1016/j.cell.2007.12.033.

Sangiorgi, E. and Capecchi, M. R. (2008) 'Bmi1 is expressed in vivo in intestinal stem cells', *Nature Genetics*. Nature Publishing Group, 40(7), pp. 915–920. doi: 10.1038/ng.165.

Sansom, O. J. *et al.* (2006) 'Loss of Apc allows phenotypic manifestation of the transforming properties of an endogenous K-ras oncogene in vivo.', *Proceedings of the National Academy of Sciences of the United States of America*, 103(38), pp. 14122–7. doi: 10.1073/pnas.0604130103.

Sato, T. *et al.* (2009) 'Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche.', *Nature*. Nature Publishing Group, 459(7244), pp. 262–5. doi: 10.1038/nature07935.

Sato, T. *et al.* (2011) 'Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts $watermark-text $watermark-text $watermark-text', *Nature*, 469(7330), pp. 415–418. doi: 10.1038/nature09637.

Sauer, B. and Henderson, N. (1988) *Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1*, *Proc. Nail. Acad. Sci. USA*. doi: 10.1073/pnas.85.14.5166.

Sauer, B. and McDermott, J. (2004) 'DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages.', *Nucleic acids research*, 32(20), pp. 6086–95. doi: 10.1093/nar/gkh941.

Schepers, A. G. *et al.* (2012) 'Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 337(6095), pp. 730–5. doi: 10.1126/science.1224676.

Schonhoff, S. E., Giel-Moloney, M. and Leiter, A. B. (2004) 'Neurogenin 3-expressing progenitor cells in the gastrointestinal tract differentiate into both endocrine and non-endocrine cell types', *Developmental Biology*, 270(2), pp. 443–454. doi: 10.1016/j.ydbio.2004.03.013.

Schwitalla, S. *et al.* (2013) 'Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties.', *Cell*, 152(1–2), pp. 25–38. doi: 10.1016/j.cell.2012.12.012.

Shih, I.-M. *et al.* (2001) 'Top-down morphogenesis of colorectal tumors', *Proceedings of the National Academy of Sciences*, 98(5), pp. 2640–2645. doi: 10.1073/pnas.051629398.

Shoshkes-Carmel, M. *et al.* (2018) 'Subepithelial telocytes are an important source of Wnts that supports intestinal crypts', *Nature*, 557, pp. 242–248. doi: 10.1038/s41586-018-0084-4.

Shroyer, N. F. *et al.* (2007) 'Intestine-Specific Ablation of Mouse atonal homolog 1 (Math1) Reveals a Role in Cellular Homeostasis', *Gastroenterology*, 132(7), pp. 2478–2488. doi: 10.1053/j.gastro.2007.03.047.

Simons, B. D. (2016) 'Deep sequencing as a probe of normal stem cell fate and preneoplasia in human epidermis.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 113(1), pp. 128–33. doi: 10.1073/pnas.1516123113.

Sinha, M. and Lowell, C. A. (2017) 'Efficiency and Specificity of Gene Deletion in Lung Epithelial Doxycycline-Inducible Cre Mice.', *American journal of respiratory cell and molecular biology*. American Thoracic Society, 57(2), pp. 248–257. doi: 10.1165/rcmb.2016-0208OC.

Sinn, E. *et al.* (1987) ‘Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: Synergistic action of oncogenes in vivo’, *Cell*. Cell Press, 49(4), pp. 465–475. doi: 10.1016/0092-8674(87)90449-1.

Snippert, H. J. *et al.* (2010) ‘Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells.’, *Cell*, 143(1), pp. 134–44. doi: 10.1016/j.cell.2010.09.016.

Snippert, H. J. *et al.* (2014) ‘Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion.’, *EMBO reports*, 15(1), pp. 62–9. doi: 10.1002/embr.201337799.

Soriano, P. (1999) ‘Generalized lacZ expression with the ROSA26 Cre reporter strain’, *Nature Genetics*, 21, pp. 70–71. doi: https://doi.org/10.1038/5007.

Sottoriva, A. *et al.* (2013) ‘Single-molecule genomic data delineate patient-specific tumor profiles and cancer stem cell organization.’, *Cancer research*. American Association for Cancer Research, 73(1), pp. 41–9. doi: 10.1158/0008-5472.CAN-12-2273.

Sottoriva, A. *et al.* (2015) ‘A Big Bang model of human colorectal tumor growth’, *Nature Genetics*. Nature Publishing Group, 47(3), pp. 209–216. doi: 10.1038/ng.3214.

De Sousa E Melo, F. *et al.* (2013) ‘Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions’, *Nature Medicine*. Nature Publishing Group, 19(5), pp. 614–618. doi: 10.1038/nm.3174.

Spicer, S. S. (1960) ‘A CORRELATIVE STUDY OF THE HISTOCHEMICAL PROPERTIES OF RODENT ACID MUCOPOLYSACCHARIDES’, *Journal of Histochemistry & Cytochemistry*, 8(1), pp. 18–35. doi: 10.1177/8.1.18.

Stamp, C. *et al.* (2018) ‘Predominant Asymmetrical Stem Cell Fate Outcome Limits the Rate of Niche Succession in Human Colonic Crypts’. doi: 10.1016/j.ebiom.2018.04.017.

Sternberg, N. and Hamilton, D. (1981) ‘Bacteriophage P1 site-specific recombination: I. Recombination between loxP sites’, *Journal of Molecular Biology*. Academic Press, 150(4), pp. 467–486. doi: 10.1016/0022-

2836(81)90375-2.

Su, L. *et al.* (1992) ‘Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene’, *Science*. American Association for the Advancement of Science, 256(5057), pp. 668–670. doi: 10.1126/science.1350108.

Takeda, N. *et al.* (2011) ‘Interconversion Between Intestinal Stem Cell Populations in Distinct Niches’, *Science reports*, 334, pp. 1420–1424. doi: 10.1126/science.1210125.

Taylor, R. W. *et al.* (2003) ‘Mitochondrial DNA mutations in human colonic crypt stem cells.’, *The Journal of clinical investigation*. American Society for Clinical Investigation, 112(9), pp. 1351–60. doi: 10.1172/JCI19435.

Tetteh, P. W. *et al.* (2016) ‘Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters’, *Cell Stem Cell*. doi: 10.1016/j.stem.2016.01.001.

The Cancer Genome Atlas Network (2012) ‘Comprehensive molecular characterization of human colon and rectal cancer.’, *Nature*. Nature Publishing Group, 487(7407), pp. 330–7. doi: 10.1038/nature11252.

Tian, H. *et al.* (2011) ‘A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable’, *Nature*. Nature Publishing Group, 478(7368), pp. 255–259. doi: 10.1038/nature10408.

Tomasetti, C., Li, L. and Vogelstein, B. (2017) ‘Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention.’, *Science (New York, N.Y.)*. American Association for the Advancement of Science, 355(6331), pp. 1330–1334. doi: 10.1126/science.aaf9011.

Tomasetti, C. and Vogelstein, B. (2015) ‘Variation in cancer risk among tissues can be explained by the number of stem cell divisions’, *Science*. American Association for the Advancement of Science, 347(6217), pp. 78–81. doi: 10.1126/science.1260825.

Tomasetti, C., Vogelstein, B. and Parmigiani, G. (2013) ‘Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation’, *PNAS*, 110(6), pp. 1999–2004. doi: 10.1073/pnas.1221068110.

Tomic, G. *et al.* (2018) 'Phospho-regulation of ATOH1 Is Required for Plasticity of Secretory Progenitors and Tissue Regeneration', *Cell Stem Cell*. Elsevier, 0(0). doi: 10.1016/j.stem.2018.07.002.

Valentin-Vega, Y. A., Okano, H. and Lozano, G. (2008) 'The intestinal epithelium compensates for p53-mediated cell death and guarantees organismal survival', *Cell Death & Differentiation*. Nature Publishing Group, 15(11), pp. 1772–1781. doi: 10.1038/cdd.2008.109.

VanDussen, K. L. *et al.* (2012) 'Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells', *Development*, 139(3), pp. 488–497. doi: 10.1242/dev.070763.

VanDussen, K. L. and Samuelson, L. C. (2010) 'Mouse atonal homolog 1 directs intestinal progenitors to secretory cell rather than absorptive cell fate', *Developmental Biology*. Academic Press, 346(2), pp. 215–223. doi: 10.1016/J.YDBIO.2010.07.026.

Vermeulen, L. *et al.* (2013) 'Defining stem cell dynamics in models of intestinal tumor initiation.', *Science (New York, N.Y.)*, 342(6161), pp. 995–8. doi: 10.1126/science.1243148.

Vincent, J.-P., Fletcher, A. G. and Baena-Lopez, L. Al. (2013) 'Mechanisms and mechanics of cell competition in epithelia', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 14(9), pp. 581–591. doi: 10.1038/nrm3639.

Vooijs, M., Jonkers, J. and Berns, A. (2001) 'A highly efficient ligand-regulated Cre recombinase mouse line shows that LoxP recombination is position dependent.', *EMBO reports*. European Molecular Biology Organization, 2(4), pp. 292–7. doi: 10.1093/embo-reports/kve064.

Wehkamp, J. and Stange, E. F. (2010) 'Paneth's disease'. doi: 10.1016/j.crohns.2010.05.010.

Westphalen, C. B. *et al.* (2014) 'Long-lived intestinal tuft cells serve as colon cancer-initiating cells', *The Journal of Clinical Investigation*, 124. doi: 10.1172/JCI73434.

Winton, D. J., Blount, M. A. and Ponder, B. A. J. (1988) 'A clonal marker induced by mutation in mouse intestinal epithelium', *Nature*. Nature Publishing Group,

333(6172), pp. 463–466. doi: 10.1038/333463a0.

Xie, M. *et al.* (2014) 'Age-related mutations associated with clonal hematopoietic expansion and malignancies.', *Nature medicine*. Nature Publishing Group, (CII). doi: 10.1038/nm.3733.

Yan, K. S. *et al.* (2012) 'The intestinal stem cell markers Bmi1 and Lgr5 identify two functionally distinct populations.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 109(2), pp. 466–71. doi: 10.1073/pnas.1118857109.

Yanai, H. *et al.* (2017) 'Intestinal cancer stem cells marked by Bmi1 or Lgr5 expression contribute to tumor propagation via clonal expansion', *Scientific Reports*. Nature Publishing Group, 7(1), p. 41838. doi: 10.1038/srep41838.

Yang, Q. *et al.* (2001) 'Requirement of Math1 for secretory cell lineage commitment in the mouse intestine.', *Science*, 294(5549), pp. 2155–2158. doi: 10.1126/science.1065718.

Young, N. P., Crowley, D. and Jacks, T. (2011) 'Tumor and Stem Cell Biology Uncoupling Cancer Mutations Reveals Critical Timing of p53 Loss in Sarcomagenesis'. doi: 10.1158/0008-5472.CAN-10-4563.

Yuan, W. *et al.* (2013) 'Conditional activation of Pik3caH1047R in a knock-in mouse model promotes mammary tumorigenesis and emergence of mutations', *Oncogene*. Nature Publishing Group, 32(3), pp. 318–326. doi: 10.1038/onc.2012.53.

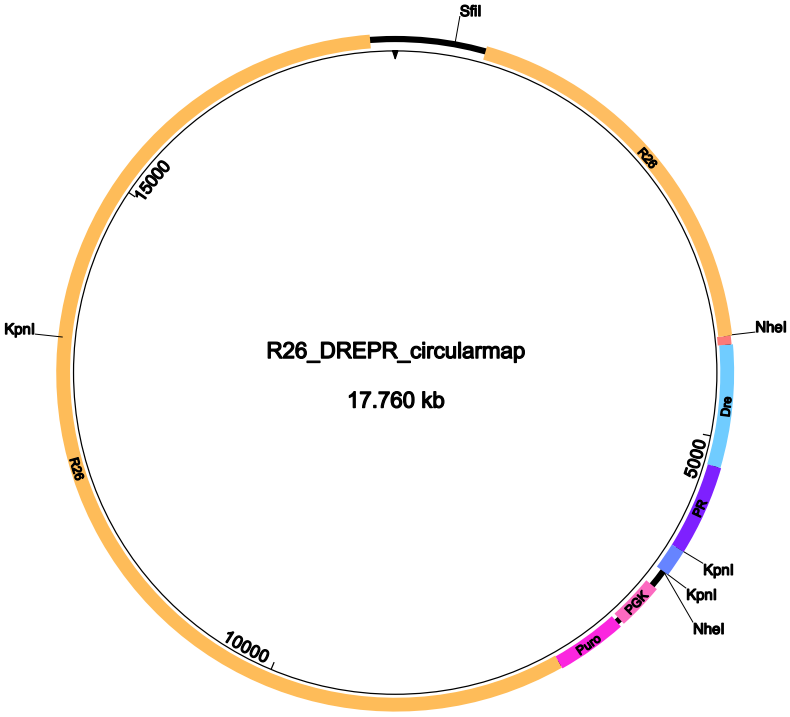
Zhu, L. *et al.* (2009) 'Prominin 1 marks intestinal stem cells that are susceptible to neoplastic transformation', *Nature*. Nature Publishing Group, 457(7229), pp. 603–607. doi: 10.1038/nature07589.

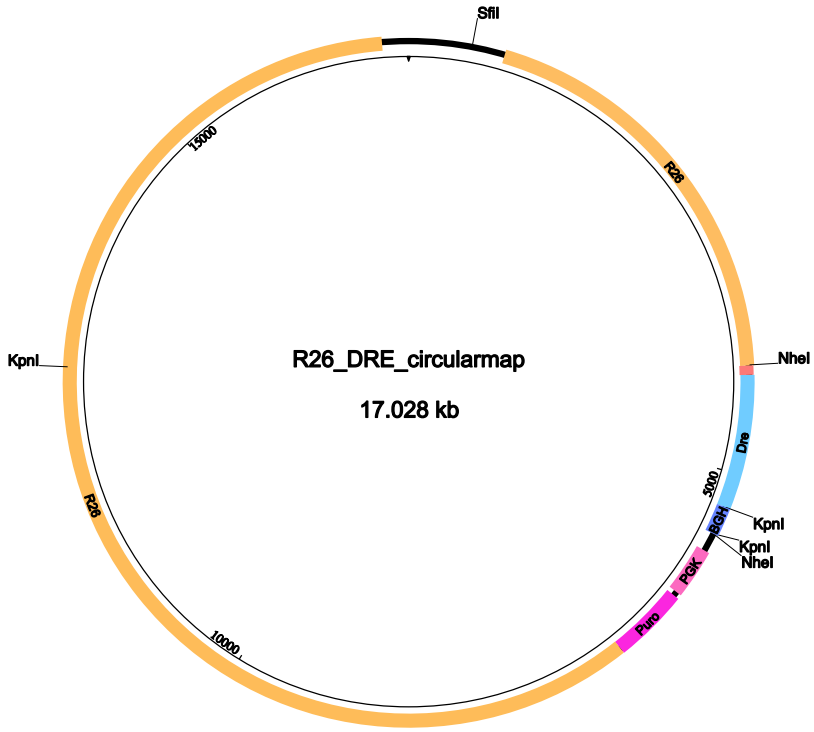
8 Appendices

Appendix 1: PCR primers

Primer name	Sequence
Infusion primers	
Dre_insert1_Fwd	CAGGTCGACTCTAGAGGATCTGCTAGCGTAGGGCGCAGTAGTCCAGGGTTTCCTTGATGA TGTCATACTTATCCTGTCCCTTTTTTTTCCACAGCTCCACCATGGGTGCCAGCGAGCTG
Dre_insert1_Rev	GAAGGCACAGGGTACCTCACACTTTCCTCTTCTTC
Dre_insert2_Fwd	GTGAGGTACCCTGTGCCTTCTAGTTGCCAGC
Dre_insert2_Rev	GAGCTCGGTACCCGGGGATCGCTAGCGGTACCCCATAGAGC
DrePr_insert1_Fwd	CAGGTCGACTCTAGAGGATCTGCTAGCGTAGGGCGCAGTAGTCCAGGGTTTCCTTGATGA TGTCATACTTATCCTGTCCCTTTTTTTTCCACAGCTCCACCATGGGTGCCAGCGAGCTG
DrePr_insert1_Rev	TGAACTTTTTGGCGCCGCTGGCGGAGTC
DrePr_insert2_Fwd	CAGCGGCGCCAAAAAGTTCAATAAAGTCAGAGTTG
DrePr_insert2_Rev	ACAGGGTACCTCAGCAGTACAGATGAAGTTG
DrePr_insert3_Fwd	GTACTGCTGAGGTACCCTGTGCCTTCTAGTTG
DrePr_insert3_Rev	GAGCTCGGTACCCGGGGATCGCTAGCGGTACCCCATAGAGC
Dre and DrePr mESC verification primers	
F1	AGAGTCCTG-ACCCAGGGAAGACATT
R1	CATCAAGGAAACCCTGGACTACTGCG
F2	GTCACCGAGCTGCAAGAACTCT
R2	GGTGGTGGTGGTGGCATATAC-ATT
Kras LSL and FSF recombination primers	
Kras_P1	gtctttccccagcacagtgc
Kras_P2	ctcttgcctacgccaccagctc
Kras_P3	agctagccaccatggcttgagtaagtctgca
Atoh1 genotyping primers	
Atoh1_Forward	TTTGTTGTTGTTGTTCGGGG
Atoh1_Reverse	TCTTTTACCTCAGCCCACTCT T

Appendix 2: Vector maps





Appendix 3: Genotyping probes for Transnetyx

Probe name	Identifies	Forward	Reverse	Reporter 1	Reporter 2
tdTom	R_tdTom^{fsf} and R_tdTom^{lsl}	AGATCCACCAGGCCCTGAA	GTCTTGAA-CTCCACCAGGTAGTG	CCGCCGTCCTTCAGC	NA
R26_WT	R_tdTom wt	TTCCCTCGTGATCTGCAACTC	CTTTAAGCCTGCCCAGAAGACT	CCGCCCATCTTCTAGAAAG	NA
R26_3WT	R_Dre wt	CTCTTACACTAAGCAATAAAGAAATA AAAATTGAACTTCT	CTGCAGACTTAGCTTTCAGCTTTG	ACTGCTAGCTTTACTTAACTTT	NA
Dre-2 Tg	R_Dre	GATGAGGATGTCTGAGATCTCTAGGA	GGTCTTGGTGTGGGAGATGTG	CCAGACCGGAGACACC	NA
Dre-PGR Tg	R_Dre^{Pr}	GTCTGTGGGAGCCAACTCTAG	CAGTGCTCTCACAACTCTGACTTTA	TTTTTGGCGCCGCTGGCG	NA
G12D mut	Kras^{FrtG12D} and Kras^{loxG12D}	GGCCTGCTGAAAATGACTGAGTATA GTATAGCATACATTATACGAAGTTATC	CTGTATCGTCAAGGCGCTCTT	TTGGAGCTGGTGGCGT	TGGAGCTGATGGCGT
Cyp1a1-1 KO	CypFlpO	CTAGGG	CCCAGAATCCAAGGCAGAATACG	CCGTCGACAGCTGTC	NA
Cyp1a1-1 WT	Cyp1a1 WT	GTATGGACTTCCAGCCTTCGT	CCAAGGCAGAATACGGTGACA	CCAGGAGCAGCTCTGT	NA
FlpO	Villin_FlpO	CAAGCTGGTGCAGAACAAGTAC	TGGACACGCTTGTCTTGGT	CTGGGCGTGATCATTC	NA
Atoh1-5 WT	AtohCre^{ERT2}	GCCAGCACCTCCTCTAACAC	CTCAGCCCACTCTTCTGCAT	TTGCAGTGCGATGTCC	NA
DBC-1 KO	Ctnnb$^{lox(ex3)}$	CGTCCAGCCAAGCTTAGGAT	CGGTGGCTTGCTGATTATTTCAC	CCCCTCGACTTCTGC	NA
DBC-1 WT	Ctnnb WT	GCAAGAGCAAGTAGCTGGTAAAG	CACAGCCCTGTCAAGAAACTTAATG	CTAGCTTCCAAACACAAATG	NA
AhCreER Tg	AhCre^{Ert2}	ATCTCGACCATGCCCAAGAAG	GCAAATTTTGGTGTACGGTCAGT	AAGAGGAAGGGTGTCCAATTT	NA
Ahr-1 MUT	Ahr mutation	CTTTTGCAAGAGCATCAAGTTCCT	CAAACCCTCCATCAGTACAGTACA	CCTCACGTTAGCCC	CCTCACCTTAGCCC
DTR tg	Lgr5-GFP-DTR	ACGGAGAATGCAAATATGTGAAGGA	ACACCTCTCTCCATGGTAACCT	CCTCCTGCATCTGCC	NA
eGFP	Lgr5-GFP-DTR	CGTCGTCCTTGAAGAAGATGGT	CACATGAAGCAGCACGACTT	CATGCCCGAAGGCTAC	NA

Appendix 4: Antibodies for IHC

Antibody	Manufacturer	Catalogue Number	Raised Species	Dilution	Antigen retrieval	Temp for antigen retrieval
BrdU	Abcam	Ab1893	Sheep	2.6 ug/ml	Proteinase K, 10'	37 C
Lysozyme	Dako	A0099	Rabbit	1;2500	Proteinase K, 10'	37 C
tdTom	Abcam	Ab34771	Rabbit	1;100	Sodium Citrate, 20'	100 C
GFP	Abcam	Ab13970	Chicken	10 ug/ml	Proteinase K, 20'	37 C
Beta-Catenin	BD Biosciences	610154	Mouse	0.25 ug/ml	Sodium Citrate, 20'	37 C

Appendix 5: Script for random clonal localization

```
#!/usr/bin/perl -w
# Simulation of random contiguation



print "\nEnter total number of crypts\n";
$crypts = <STDIN>;
chomp($crypts);

print "\nEnter number of marked crypts\n";
$marked = <STDIN>;
chomp($marked);

print "\nEnter number of runs\n";
$runs = <STDIN>;
chomp($runs);

$iteration =1;
while ($iteration <= $runs){
        # Create the array of cells
        @CELLS = ();
        $cells = 1;
        while ($cells <= $crypts){
                push(@CELLS,"-");
                $cells = $cells +1;
        }

        # Randomly mark some of these cells
        $count = 1;
        while ($count <=$marked){
                $marked_cell = int(rand($crypts));
                if      ($CELLS[$marked_cell] eq "@"){
                }else{
                        $CELLS[$marked_cell] = "@";
                $count = $count +1;
                }
        }


        print "Run $iteration\n";
        # Print array of marked cells
        foreach $val (@CELLS){
                print "$val ";
        }
        print "\n\n";
        $iteration = $iteration +1;
}
exit;
```

Appendix 6: R-Script for Prediction of PR and λ in Atoh1 datasets

```
library(tidyverse)
library(CryptDriftR)
library(cowplot)
theme_set(theme_cowplot(font_size=20))

# Function to use neutral drift equations to get ratio
get_ratio = function(N, lambda, tau, Pr = 0.5, day = 30, remove_size = 0)
{
  prob_vec = th_PulseChase(lambda, N, tau, day, F, Pr = Pr)
  prob_vec[N+1]/sum(prob_vec[(2+remove_size):(N)])
}

find_closest_val <- function(y_th, x_th, x_find){
  ind_min = which.min(abs(x_th - x_find))
  y_th[ind_min]
}

ratios_atoh      = c(0.824, 0.688, 0.598)
ratios_atoh_kras = c(1.93150685, 2.75862069, 1.84810127, 1.43661972 )

size_remove = 1
N_use       = 7
tau_use     = 1


## First WT
```

```
mean_ratio_atoh = mean(ratios_atoh)
se_ratio_atoh   = sd(ratios_atoh)/sqrt(length(ratios_atoh))

tbl_th = tibble(lambda = seq(0.1, 0.4, length.out =  25),
                ratio  = lambda %>% map_dbl( ~ get_ratio(N_use, .x, tau_use, remove_size = size_remove)))
pp_WT = tbl_th %>%
  ggplot(aes(x = lambda, y = ratio)) + geom_point(size = 4) +
  geom_hline(yintercept = mean_ratio_atoh, col = "red") +
  ggtitle("Predicted ratio varying lambda, N = 7")

# Get estimates of lambda ==========
tbl_th = tibble(lambda = seq(0.1, 0.4, length.out =  500),
                ratio  = lambda %>% map_dbl( ~ get_ratio(N_use, .x, tau_use, remove_size = size_remove)))

lambda_est = find_closest_val(tbl_th$lambda, tbl_th$ratio, mean_ratio_atoh)
lambda_min = find_closest_val(tbl_th$lambda, tbl_th$ratio, mean_ratio_atoh - 1.96*se_ratio_atoh)
lambda_max = find_closest_val(tbl_th$lambda, tbl_th$ratio, mean_ratio_atoh + 1.96*se_ratio_atoh)


## Then KRAS ------------------------------------------------------------------------------
mean_ratio_atoh_kras = mean(ratios_atoh_kras)
se_ratio_atoh_kras   = sd(ratios_atoh_kras)/sqrt(length(ratios_atoh_kras))

tbl_th_Pr = tibble(Pr = seq(0.5, 0.8, length.out =  25),
```

```
                ratio  = Pr %>% map_dbl( ~ get_ratio(7, lambda_est, tau_use, Pr =
.x, remove_size = size_remove)))



pp_KRAS = tbl_th_Pr %>%
  ggplot(aes(x = Pr, y = ratio)) + geom_point(size = 4) +
  geom_hline(yintercept = mean_ratio_atoh_kras, col = "red") +
  ggtitle("Predicted ratio varying Pr, N = 7 (For KRAS)")


# Get estimates of Pr ====================================
tbl_th_Pr = tibble(Pr = seq(0.5, 0.8, length.out =  500),
                ratio  = Pr %>% map_dbl( ~ get_ratio(7, lambda_est, tau_use, Pr =
.x, remove_size = size_remove)))


Pr_est = find_closest_val(tbl_th_Pr$Pr, tbl_th_Pr$ratio, mean_ratio_atoh_kras)
Pr_min = find_closest_val(tbl_th_Pr$Pr, tbl_th_Pr$ratio, mean_ratio_atoh_kras -
1.96*se_ratio_atoh_kras)
Pr_max = find_closest_val(tbl_th_Pr$Pr, tbl_th_Pr$ratio, mean_ratio_atoh_kras
+ 1.96*se_ratio_atoh_kras)


pdf("Plots_Sofie.pdf", 9, 7)
plot(pp_WT)
plot(pp_KRAS)
dev.off()


print(c(Pr_est, Pr_min, Pr_max))
print(c(lambda_est, lambda_min, lambda_max))
```