U.PORTO

DOUTORAMENTO

BIOLOGIA BÁSICA E APLICADA

Deciphering the molecular and genomic landscapes in the progression of serrated intestinal tumours

Miguel Gomes Silva

D

2021

Deciphering the molecular and genomic landscapes in the progression of serrated intestinal tumours

Miguel Gomes Silva

Instituto Ciências Biomédicas Abel Salazar
Faculdade de Ciências
Faculdade de Medicina

Miguel Gomes Silva

Deciphering the molecular and genomic landscapes in the progression of serrated intestinal tumours

Tese de Candidatura ao grau de Doutor em Àreas da Biologia Básica e Aplicada

Programa Doutoral da Universidade do Porto (Instituto de Ciências Biomédicas de Abel Salazar e Technische Universität München (TUM)

Orientador – Professor Doutor Roland Rad

Categoria – Diretor do Instituto de Oncologia Molecular (IMO), TUM

Afiliação – Technische Universität München (TUM)

Co-orientador – Doutor José Luís Costa

Categoria – Investigador afiliado ao grupo de investigação “Genetic dynamics of cancer cells”

Afiliação – Instituto de Investigação e Inovação em Saúde da Universidade do Porto (i3S)

Acknowledgements

All the work developed for this thesis would not have been possible without the contribution of several people that directly or indirectly were a part of it. A journey as difficult as it may be, is always easier when you don't do it alone.

First and foremost, I would like to thank the person that made everything possible. This work would have never seen the light of day if it was not for Markus Tschurstchenthaler. He taught me that all the struggles, drawbacks and frustrations can be overcome with work and tenacity and that achievements and milestones should be rewarded and celebrated. One of the people I know with the biggest heart and strongest character. The best colleague and mentor one can ask for and a true friend.

To Moritz Jesinghaus, to Katha and Emilia my extended family. Thank you for all the teachings, constant support, friendship and for always watching over me. For showing me that you can also choose family and that things will always work out the best way possible.

To my supervisor Roland Rad for giving me the opportunity to develop the work for this thesis at his lab.

To the AG Saur members that contributed to this work and helped with the experimental part in particular Antonio Zaurito and Valentina Brunner.

To all the lab members of AG Rad for making this journey easier and for welcoming me so nicely. True friendships were formed there that won't be easily broken.

To all my friends in Portugal that stood by my side through good and bad times and shortened the distance between both homes. To all my friends in Munich for showing me there is life outside the lab.

To my family in Portugal for the constant support and love.

To Federica, my wall, my support, my strength and my love. Without her help, her understanding, her care and concern I could not have done it. For showing me that every cloud has a silver lining.

To my parents. For giving me the necessary strength and love I needed to carry on and never give up. For always being there for me whenever I needed them. For showing me I was never alone. For teaching me that hard work pays off and that I am capable of achieving my goals using their example.

I would like to dedicate this thesis to my grandmother. One of the greatest human beings I will ever meet.

Resumo

O cancro colorectal (CRC) é uma condição complexa com uma etiologia heterogénea causada por uma combinação de fatores ambientais, genéticos e epigenéticos. A progressão da maioria dos casos de CRC segue os postulados do bem estabelecido "Vogelgrama" iniciado pela perda de função da proteína APC seguido de mutações nos genes *KRAS*, *SMAD4* e *TP53.* Em contraste, um subgrupo de tumores segue a "via serrillhada" de carcinogénese que começa geralmente com mutações no gene *KRAS* ou *BRAF* e que, consequentemente, levam a ativação da via de sinalização MAPK/ERK. Estes tumores representam uma entidade clinicamente mais agressiva, associada com resistência á terapia. No entanto, os mecanismos responsáveis por estas observações permanecem por identificar. Inúmeros modelos *in vitro* e *in vivo* foram recentemente gerados que possibilitam a modulação do CRC serrilhado, porém com consideráveis limitações. O estabelecimento de linhas células 2D derivadas de CRC é um processo extremamente ineficiente no qual apenas clones derivados de tumores avançados podem ser cultivados. Modelos de ratinho geneticamente modificados (GEMMs) refletem aspectos genéticos e morfológicos cruciais dos tumores de pacientes com CRC serrilhado, incluindo as mutações iniciadoras de tumorigénese ($Kras^{G12D}$, $Braf^{V637E}$, $Pik3ca^{H1047R}$) e histologia serrilhada. Porém, estes modelos são intrinsicamente limitados pelo longo tempo de formação dos tumores nos animais, ausência de estadios avançados de doença e o facto de os tumores se formarem quase exclusivamente no intestino delgado e não no colón e recto. De modo a ultrapassar os obstáculos inerentes a estes modelos, estabelecemos, no laboratório, um modelo de cancro *in vivo* baseado na transplantação ortotópica de organóides geneticamente modificados ou derivados de tumores, na submucosa do colón de ratinhos. Para tal, geramos uma extensa coorte de 341 linhas de organóides isolados de tecido normal, hiperplasia e tumores (adenomas e carcinomas) de diferentes modelos genéticos de ratinho ($Apc^{fle1-15}$, $Kras^{G12D}$, $Braf^{V637E}$, $Pik3ca^{H1047R}$). Transplantações ortotópicas de organóides derivados de estadios iniciais (adenomas) de tumores serrilhados de ratinho permitiu a continuidade da progressão tumoral *in vivo* e levou a formação de tumores avançados (carcinomas) na maioria dos animais e formação de metástases num subgrupo de ratinhos injetados. Detectamos que a ativação da via de sinalização Wnt causada por mutações no gene *Ctnnb1* ou *Apc* colabora com as mutações oncogénicas $Kras^{G12D}$ e $Braf^{V637E}$ em estadios iniciais da doença. A progressão dos adenomas iniciados no contexto $Kras^{G12D}$ é mediada pelo aumento de dose oncogénica (iGD) que juntamente com mutações nos genes *Cdkn2a* e *Trp53* medeiam a transição de adenoma para carcinoma. A nível da transcrição observamos a ativação de programas relacionados com transição epitelial-mensenquimal

(EMT) e inflamação, em estados iniciais da doença que estão presentes durante a progressão tumoral. Em estadios avançados, as células tumorais ativam vias de sinalização que promovem a sua interação com células do sistema imunitário presentes no microambiente tumoral (TME). Paralelamente, tumores iniciados pela mutação *Braf*^{V637E} também mostram ativação de EMT e inflamação na transição de hiperplasia para adenoma. No entanto a transição de adenomas para carcinomas é caracterizada por uma forte ativação de vias de sinalização dependentes do gene *Myc* e do ciclo celular. Para além de EMT e inflamação, no contexto *Pik3ca*^{H1047R} a formação de adenomas é mediada por uma resposta dependente de interferão tipo I e II. Adicionalmente, injeções ortotópicas usando organóides derivados de carcinomas humanos, em ratinhos NSG, confirmou a viabilidade de usar o ratinho como um modelo *avatar* para o estudo de CRC no contexto humano. Mostramos também que, usando este modelo, é possível recapitular no ratinho aspetos de doença localizada e distante (metástases), e que o perfil genético e transcriptómico dos organóides derivados de ratinhos transplantados refletem os perfis das linhas parentais humanas usadas para implantação.

As ferramentas e recursos gerados no âmbito desta tese permitiu a modulação de aspetos cruciais de CRC humano e de ratinho e possibilitou a identificação dos eventos biológicos que ocorrem em cada estadio de progressão de doença, em *backgrounds* genéticos distintos.

Abstract

Colorectal cancer (CRC) is a complex condition with a heterogeneous aetiology, caused by a combination of environmental, genetic and epigenetic factors. Progression of most CRC cases follows the postulates of the “Vogelgram” which is well characterized as it is initiated by the loss of a functional APC protein followed by *KRAS*, *SMAD4* and *TP53* mutations. In contrast, a subset of cancers follows the “serrated pathway” through initiating oncogenic mutations of either *KRAS* or *BRAF* that leads to an upregulation of MAPK/ERK pathway. These tumours are clinically more aggressive and associated with therapy resistance nonetheless, the reasons or mechanisms leading to these findings remain partially elusive. Several *in vitro* and *in vivo* models have been generated in recent years to model serrated CRC, however with considerable limitations. The establishment of standard 2D CRC cell lines is a very inefficient process and limits the culture to culture-competent cell clones from advanced disease. Genetically engineered mouse models (GEMMs) reflect key genetic and morphological features of human serrated tumours including the initiating genetic events ($Kras^{G12D}$, $Braf^{V637E}$, $Pik3ca^{H1047R}$) and serrated histology. However, these models are limited by the animals’ long tumour latency, the lack of advanced disease stages and, finally, the location of the tumours in the small intestine rather than in the colon and rectum.

In order to overcome the aforementioned caveats of both settings (*in vitro* and *in vivo*) we established a cancer mouse model based on the orthotopic transplantations of genetically engineered or tumour-derived organoids into the colon submucosa of recipient mice. We generated a large cohort accounting for a total of 341 murine normal tissue, hyperplastic and tumour-derived organoids from distinct disease stages in different backgrounds ($Apc^{fle1-15}$, $Kras^{G12D}$, $Braf^{V637E}$, $Pik3ca^{H1047R}$). Orthotopic transplantation of serrated adenoma-derived organoids in the mouse colon resumed tumour progression and led to the formation of advanced tumours (carcinomas) in the majority of animals and distant disease (metastases) in a subset of mice. We found the upregulation of Wnt pathway either by *Ctnnb1* or *Apc* mutations, to cooperate with oncogenic $Kras^{G12D}$ and $Braf^{V637E}$ in early stages of tumorigenesis. Genetic progression of $Kras^{G12D}$-driven adenomas is mediated by increased gene dosage (iGD) that together with *Cdkn2a* deletion and *Trp53* mutations drive the adenoma-carcinoma transition. Transcriptionally, we observed epithelial to mesenchymal transition (EMT) and inflammation-related programs to be activated in initial stages of disease and to follow tumour progression. In advanced disease, cancer cells activate immune signalling pathways that mostly rely on their interaction with immune cells from the tumour microenvironment (TME). In a similar way, $Braf^{V637E}$-driven tumours also upregulate EMT and inflammatory programs during the transition from hyperplasia to

adenoma however the transition from adenomas to carcinomas is characterized by a strong activation of Myc signalling and cell cycle effectors. Finally, in addition to EMT and inflammation, in the *Pik3ca*^{H1047R} setting, interferon response seems to mediate malignant transformation. In addition, orthotopic injections of carcinoma-derived organoids from human patients into NSG mice showed the feasibility of using the mouse as an avatar model for the study of human CRC. We were able to recapitulate aspects of both localized and metastatic disease in the mouse. Furthermore, genetic and transcriptomic analysis of implantation derived organoids (IDOs) lines reflected the landscapes and profiles from the parental injected carcinoma lines.

The unique tools and resources generated in the scope of this thesis offer the possibility to modulate aspects of both murine and human CRC and allowed the identification of the biological events occurring in each stage of disease progression in distinct genetic backgrounds.

List of abbreviations

APC	Adenomatous polyposis coli
BMP	Bone morphogenic protein
BRAF	v-Raf murine sarcoma viral oncogene homolog B
CDKN2A	Cyclin-dependent kinase inhibitor 2A
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
CMS	Consensus molecular subtypes
CNV	Copy number variation
CRC	Colorectal cancer
CRIS	CRC intrinsic subtypes
CRISPR	Clustered regularly interspaced short palindromic repeats
DNA	Deoxyribonucleic acid
ECM	Extracelular matrix
EEC	Enteroendocrine cells
EGF	Epidermal growth factor
EMT	Epithelial to mesenchymal transition
ERK	Extracellular receptor kinase
GEMM	Genetically engineered mouse model
GTP	Guanosine triphosphate
GDP	Guanosine diphosphate
IDO	Implantation-derived organoids
iGD	Increased gene dosage
IHC	Immunohistochemistry
ITH	Intratumour heterogeneity
KRAS	Kirsten rat sarcoma viral oncogene homolog
MAPK	Mitogen-activated protein kinase
MMR	Mismatch repair
MSI	Microsatellite instability
OIS	Oncogene-induced senescence
lcWGS	Low coverage whole genome sequencing
PCA	Pincipal componente analysis
PDO	Patient derived organoids
PDOX	Patient derived organoids-based xenograft
PDX	Patient derived xenograft
LOH	Loss of heterozygosity
PI3K	Phosphoinositide 3-kinase
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
TAM	Tamoxifen

TGFβ	Transforming growth factor beta
TP53	Tumour protein P53
TSA	Traditional serrated adenoma
TSG	Tumour supressor gene
SBS	Single base substitution
sgRNA	Single guide ribonucleic acid
SNV	Single nucleotide variant
SSA	Sessile serrated adenoma
VAF	Variant allele frequency
WES	Whole exome sequencing
WGS	Whole genome sequencing
Wnt	Wingless-related integration site
wt	Wild-type

List of figures

Table of Contents

1 Introduction

1.1 Colorectal cancer: Disease overview

Colorectal cancer (CRC) is a complex condition with a heterogeneous aetiology, caused by a combination of various factors with environmental, genetic and epigenetic contributions (11). In fact, CRC is generally viewed not as a single disease, but rather as a group of diseases, each with a different profile that is unlikely to be exactly recapitulated by any other. It is the second most common diagnosed cancer in women and third most common in men, however both incidence and mortality are lower in women than in men (12). Highest incidence values are reported in developed countries such as in the US, Europe and Australia (**Figure 1**). In these countries, numbers are steadily decreasing which can be attributed to the establishment of early screening programs such as colonoscopy and changes in dietary behaviours. On the other hand, in developing countries numbers continue to rise. In fact, it is expected that by 2030, there will be approximately 2.2 million new cases and 1.1 million cancer deaths attributed to CRC worldwide (13). Additionally, a worrying high percentage of new cases in young individuals (<50 years) has been recently reported, mostly for rectal cancer and left sided colon cancer (14), which further emphasizes the importance of early cancer prevention.

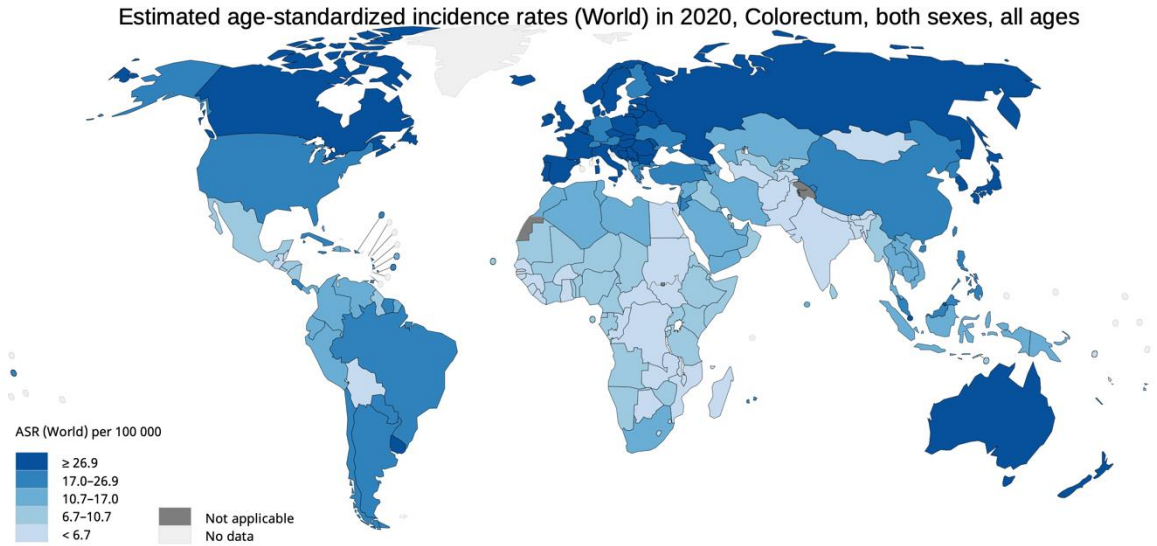


Figure 1 | CRC incidence rates in 2020 in the world (Adapted from (1))

Most CRC cases start with an aberrant crypt that gives rise to a polyp and can ultimately progress to CRC in a process that spans approximately 10-15 years in the sporadic setting (15). Stem cells at the base of the intestinal crypt are believed to be at the center of this process. Through genetic and epigenetic insults these cells activate

oncogenes and inactivate tumour suppressor genes (TSG) leading to malignant transformation (16). Therefore, limiting the growth of a pre-neoplastic lesion through screening approaches can prevent disease progression. Modern CRC screening techniques, such as colonoscopy, can efficiently detect early-stage disease and remove precancerous lesions. It is estimated that nearly 60% of CRC deaths could be prevented if population 50 years or older were screened routinely (17). The choice of the strategy for CRC treatment depends mostly on the stage and mutational status of the tumour. Surgery remains the most successful treatment option for CRC and it is recommended, whenever possible, for all disease stages due to its curative intent (18). When treating an early stage localized disease, surgery has a greater chance of cure, nonetheless, 20-25% of patients present metastases at the time of diagnosis, and approximately 50-60% of the remainder, will develop metastases, contributing to the high mortality rates of CRC registered every year (19). Chemoradiotherapy is the most common first line therapy for CRC with the goal of tumour downsizing to render surgery possible (20).

Although most CRC cases are sporadic, inherited cancer syndromes account for up to 5-10% of all cases (21). In addition, there are a number of cases (around 25%) in which a family history does not follow a classical Mendelian pattern, presumably because of incomplete penetrance and/or multifactorial effects (22). The genetic and epigenetic events that lead to the onset and progression of CRC will be discussed in the upcoming chapters of this thesis.

1.2 The intestine composition: a single-cell perspective

The intestinal epithelium is composed of millions of crypt-villus units. Each unit is composed of a villi which is no more than a protrusion of the intestinal wall into the lumen of the gut, surrounded by several invaginations (crypts) (7).

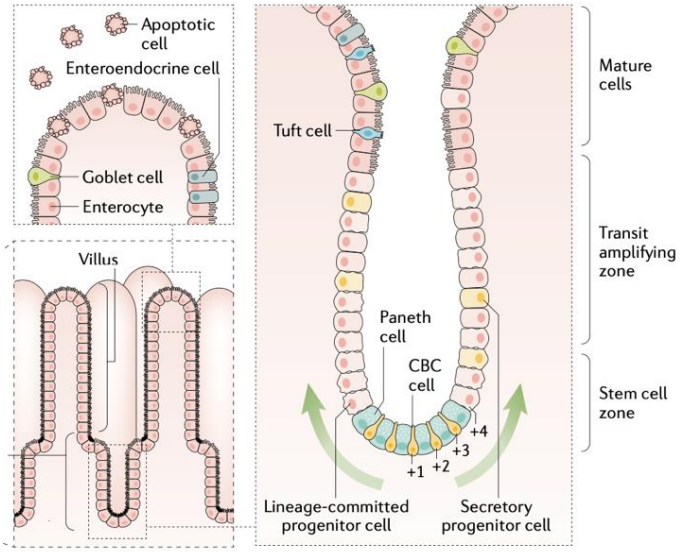


Figure 2 | Crypt-Villus organization in the small intestine. At the bottom of the crypt (stem cell zone) stem cells are maintained by a plethora of signalling cues provided by several specialized cell types that favour, above other outcomes, proliferation. As stem cells move up the crypt there is a shift in the signals towards differentiation into the different intestinal lineages (transit amplifying zone). This area contains cells in transition from stem cell to a differentiated state. The villus (mature cells) are composed of differentiated cells including all cells from the intestinal lineage (Paneth, goblet, enteroendocrine and tuft cells, enterocytes and M cells) that usually after 5-6 days undergo apoptosis and are shed from the epithelium into the intestinal lumen (adapted from (7)).

The presence of villi in the intestine increases the area of absorption in the digestive process but at the same time these structures are subjected to high levels of mechanical stress together with other physical and chemical insults (e.g., pH and temperature changes). In order to avoid or minimize the accumulation of damage in the epithelium, that can lead to infection or transformation of normal cells to neoplastic cells, mature cells that are in the villus are replaced every 3-5 days (23). Intestinal crypts on the other hand are well protected from the abrasive processes of the intestine since they sit on invaginations of the intestinal wall. At the base of the crypt, a pool of rapid proliferating stem cells give rise to all cell lineages found in the intestinal epithelium. These cells are continuously pushed out of crypt by the next generation of epithelial cells while moving towards the upper part of the villus. During their journey, stem cells differentiate into one of six cell types found in the intestinal epithelium, and eventually go into apoptosis and are shed into the lumen once they reach the tip of the villus (**Figure 2**). The mature cell types found in the intestine can be separated into absorptive (enterocytes and M cells) and secretory (Paneth, goblet, enteroendocrine and tuft cells) lineages (7). Each different cell type has a specific role for maintaining gut homeostasis while promoting one of the most important biological processes in any living organism: food digestion. The most numerous cell type in the intestine are enterocytes (~80%) which are primarily responsible for nutrient absorption, including ions, water, sugars, and amino acids (24). M cells (<1%) have a central role in the human adaptive immune response. They are responsible for the uptake and presentation of antigens from the intestinal lumen to the underlying mucosal immune system such as Peyer's patch and other gut associated lymphoid tissues (GALTs). Here, dendritic cells and macrophages can directly activate T and B cells which can be directed to the intestinal lumen by M cells (25). Goblet cells present the first line of defence of the intestinal mucosa. Morphologically they show a goblet, cup-like shape and specialize in the synthesis and secretion of mucus. Their cytoplasm is mostly composed of mucin granulae containing mucins and other mucus components such as MUC2, FCGBP, CLCA1, ZG16, and AGR2 (26). Enteroendocrine cells (EECs) (<1%) are involved in nutrient sensing, and can regulate a range of signalling pathways involved in the digestive process. Additionally, they produce a plethora of gut hormones such as secretin, pancreozymin and enteroglucagon crucial for the processes of insulin secretion and regulation of appetite (27). Tuft cells are a very rare cell type making up for <0.4% of the cells in the intestine. They play an important role in the immune response particularly against helminth parasites (28). Unlike all the other cells types in the intestine, Paneth cells (~3-8%) move down the crypt as they differentiate, and their main functions are both protection and nurturing of stem cells at the crypt base. They are located in small depressions of the mucosa of the small intestine called crypts of Lieberkühn. Moreover, they produce and secrete an abundance of antimicrobial peptides

and immunomodulating proteins including lysozyme, α-defensins and phospholipase A2 that help regulate the composition of the intestinal flora (29).

The cell fate of each lineage is a tightly regulated process that depends mostly on Wnt, Notch and BMP signalling in the crypt (**Figure 3**). Here, approximately 15 stem cells are interspersed with Paneth cells and every stem cell touches at least one Paneth cell (7). The number and position of Paneth cells in the crypt is not random and it can affect the fate of stem cells. In fact, it was reported that stem cells located at the center of the crypt surrounded by Paneth cells were more likely to become the dominant clone of that crypt in compared to cells that only contacted with one Paneth cell located at the border of the crypt (30). Paneth cells secrete WNT3, EGF and the Notch ligands DLL1 and DLL4, which make up the core signalling cues for the maintenance and survival of stem cells (31). Given the close proximity between these two cell types, signal transfer occurs by direct cell-to-cell contact. At the same time mesenchymal cells residing beneath the epithelium also play an important role in maintaining stem cell homeostasis by providing physical support and releasing soluble factors such as Wnt, R-spondins, BMPs, BMP antagonists and cytokines.

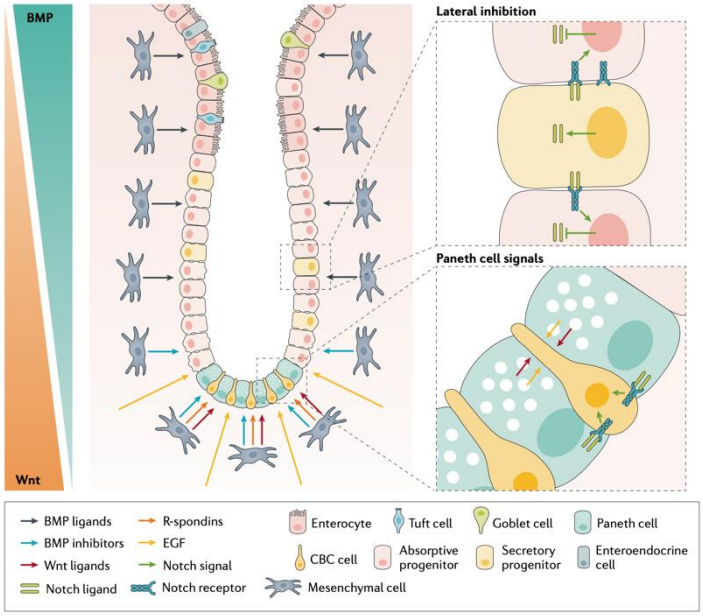


Figure 3 | Signalling cues in the intestinal crypt. Gut homeostasis is a very tightly controlled mechanism that depends on the contribution of 4 main signalling pathways (Wnt, NOTCH, BMP and TGFβ) that ensure a decreasing gradient of proliferation from the bottom of the crypt to the villus typ and an increasing differentiation gradient on the opposite direction. While some signals are provided by diffusion through mesenchymal cells, others like Paneth cells, require cell-to-cell contact for proper signal transduction to occur. Disruption of this balance often leads to disease such as CRC for instance where a mutation in APC leads to persistent activation of Wnt signals causing proliferation to be sustained and preventing differentiation of the stem cells (Adapted from (7)).

A high Wnt signalling is maintained at the bottom of the crypt by Paneth cells and mesenchymal cells through the secretion of WNnt agonists and R-spondin and the strength of the signal decreases as we travel up the crypt. This signalling aims to promote proliferation and maintain a stem-like phenotype at the crypt base. Conversely, a gradient of increasing BMP stands in opposition to Wnt as we follow this direction, in order to stimulate differentiation of stem cells into one of the six mature cell lineages found in the intestine (7).

The aforementioned signalling pathways are of the utmost importance for crypt integrity maintenance, and formation of a healthy intestine. It is therefore not surprising to find them dysregulated in CRC. In order to understand the biological consequences of disrupted signalling in disease, one must first look at the normal physiological setting. The next subchapter will focus on the pathways that are predominantly active in the human intestine, which are subsequently often dysregulated in CRC.

1.3 Signalling pathways in the intestine

1.3.1 Wnt pathway

It is the most important pathway when it comes to regulation of proliferation and differentiation of stem cells at the bottom of the crypt. In its active state, Wnt ligands (such as WNT3a, WNT1 and WNT5a) bind to the Frizzled–LRP5–LRP6 receptor complex located at the cell membrane and inhibit the degradation of β-catenin by the adenomatous polyposis coli (APC) destruction complex. β-catenin is then free to translocate to the nucleus where it binds T cell factors (TCFs) and regulates gene expression (7). At the same time, and also at the cell membrane level, three members of 7-TM receptors, Lgr4, Lgr5, and Lgr6 bind R-spondins with high affinity through their N-terminal ectodomain (32). In fact *Lgr5* was shown to mark adult stem cells in a variety of tissues including the intestine, and it is currently the *bonafide* marker of intestinal crypt stem cells (33). Concomitantly, genetic ablation of *Lgr5* in the intestine leads to disruption of the stem cell compartment (34). RNF43/ZNRF3 are two E3 ligases that act on Frizzled molecules leading to turnover of these receptors and thus serve as negative-feedback regulators of Wnt signal strength. Activation of Lgr receptors leads to clearance of the RNF43/ZNRF3 molecules making the Frizzled receptors available, ultimately enhancing Wnt pathway activation. The absence of Wnt ligands and R-spondins (in the intestine the most important are RSPO2 and RSPO3) (35) renders the pathway inactive. In this context, β-catenin is degraded by a destruction complex composed of the tumour suppressor proteins APC and AXIN, and two constitutively active serine-threonine kinases - CK1a/d and GSK3a/b. These two molecules sequentially phosphorylate Axin-bound β-catenin and mark it for proteossomal degradation, after ubiquitination by b-TrCP (**Figure 4**) (32). At the crypt base, Wnt signalling is provided mostly by Paneth cells in the form of WNT3a, by direct cell-to-cell contact and also from the cells from the mesenchyme (36). Interestingly, intestinal crypts developed an elegant way of controlling Wnt signalling gradient along the crypt regarding their proliferative state. In brief, membrane bound WNT3a in stem cells provided by Paneth cells is halved everytime there is a mitotic division and the cell moves up the crypt. With subsequent divisions, the amount of WNT3a decreases as we leave the crypt. However, this gradient is lost once proliferation stops at

the crypt level. In sum, when proliferation is fast many cells move upwards in the crypt decreasing the amount of Wnt in the stem cell zone leading to reduced proliferation in this compartment. If proliferation is slow, higher levels of Wnt in the crypt favours controlled symmetrical stem cell division, thereby shortening the Wnt gradient. This negative feedback loop regulation of proliferation and Wnt signalling at the crypt is crucial for intestinal homeostasis. However, depletion of WNT3a in the intestinal epithelium show no adverse phenotype in mice (37) which suggests the existence of another source of Wnt that remains to be identified.

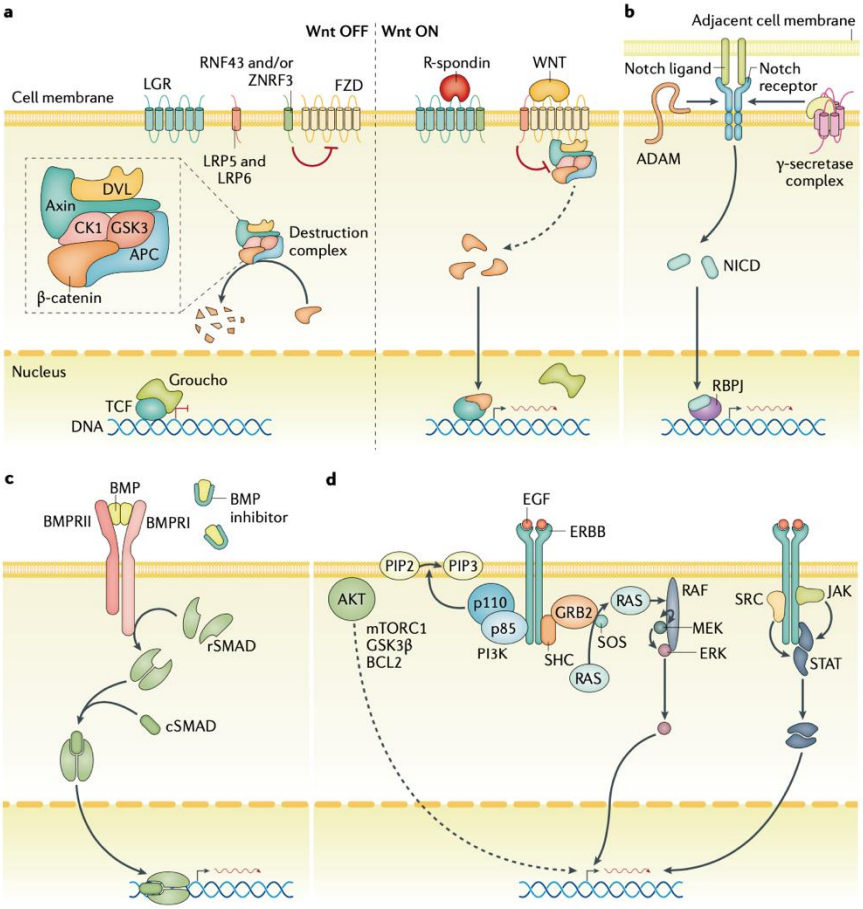


Figure 4 | Main signalling pathways in the intestine. In the absence of Wnt agonists and R-spondin, FZD receptor is degraded by ZNRF3 and RNF43 leading to clearing of β-catenin by the destruction complex (Wnt OFF). In the presence of Wnt ligands, the destruction complex is inactive and cannot degrade β-catenin which is then free to translocate to the nucleus where it dislocates Groucho, binds TCF and drives transcription (Wnt ON) (**a**). Binding of Notch ligands to Notch receptors activates Notch signalling. Notch intracellular domain (NICD) is then cleaved to the cytoplasm and then to the nucleus where it binds RBJP and activates transcription (**b**). Binding of BMP dimerizes type I and II BMP receptors which then phosphorylates and dimerizes rSMADs. These in turn bind common SMADs (like cSMAD) and this complex translocates to the nucleus where it regulates gene expression (**c**). Binding of EGF to its receptor activates the conversion of PIP2 to PIP3 by PI3K, which subsequently activates AKT. AKT activates mTORC1, GSK3β and BCL2 that regulate several crucial cellular processes. At the same time EGF binding activates both MAPK as well as JAK/STAT pathways in order to promote proliferation and counteract apoptosis (**d**) (Adapted from (7)).

1.3.2 Transforming growth factor (TGF)β pathway

BMPs are members of the TGFβ family of ligands, which bind to BMP type I receptors and phosphorylate SMAD1, SMAD5 and SMAD8, also known as receptor-regulated SMAD proteins (R-SMADs). The latter bind to SMAD4 and translocate to the nucleus where they regulate the expression of several genes that trigger differentiation in stem cells, cell cycle arrest in epithelial cells or homeostatic disruption in immune and vascular cells. BMP2 and BMP4 are the most common ligands in the intestine and are provided by mesenchymal cells. BMP signalling at the crypt level acts in order to counteract the proliferative signals provided by Wnt pathway, at the same time promoting differentiation of stem cells to the different intestinal mature cell lineages found in the upper part of the crypt (38). In fact, genetic ablation of *Bmpr1a*, the main receptor for BMPs in the intestinal epithelium of mice leads to expansion of the stem cell compartment resulting in the formation of benign polyps in these animals (39). Because differentiation of stem cells at the crypt base would disrupt the intestinal epithelium, there are a number of factors that act to prevent this from happening, ensuring at the same time that the Wnt/BMP gradient is maintained. The main effectors of this process are myofibroblasts that secrete Gremlin 1, Gremlin 2, Chordin-like 1 and most importantly Noggin that inhibit BMP signalling at the crypt base.

1.3.3 Epidermal growth factor (EGF) pathway

EGF signalling is one of the most important pathways at the crypt level given the variety of cellular processes it controls. Not surprisingly, EGF pathway is commonly disrupted in CRC. Provided by Paneth cells, EGF binds to its receptor ERBB1 which is highly expressed in intestinal stem cells (29). The binding of EGF with its receptor allows binding of phosphoinositide 3-kinase (PI3K) which converts PIP2 into PIP3. AKT is then activated by PIP3, and is able to regulate cell proliferation and survival through interaction with its downstream targets such as mTORC1, BCL2 or GSK3 (7). In fact, PI3K activation was found to be crucial for G1 cell cycle progression of intestinal epithelial cells (40). Further line of evidence suggests that *Pik3ca*, encoding for the class IA PI3K catalytic subunit p110α, is essential for embryonic development. Conversely, *knockout* of *Pik3ca* ablates proliferation in embryonic cells (41).

At the same EGF can activate mitogen activated protein kinase (MAPK) pathway which regulates processes such as cell proliferation, cell differentiation, and cell death in a very conserved manner across different living organisms (42). This signalling cascade starts with the activation of Ras small G protein molecules. It possesses an active GTP-binding and an inactive GDP-binding conformation. In its active form, Ras can activate its

downstream effectors of the pathway. Thus, the signalling continues through sequential activation of three protein kinases known as MAPK kinase kinase (MAP3K), MAPK kinase (MAPKK), and MAPK. Specifically in the ERK cascade Raf proteins act as MAP3K, MEK as MAPKK and ERK proteins as MAPK forming the Ras-Raf-MEK-ERK pathway (**Figure 4**) (43). Activation of MAPK molecules occurs through a series of phosphorylation and dephosphorylation events that leads to conformational changes that culminate in the stimulation of their kinase activity toward the MEK1/2 dual-specificity kinases. Finally, the ERK family of kinases (ERK1/2) are phosphorylated and are able to translocate to the nucleus in order to bind transcription factors that regulate genes that control cell proliferation, differentiation apoptosis and transcription (43). Up to date, several isoforms of Ras molecules were identified, however in cancer the most frequent mutations occur in the *kirsten rat sarcoma viral oncogene* (*KRAS)* gene which are involved in up to 96% of pancreatic ductal adenocarcinomas (PDACs), 52% of colorectal carcinomas, 32% of lung adenocarcinomas, and less frequently in other tumour types (44). The downstream family of Raf kinases is composed of three subtypes: Raf-1, A-Raf and B-Raf which shows the strongest kinase of activity. Moreover, mutations in *BRAF* are reported in about 8% of all human cancers. Mutations in both *KRAS* and *BRAF* lead to constitutive activation of MAPK pathway without the need of an extracellular stimulus (45).

1.3.4 Notch pathway

In the intestine, Notch signalling blocks differentiation of stem cells to the secretory lineage at the crypt base. In fact, Notch is the main regulator of the ratio of secretory to absorptive cells in the intestine. Unlike other molecules, Notch signals are transmitted between adjacent cells and activity in one cell can promote distinct functions in nearby cells. This process, called lateral inhibition, is key for the regulation of stem cell function in the intestine (**Figure 3**) (46). Therefore, Notch signalling requires always a signal-sending and a signal-receiving cell, which have to be physically close. Four Notch receptors (NOTCH 1-4) act as binding sites for five Notch ligands (Delta-like (DLL) 1, 3, 4 and Jagged (JAG) 1, 2). This interaction leads to the proteolytic cleavage of the transmembrane receptor and subsequent release of an intracellular signalling fragment of the receptor NICD (Notch intracellular domain), which then translocates to the nucleus. Here, NICD interacts with the DNA-binding protein RBPJ, and activates the transcription of Notch target genes such as basic loop helix transcription factor HES1 and Olfm4 (47). HES1 represses the transcription of another basic loop helix transcription factor, ATOH1. Downregulation of *ATOH1* expression in the signal-receiving cell represses its own ability to express Notch ligands. In sum, a cell presenting Notch ligands is able to induce Notch signalling in the surrounding

cells but prevents activation of its own Notch signalling. This elegant binary activation system is crucial for the establishment and maintenance of the correct mature cell lineages ratios in the intestine (48). At the crypt base, Paneth cells provide stem cells with DLL1 and DLL4 Notch ligands. When lateral inhibition is removed, stem cells are converted to secretory cells once more showing the role of Notch signalling in intestinal homeostasis maintenance (7). However, the fact that some studies suggested that Paneth cells are not necessary for intestinal epithelial cell maintenance (49, 50) combined with the fact that the colon is devoid of Paneth cells, suggests that other sources of Notch signals should exist.

1.4 CRC carcinogenesis

1.4.1 Molecular pathways

Colorectal carcinogenesis requires a normal cell to accumulate multiple genetic and epigenetic alterations, and establish successive clones, each characterized by a growth advantage. The cell of origin for the majority of CRCs is currently presumed to be a stem cell or stem-cell-like cell residing at the base of intestinal crypts (51). The adenoma-carcinoma sequence model was initially proposed by Fearon and Vogelstein, who suggested that colorectal carcinogenesis is a result of a group of four premises: i) the activation of oncogenes, and/or inactivation of tumour suppressor genes; ii) the formation of a malignant tumour requires the occurrence of mutations in at least four to five genes (whereas fewer changes suffice for benign phenotype); iii) the total accumulation of changes (rather than their order of occurrence) defines the tumour's biologic behaviour; iv) in some cases, mutant suppressor genes exert a phenotypic effect, even in the heterozygous state, which may point to the idea that some tumour suppressor genes may not be "recessive" at the cellular level (52). In sum, this model highlights the disruption of the pathways contributing to intestinal homeostasis that were discussed in the previous subchapter of this thesis. Inactivation of *adenomatous polyposis coli* (*APC)* gene is the initiating event of colorectal carcinogenesis, leading to the formation of adenomas from normal colonic mucosa. This is followed by oncogenic mutations in *KRAS* gene, which promotes adenomatous growth, allelic loss of chromosome 18q, mutation in *PIK3CA* gene and inactivation of deleted in colorectal cancer (DCC) and *SMAD4* genes, both involved in the transforming growth factor (TGF)-β pathway. Of note, *PIK3CA* encodes for the p110 alpha catalytic subunit of PI3K and it is mutated in 20–30% of human CRC. The transition from late adenoma to carcinoma is mediated by inactivation of *TP53* and loss of heterozygosity (LOH) of chromosome 17p (11, 52). Until today this model holds true for most CRC cases (classical pathway). Recent advances in the area of cancer research led to a better understanding of the roles of the different genes and pathways involved in

intestinal carcinogenesis, which in turn led to the need to include new findings to the classical progression model. Mounting evidence shows that most cancer types result from an imbalance between mutation development and the mechanisms responsible for cell cycle maintenance, where epigenetics also play a central role. The lack of the ability of a cell to control and correct mutations in their DNA leads to genomic instability. Three main pathways, rather than a specific sequence of events, have been identified that contribute to this state: chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylator phenotype (CIMP) (11, 53).

Chromosomal instability (CIN)

The CIN, or “suppressor”, pathway accounts for nearly 70-85% of CRCs and it is characterized by gain or loss of whole chromosomes or chromosomal regions, activation of oncogenes (*KRAS* and *BRAF*) and inactivation of TSGs (*APC* and *TP53*) (54). Additional genetic mutations were found also in *TGFBR* family of genes and *PIK3CA* (55). As a result, aneuploidy, chromosomal amplifications and a high frequency of loss of heterozygosity (LOH) events are features of this type of tumours. Causes leading to CIN are currently elusive, however it was suggested they may be related to defects in chromosome segregation, telomere dysfunction or defects in the DNA repair mechanisms (54). Some of the main karyotypic alterations present in CIN-driven tumours are amplifications on chromosomes 7, 8q and 13q, deletions on chromosomes 1, 4, 5, 8q, 18q and 17p, which contains *TP53* gene, and focal gains or losses in regions containing important cancer genes such as *VEGF*, *MYC* and *PTEN*. Nonetheless, only a minority of CRCs characterized by this pathway have a full complement of these molecular abnormalities. Other genetic events might overlap or bypass these steps, in order to deliver the necessary biological advantages to the tumour (11, 56).

Microsatellite instability (MSI)

The MSI, or “mutator” pathway is the other main mechanism for genomic instability in CRC and accounts for 10-15% of sporadic cases (56). Spread throughout the genome, there are short repeat nucleotide sequences of 1–6 nucleotides that are prone to errors during replication, due to its repetitive nature (57). These sequences (microsatellites) are present in both protein-coding and non-coding regions of the DNA. Therefore, replication errors can also occur within the gene sequence. The DNA mismatch repair (MMR) system recognizes and repairs these errors during DNA replication. Members of the MMR system identified so far include *MLH1, MSH2, MSH6, PMS2, MLH3, MSH3,* PMS1 and *Exo1* and, together, they form a strand specific protein complex responsible for the repairing of DNA

errors (58). However, if an abnormality is present in the genes encoding the proteins that make up this complex, these errors will not be repaired. This can lead to MSI due to incorrected mismatch errors that in turn leads to frameshift mutations at specific oncogenes or TSGs such as in *TGFβR2, EGFR* and *BAX* (among many others) (59). In order to provide uniformity of definition of MSI tumours, and aid clinicopathological practice, a standardized panel of microsatellites was created (60). Although some clinical laboratories use different markers (61), a tumour continues to be classified as follows: Microsatellite high (MSI-H) if 30% or more markers show instability; Microsatellite low (MSI-L) if 10-29% of markers are unstable, and Microsatellite stable (MSS) if no marker is unstable (53). Of note, patients with Lynch syndrome, a hereditary disease which confers a high risk for CRC development, show frequently MSI tumours due to germline mutations in MMR genes. On the other hand, sporadic MSI CRC frequently display loss of MMR function due to *MLH1* silencing caused by aberrant DNA methylation (11). Additionally, it was shown that MSI-H sporadic tumours appear to be strongly related with methylation of *MLH1* whilst MSI-L lesions may be more associated with *MGMT* methylation, *KRAS* mutation and low frequency of 5q LOH (56).

CpG island methylator phenotype (CIMP)

Finally, the CIMP phenotype shows the implication of epigenetics in the CRC carcinogenesis. Epigenetic alterations refer to changes in gene expression or function, without affecting the DNA sequence. In humans this is usually caused either by DNA methylation, histone modifications and miRNA interference (62). DNA methylation can occur at the promoter region of genes in regions called CpG sites. These regions (approximately 40-50% of gene promoters in mammals) contain high levels of cytosine-guanine pairs with phosphate bonds present in the genome, and are unmethylated in normal conditions to allow gene expression. Moreover, methylation at CpG sites can occur physiologically, to silence particular genes, such as in X-chromosome inactivation, while decontrolled methylation may occur pathologically as an important step in carcinogenesis. In CRC, the influence of epigenetics is seen both by global hypomethylation of the genome which is associated with genomic instability and simultaneous hypermethylation of the promoter region of specific genes (63). This provides an alternative mechanism for loss of function of TSGs, and dysregulation of crucial molecular pathways for homeostasis including Wnt, TGFβ and NOTCH (53). According to the CIMP status, CRC can be classified as CIMP high (CIMP-H) (around 20% of CRCs) or CIMP low (CIMP-L), relative to the methylation status of a significant proportion of marker genes previously defined (64, 65). An activating mutation in the *BRAF* gene ($BRAF^{V600E}$) is thought to be the initial step of this pathway resulting in an inhibition of apoptosis in colonic epithelial cells. However, CIMP tumours can also present a mutation in *KRAS* in absence of *BRAF* mutation. This is,

generally, preceded by other alterations such as upregulation of DNA methyltransferases (*DNMT3B* or *DNMT1*) and methylation of *MLH1* promoter that leads to MSI (53, 66). Clinically, CIMP tumours tend to be more proximally located, correlate with female gender and age and are usually poorly differentiated. Molecularly, CIMP tumours are associated with *BRAF* or *KRAS* mutations as well as MSI-H status (11). Conversely these tumours usually do not present *TP53* mutations (67).

1.4.2 Main genetic drivers in CRC

Adenomatous polyposis coli (APC)

The most common mutation in CRC occurs in the *APC* gene, resulting in a truncated form of the protein. The *APC* gene is located on chromosome 5q21-q22 and consists of 21 exons that encode a 310 kDa protein. Inactivation of *APC* by mutation or allelic losses is reported in a high percentage of CRC cases (~80-90%) and causes Wnt signalling to be inappropriately and constitutively activated. In most sporadic colorectal adenomas and cancers, both copies of *APC* are inactivated by somatic mutations, deletions or LOH events (68). Germline mutations in *APC* lead to several hereditary syndromes that increase the probability of developing CRC. For instance, in patients with familial adenomatous polyposis (FAP), germline mutations in one *APC* allele are present in all of their cells, so the initiation of the adenoma-carcinoma sequence, in these cases, follows inactivation of the remaining normal copy of *APC* usually by point mutation (56). Within the first 3 decades of life patients can develop up to 1000 precancerous polyps in the colon that, if not removed (usually by total colectomy), will progress to cancer (69). Loss of *APC* has also been suggested to affect cell adhesion (70) cell cycle progression (71) and DNA repair (72).

Kirsten rat sarcoma viral oncogene homolog (KRAS)

KRAS encodes a GTP-binding protein involved in signal transduction present in many molecular pathways, which is activated in response to extracellular signals. The transition between its active GTP state to its inactive GDP-bound state is controlled by GTPase-activating proteins and guanine nucleotide exchange factors (73). When mutated (mainly in codons 12 and 13 of exon 2) it constitutively activates MAPK signalling pathway, which controls central cellular processes such as growth, differentiation, survival, apoptosis, cell motility and proliferation as previously stated (44). Mutations in *KRAS* are considered an early event in malignant transformation and are present in approximately 30-60% of human CRCs. It is proposed that these mutations may be of importance in the transition from adenoma to carcinoma through activation of downstream targets including *BCL-2, E2F4* and *MMP1*. EGFR-targeting antibodies (cetuximab and panitumumab) are routinely

used for the treatment of metastatic CRC in combination with conventional chemotherapy. *KRAS* mutational status is considered a predictive factor in the response to anti-EGFR therapy. When a mutation is present, *KRAS* and its downstream effectors become independent of upstream EGFR signalling, rendering the therapy ineffective. On the other hand, wt *KRAS* tumours will benefit from anti-EGFR therapy as signalling can happen through the EGFR-KRAS axis and may, thus, be inhibited (74). However, only 30% of wt *KRAS* patients respond to anti-EGFR therapy as mutations in other genes of the *RAS* family (*NRAS*, *HRAS*) and *BRAF* might provide an alternative mechanism for activation of this pathway.

v-Raf murine sarcoma viral oncogene homolog B (BRAF)

The *BRAF* protooncogene is located in chromosome 7 and is composed of 18 exons. The most frequent somatic alteration (~95% of mutated samples) in *BRAF* is a point mutation leading to a valine to glutamic acid change at codon 600 (T1799A encoding $BRAF^{V600E}$), which results in a several hundred-fold increased activity of its protein's kinase domain. This causes sustained activation of the MEK1/2 / ERK1/2 mitogen-activated kinase (MAPK) signalling cascade (75). *BRAF* is mutated in 5-15% mCRC and is associated with poor prognosis (11). Additionally, mutations in this gene are present in sessile serrated adenomas (SSA) and associated with MSI and CIMP phenotypes. Interestingly this same BRAF activating mutation is also commonly found in other cancer entities including melanoma and papillary thyroid cancer which supports a context-dependency for the biological effect of mutant BRAF. This hypothesis becomes stronger when considering that Vemurafenib, a FDA approved drug targeting V600E-mutant BRAF leads to improved rates of overall and progression-free survival in melanoma patients expressing this mutation (76), but shows very limited effect in CRC patients harbouring the exact same mutation (77). Additionally, unlike in the *KRAS*-mutant setting, there is no difference in anti-EGFR therapy´ efficiency (78). Based on these data, *BRAF* inhibitors alone seem to have insufficient clinical activity in patients with *BRAF*-mutant CRC and several clinical trials are ongoing where the effect of the combination of 2 or more targeted drugs or conventional chemotherapeutics is tested (79).

1.4.3 Clinical pathways of CRC

Cancer is, by definition, a genetic disease. Therefore, understanding the molecular events that lead to the transformation of a normal to a neoplastic cell is crucial to predict disease behaviour and to be able to treat it in an appropriate time window. CRC initiation and progression typically follows two distinct pathways based not only on their molecular

features but also on their morphological and clinical presentations. The “traditional pathway” accounts for nearly 70-90% of CRC cases and, in its sporadic setting, follows the classical adenomatous polyp-adenoma-carcinoma sequence discussed in the previous subchapter, initiated by an *APC* mutation (11). Tumours arising from this pathway are usually MSS (15). In 1990 another type of lesions were identified by Longacre and Fenoglio-Presise (80) characterized by a “saw-toothed” infolding of the crypt epithelium. These called “serrated polyps” constitute the precursor lesions of another pathway by which roughly 30% of CRCs can develop: the “serrated pathway” (81). Serrated polyps can be divided mainly into 4 subtypes (Hyperplastic polyps, SSA, traditional serrated adenoma (TSA) and mixed serrated polyps). Usually, cancers arising from the serrated pathway derive from SSAs and TSAs since they can present dysplasia, in contrast with the remaining types of polyps (81). SSAs account for 2-9% of all polyps and are usually located in the proximal colon, while TSAs are even rarer making up for ~1% colon polyps and usually arise in the left colon (82, 83). Serrated lesions usually present CIMP, *BRAF* mutations and to a lesser extent *KRAS* mutations and these are thought to be the initial events in this pathway. In this setting, *BRAF* and *KRAS* mutations are mutually exclusive. Interestingly, 80-90% of SSAs and 30-60% of TSAs express mutant *BRAF*, while *KRAS* is often mutated only in <10% of SSAs and around 80% of TSAs (81, 84). Thus, *BRAF* mutations are more commonly associated with SSAs and *KRAS* mutations to TSAs. Activation of oncogenic *BRAF* and *KRAS* induces, initially, proliferation of intestinal epithelial cells followed by cell senescence. For senescence to be bypassed and proliferation resumed, loss of TSGs such as *p16INK4a*, *IGFBP7*, *MGMT* and *TP53* need to follow oncogene activation. Inactivation of the aforementioned targets often results from methylation at the promoter region in this pathway (85). Wnt pathway dysregulation is also reported in serrated lesions mostly when dysplasia is present. However, a different mechanism than the one observed in CRCs arising from the “classical pathway” was proposed. In the former, β-catenin nuclear accumulation might be achieved by methylation-induced silencing of mutated in colorectal cancer (*MCC*), a gene coding for a protein known to interact directly with β-catenin and regulate Wnt pathway (86). In addition to these alterations, SSAs can also present CIMP-H, highlighting the strong correlation between *BRAF* mutation and the CIMP phenotype (87). In line with this, *MLH1* is commonly inactivated by promoter methylation which, as discussed before, can lead to MSI and fuel tumour progression. On the other hand, oncogenic *KRAS*-driven lesions are usually CIMP-L and MSS (81). According to Jass JR. (2007), serrated lesions can be classified as such: 1) *KRAS* mutant, CIMP-Low, MSS/MSI-Low; 2) *BRAF* mutant, CIMP-H, MSI-H; and 3) *BRAF* mutant, CIMP-Low, MSS/MSI-Low (88).

Currently, the classification of “serrated lesions” remains still a challenge due to the variety of molecular features that characterize these tumours and the difficulty to associate

them to sometimes very similar histological profiles. The fact that different genes can be altered in different clinical presentations of the same “general” disease (CRC) and at different times of disease progression indicates that not just the combination of molecular alterations but also the timing at which they occur is critical for pathway allocation and thus clinical management of the patient. Tumours arising from the “serrated pathway” are clinically more aggressive and associated with therapy resistance (15). Understanding which events lead to the progression of this subset of tumours is of the utmost importance in order to be able to prevent the development of a serrated polyp to advanced serrated adenocarcinoma.

1.4.4 Molecular classification of CRC

CRC can be classified by mainly two molecular pathological classification systems, based on their genetic mutational landscape (55) or gene expression data (89). In the former, using whole exome sequencing (WES), along with DNA copy number variation (CNV) analysis, promoter methylation, messenger RNA (mRNA) and microRNA (miRNA) data from 276 DNA samples from cancer patients, the authors could divide human CRCs in mainly three groups based on their mutation rate – ultra-mutated, hypermutated and non-hypermutated tumours. Hypermutated tumours (~16%) are usually MSI as a result of MMR deficiency, often by *MLH1* promoter methylation, and show a high mutation rate of 12–40 mutations/Mb. Ultra-mutated tumours comprise only about 3% of all CRCs and show an extremely high mutation rate (>40 mutations/Mb) due to mutations in the DNA Polymerase Epsilon or Delta 1 (POLE or POLD1) exonuclease domain (proofreading) mutations (EDM). Inactivating mutations in *POLE* or *POLD1* prevents the cell to correct the misincorporation of nucleotides during DNA replication or repair (90). Non-hypermutated tumours (~84%) are MSS and show a strong association with the CIN pathway of CRC progression, described before. As such, these lesions are characterized genetically by high frequency of DNA somatic CNVs and mutations in genes that lead to the activation of the Wnt, MAPK and PI3K pathway, or inactivation of TGF-β and p53 inhibitory pathways (53, 55).

In 2015 a new classification was proposed by integrating transcriptomics data together with genetic and genomic information of a set of 18 different CRC (**Figure 5**) (89).

CMS1 MSI immune	CMS2 Canonical	CMS3 Metabolic	CMS4 Mesenchymal
14%	37%	13%	23%
MSI, CIMP high, hypermutation	SCNA high	Mixed MSI status, SCNA low, CIMP low	SCNA high
BRAF mutations		*KRAS* mutations	
Immune infiltration and activation	WNT and MYC activation	Metabolic deregulation	Stromal infiltration, TGF-β activation, angiogenesis
Worse survival after relapse			Worse relapse-free and overall survival

Figure 5 | CMS classification of CRC (adapted from (89)).

Briefly, CMS1 tumours often show MSI and CIMP-H features and are commonly associated with *BRAF* mutations and immune infiltration. CMS1 lesions are more often found in the proximal colon and therefore associated with the "serrated pathway" of CRC initiation and progression (91). The CMS2 subtype is the most frequent type and is characterized by an epithelial gene-expression profile. Given its high levels of CIN and Wnt pathway activation, they are recognized as the classical type of CRC. The CMS3 or metabolic group usually does not show CIMP, and is more associated with *KRAS* mutations and with an increase in metabolic pathways. Finally, the CMS4 mesenchymal type highlights the influence of the tumour microenvironment in cancer progression. These lesions show the presence of a large stromal compartment, with a strong epithelial-mesenchymal transition (EMT) signature, as well as angiogenesis, TGFβ signalling and matrix remodelling at the transcriptomics level (89, 91).

Although this revolutionary classification system deepened the understanding of the cell-intrinsic and extrinsic pathways that fuel CRC it is important to note that roughly 13% of cases cannot be assigned to any CMS subtype since they represent either mixed or intermediate profiles. The concept of tumour heterogeneity tells us that a given tumour can have different regions as well as single cells within the tumour that can be classified into different subtypes (92). Therefore, sampling of the tumour might include regions with different subtype affiliations. Sampling bias is a methodological artifact that might hamper the interpretation of results not only in the research setting but also in the clinical management of cancer patients (93). It is therefore crucial to bear in mind that tissue sampling might give us only a part of the bigger picture and thus not capture the full spectrum of features that characterize that particular lesion.

More recently and using a different approach, a new CRC classification also based on transcriptional profiling was proposed (94). The authors performed patient derived xenografts (PDX) (n=515 samples from 214 patients) using both primary tumours and

metastases into mice with the aim of reducing the stromal contamination inherent to many human tumours. Mouse tumours were then transcriptionally profiled in order to find gene signatures that would allow the separation of samples into different groups. Using a set of 565 genes the authors successfully divided this cohort into 5 “CRC intrinsic subtypes” (CRIS) subgroups: CRIS-A tumours are enriched for either *BRAF* mutations and MSI phenotype, or *KRAS* mutations and MSS phenotype; CRIS-B lesions are associated with poor prognosis and high TGFβ signalling (although unrelated to the CMS4 type from (89)); CRIS-C is associated with tumours that are dependent on EGFR signals and that are sensitive to anti-EGFR treatment; CRIS-D more associated with Wnt pathway activation and *IGF2* overexpression; CRIS-E includes CIN positive tumours often mutated for *KRAS* and *TP53* (94).

1.5 Model systems to study CRC

1.5.1 Genetically engineered mouse models (GEMM)

To better understand and study CRC, a plethora of mouse models were generated over the years that allow, for instance, the study of cancer in a chemically induced setting, or by tissue-specific activation of a given oncogene, or loss of TSG (95). Since this has been extensively reviewed elsewhere (95, 96), in this section only the mouse models used in the scope of this thesis will be addressed. The generation of the first genetically engineered mouse model (GEMM) to model CRC included the application of N-ethyl-N-nitrosourea that leads to nonsense mutations in codon 380 of the *Apc* gene (Apc^{Min}). Apc^{Min} mice develop a large number of adenomas in the small intestine after 4-5 months (97). This model was considered a milestone in cancer research and many groups still use it. Several additional models were generated over the next years that made use of the *Cre-loxP* system allowing the tissue-specific and conditional *knockout* of TSGs or activation of oncogenes (98). Of note, Cheung *et al.,* generated a CRC GEMM by deleting exon 1-15 of the *Apc* gene (hereby called $Apc^{+/fle1-15}$) leading to its complete deletion. When compared to Apc^{Min} mice, $Apc^{fle1-15}$ mice develop more and larger tumours which reflected a shorter survival in the latter group (5). This setting provides a clean genetic model in a way that a phenotype is given specifically by the genetic insult elicited in these animals. In order to avoid abundant distribution of mutations in the whole organism, and to be able to study the effect of a gene in a specific tissue, it is possible to direct transgene expression by a tissue specific gene promotor, thanks to the *Cre-loxP* system (98). By using a Cre recombinase that is under the IEC-specific promoter *Villin* (hereafter referred to as *Vil-Cre*) recombination can be effectively activated specifically in the intestine (99). Thus, *Vil-Cre*; $Apc^{fle1-15}$ mice recapitulate key genetic features of the classical pathway of CRC initiation and progression.

The effect of the Cre recombinase can also be temporally controlled using a *Vil-CreER^{T2}* line which activates recombination only in the presence of tamoxifen (100). Since the generation of *Apc*-deficient mouse models a large number of additional models were generated to interrogate the effect of a specific gene or pathway, or the combinatorial effect of different genes in CRC (96).

However as previously discussed, CRC can arise through different pathways which have underlying distinct initiating events independent of *APC*. In an effort to study the route of progression of cancers arising through the "serrated pathway" of CRC initiation *in vivo*, models expressing mutated *Kras* (8) or *Braf* (9) in intestinal epithelial cells were generated. In fact, the progeny of *Villin-Cre* mice crossed with *Kras$^{G12D/+}$* mice (*Vil-Cre; Kras^{G12D}*) showed elongation of the intestine and hyperplastic transformation uniformly spread throughout both small (SI) and large intestine (LI) with a similar histology to the one found in human serrated precursor lesions. Concomitantly, after 1 year, 25% of mice develop ~1 adenoma in the SI and no tumours were observed in the colon (8). Interestingly, *p16ink4a* is upregulated specifically in the LI and not in the SI as a result of *Kras* activation, leading to oncogene induced senescence (OIS). OIS is a form of stable growth arrest that can occur *in vitro* (101) and *in vivo* (101) and can explain the absence of colon tumours in this mouse model. Remarkably, genetic ablation of *Ink4a/Arf* in a *Kras* mutant background (*Vil-Cre; Kras$^{G12D/+}$; Ink4a/ Arf$^{-/-}$*) led to tumour formation in the colon in half of the mice, with most of the cases showing features of invasive adenocarcinomas. Consistent with the status of *KRAS*-mutant human tumours, also the lesions from *Vil-Cre; Kras$^{G12D/+}$; Ink4a/ Arf$^{-/-}$* mice were MSS/MSI-L, CIMP-negative and were shown to grow independent of Wnt pathway deregulation since no genetic alterations in *Apc* or *Ctnnb* were present (8).

Another elegant mouse model that greatly improved the knowledge of the biology behind serrated intestinal lesions was published in 2013 by our group (9). Here, the authors generated a Cre-inducible *Braf knockin* allele, that leads to the production of the V637E mutant BRAF mouse protein. This genetic variant (*Braf^{V637E}*) in mouse exon 18 is at the orthologous position of the human *BRAF^{V600E}* mutation in exon 15, which is the most common *BRAF* mutation in human tumours (102). Crossing mice harbouring this mutation with *Vil-Cre* expressing mice allowed the expression of the *Braf^{V637E}* (*Vil-Cre; Braf$^{LSL-V637E/+}$*) mutation specifically in the intestine. As for the case of *Vil-Cre; Kras$^{G12D/+}$* mice, also the *Braf* mutation leads to intestinal hyperplasia that spans all along the intestine. Furthermore, after 10 months all animals develop adenomas, the large majority of them in the SI, with features of human TSAs. Fourteen percent of mice had invasive carcinomas at this timepoint. Human TSAs, as most of sporadic intestinal lesions develop in the colon and are mostly associated with *KRAS* mutations, although *BRAF* mutations can also occur. One explanation for this might be the fact that most tumours in these mice develop in the SI

which might lack the necessary morphologic features for SSAs to develop. On the other hand, 40% of the *Braf* mutant mouse tumours were MSI-H reflecting the human clinical situation (64). Interestingly, Wnt pathway seems be activated progressively when comparing tumours (adenomas and carcinomas) with hyperplastic tissue. By analysing the mutation spectra of *Braf*-induced tumours the authors showed that tumours often harbour mutations in known Wnt pathway genes, including *Apc*, *Ctnnb1* and *Lrp1b*, suggesting that Wnt pathway dysregulation is an early and necessary event in *Braf*-driven tumorigenesis. Furthermore, genetic inactivation of *Trp53* in this model further increased the percentage of carcinoma-bearing mice from 14% to 56% (9) indicating that *Trp53* does not affect tumour initiation but has a role in tumour progression to invasive states. Concomitantly, inactivation of *Cdkn2a* in addition to mutant *Braf* leads to similar results as the aforementioned.

PIK3CA mutations occur in 20–30% of human CRC. p110α, encoded by *PIK3CA*, is the catalytic subunit of the ubiquitously expressed class IA PI3Kα and 3 main mutations (E542K E545K and H1047R) have been identified in a variety of cancers (103). Specifically, transgenic expression of H1047R mutation has been shown to be tumorigenic in different cancer entities (104-106). More recently this was also confirmed in a colon tumour mouse model where the expression of *Pik3ca*^{H1047R} is controlled by a Cre recombinase specifically expressed in the distal part of the intestine and colon (107). To interrogate the effect of this mutation in murine intestine in a comparable way to the aforementioned mouse models (*Vil-Cre; Apc*$^{fle1-15}$, *Vil-Cre; Kras*^{G12D} and *Vil-Cre; Braf*^{V637E}) we generated a new mouse model where mutant *Pik3ca* is selectively expressed in the small intestine and colon (unpublished model). We generated, previously in the lab, an oncogenic *Pik3ca*^{H1047R} (encoding p110α^{H1047R}) allele silenced by a lox-stop-lox (LSL) cassette as a *knockin* at the mouse *Rosa26* locus (hence LSL-*Pik3ca*^{H1047R} mouse line) (106). To activate the expression of p110α^{H1047R} and thus PI3K signalling, specifically in the intestine, these animals can be crossed with *Villin-Cre* expressing mice (*Vil-Cre*; *Pik3ca*^{H1047R}). The resulting mice from this crossing develop widespread hyperplasia in the small intestine and tumours with similar morphology to what is observed for the other serrated models included in the study.

All the aforementioned GEMMs are based on the activation of a genetic feature in the germline, which means that the oncogene activation or loss of TSG starts at the birth of the animal. In human sporadic CRC, mutations are acquired somatically in intestinal stem cells and cancer usually develops within 10-15 years (15). Therefore, these mouse models do not reflect the human counterpart of sporadic tumour initiation. A way to overcome this constraint is given by the use of CreER^{T2} recombinases. Here, Cre is fused to mutated hormone-binding domains of the estrogen receptor and are, by default, inactive. However, they can be activated by the synthetic estrogen receptor ligand 4-hydroxytamoxifen (also known as Tamoxifen (TAM)), allowing for external temporal control of Cre activity (108).

Thus, by crossing the *Vil-CreER^{T2}* allele with either *Apc*$^{fle1-15}$, *Kras^{G12D}*, *Braf^{V637E}* or *Pik3ca^{H1047R}* -expressing mice it is possible to somatically activate the genetic alteration after the birth of the animal, upon TAM treatment.

Although the use of GEMMs has given us unprecedent knowledge of the effect of specific genes (and combinations of genes) in CRC initiation, and progression, it is important to be aware of certain limitations and constraints of their use. Cancer develops through an initial mutation and stepwise accumulation of additionally genetic or epigenetic insults. Mouse models can only partially reflect this by the combination of a constitutively active mutation together with an inducible one. Moreover, the number of mutations one can induce in a mouse model are limited, mostly due to the reduced lifespan resulting from the phenotype (*Car1$^{CreER/+}$; Apc$^{fl/fl}$; Kras$^{LSL-G12D/+}$; Trp53$^{-/-}$; Smad4$^{fl/fl}$* mice, for instance, show a lifespan of approximately one month). Finally, GEMMs are time consuming and extremely expensive to generate and maintain. In order to achieve the desired genotype, extensive breeding is necessary which also yields a lot of “reject” mice that are neither used for further breeding nor research purposes. This becomes even more evident in the models described above as mice with one driver event (*Kras* or *Braf* activation, *Apc* loss) take usually more than one year to develop tumours (5, 8, 9).

1.5.2 Orthotopic transplantation models

There are a variety of different transplantation mouse models in regards to the nature of the implanted sample (mouse vs. human) but also to the main experimental question (e.g. immunocompetent vs immunodeficient mice). A syngeneic tumour transplant is characterized by tumour tissue or cells engrafted in mice from the same strain, whereas a xenogeneic transplant implies that the transplantation is done in mice from different strains or using material from different organisms (e.g. human origin) (96). Transplanting tissue from a different organism or different mouse strains into a mouse host would elicit an immune reaction against the foreign material, ultimately causing tissue rejection. To overcome this issue, immunodeficient mouse strains have been generated, such as *Rag2$^{-/-}$, Il2rg$^{-/-}$* (RagIL) or *Prkdc^{SCID}, Il2rg$^{-/-}$* (NSG) mice. NSG mice have a severe combined immune deficiency mutation (*scid*) in the DNA repair complex *Pkrdc* which leads to B and T cell deficiency, and genetic ablation of *Il2* receptor common gamma chain, leading to dysfunctional NK cells. RagIL mice have both the *Rag2* and *Ilrg2* genes deleted which leads to the same phenotype as in NSG mice. Several groups have shown the feasibility of orthotopically injecting CRC cells (109-111) or tumour tissue (112) into the caecal wall of recipient mice. The origin of the tissue can be either from mouse or also from cancer patients (PDX). These models often lead to the development of distant metastases making

it a suitable model to study advanced cancer. However, it should be taken into account that since human CRC occurs in the colon almost exclusively, these models cannot be considered as genuinely orthotopic. Additionally, the implanted material is in contact with the abdominal cavity, therefore it cannot be ruled out that metastases formation in other organs might be due to intra-abdominal cell spillage. Lastly, these methods require opening of the abdomen of the animal which is an invasive procedure and might cause associated morbidities to the animal. To address these negative points associated with caecal implantation, colonoscopy-based transplantation techniques have been developed in the past decade allowing injection of cancer cells directly in the colon of mice with the help of a needle (2, 113-115). Here, scientists were also able to induce metastases upon CRC cell transplantation in the colon submucosa of mice, in a solution often containing Matrigel to avoid cell spillage. Moreover, surgery is not needed for this procedure and mice are simply anesthetized for some minutes while colonoscopy takes place (2).

Another type of orthotopic model is currently used without the need for injections in the colon of recipient animals. Here, tumour cells can be inoculated in the colon by an enema approach and were also shown to generate local tumours and distant metastases (116, 117). However, this model requires mucosal injury (inflammation) previous to inoculation of the cells which also does not quite recapitulate most cases of CRC. In fact, a systematic comparison of all 3 orthotopic models (intracaecal injection, submucosal injection and rectal enema) revealed the submucosal injection model to be the most efficient in primary tumour and metastases formation as well as the one with less adverse effects to the recipient mice (118).

1.5.3 Organoids

Cell lines have been extensively used over the past decades and provided us with not only crucial knowledge about the biology of a wide variety of cancer types, but also with a model through which disease could be modulated *in vitro*. For many years, 2D cancer cell lines were the mainstay of cancer research and although still very commonly used, their application has several drawbacks. For instance, the generation of primary cell lines from human tissue is an extremely inefficient process that requires extensive adaptation and selection of the cells to culture conditions. Moreover, in many cases, the cell line does not reflect the heterogeneity from the patient´s tumour due to the fact that only some clones are able to survive and expand *in vitro*. Lastly, it is not possible to establish primary cell lines from normal healthy tissue specimens making it impossible to have normal tissue-derived control cell lines (119). In 2009 however, Sato et al., completely revolutionized the field of cancer research by describing a novel 3D culture system in which intestinal stem cells were

cultured in Matrigel and could grow as organoids (120). Matrigel is a solubilized basement membrane prepared from a mouse sarcoma tumour cell line very rich in extracellular matrix (ECM) proteins such as collagen IV, laminin and entactin that serves as scaffold for organoids in culture. Although this protocol was first established for mouse tissue, 2 years later it was also described for human (121). In brief, the process includes the isolation of the $LGR5^+$ adult stem cells located at the base of the intestinal crypt which are then embedded in Matrigel and provided with the necessary niche factors mimicking the *in vivo* stem cell niche. By supplementing the serum-free medium with EGF, Noggin and R-spondin (thereby called ENR medium), stem cells can be grown into organotypic polarized epithelial structures with a proliferative crypt and differentiated villus compartments, in resemblance to the human/mouse intestine (120). Additionally, not only the presence but also the percentage of each mature cell lineage found in the intestine is preserved in this 3D system. Another great feature of intestinal organoids is their plasticity regarding cell composition. It was shown that the percentage of a specific cell lineage varies within an organoid and can be modulated by culture conditions. For instance, addition of Wnt3a to the basic culture medium (WENR) leads to enrichment of stem cells in organoids (122). Another example of the feasibility of this approach comes from the study of Yin and colleagues (123) where the authors showed that the number of secretory cells in organoids could be increased by the addition of the NOTCH inhibitor DAPT to the culture medium. This feature was greatly explored in recent years to uncover additional roles of the different lineages of mature intestinal cells, and more importantly the effect of this transition in the stem cell pool (119).

Up to date, organoid lines have been established from multiple mouse and human epithelia origins including liver (124), pancreas (125), stomach (126), lung (127), and breast (128) in addition to colon (120, 121). Because organoids can be cryopreserved and remain genetically and phenotypically stable in culture, the generation of living biobanks is very attractive not just in basic research but also in the clinical setting (129). In line with this, organoids have been used as avatars for predicting individual patient therapy responses rendering possible, at the same time, the development of novel cancer therapies. In the seminal study from Vlachogiannis, G. et al., (2018) the authors showed that patient derived organoids (PDOs) recapitulate patient responses to therapy thereby refining the concept of personalized medicine (130). Additionally, organoids derived from different CRC subtypes require different culture conditions in order to be established, which may reflect the different mutational landscapes of the primary tumour (131). In fact, the genetic profiling of PDOs allows for the correlation of mutations and drug responses (129) deepening at same time the knowledge of how tumours are genetically composed. By culturing single stem cells from tissue pieces derived from different parts of the same tumour Roerink S. et al., (2018) were able to compare mutation load in cancer cells compared to normal cells from the same

individual. Additionally, valuable information could be retrieved regarding the intra tumour heterogeneity of different lesions (132). In sum, cancer organoids exhibited higher mutation burden in comparison with normal colorectal stem cells and within the same tumour, there are different mutational processes taking place in different locations. Remarkably, drug treatment of different clones from the same tumour showed distinct responses to the same drug which clearly underlies the impact of tumour heterogeneity in therapy response (132). As seen in the clinic, therapy often selects resistant clones while killing the sensitive parts of the tumour. This selective sweep renders the tumour unresponsive to the current treatment, and other therapy regimens have to be applied. This poses one of the biggest problems in clinical management of cancer patients - acquired therapy resistance - and the likely the reason why tumours cannot be completely eradicated from the patient. In fact, mounting evidence points to the fact that the acquisition of mutations happens already in the normal tissue, and that “clones” might already be forming before malignant transformation. Strikingly, in a study developed by Blokzijl et al., (2016) whole genome sequencing (WGS) was performed in normal clonal normal organoids derived from both SI and LI and liver from donors of various ages. (133). Remarkably, the mutational signatures seen in intestinal organoids which were characterized by a high abundance of deamination-induced mutagenesis resulting in C>G and T>A at CpG sites, differed from the ones in the liver. Moreover, these signatures were also shown to be present in common CRC driver genes (133).

Another advantage of the use of organoids is the possibility to genetically edit their genome. Recent studies using CRISPR/Cas9 genome editing technology showed that organoids harbouring mutations in *APC*, *KRAS*, *TP53* and *SMAD4* grew independently of medium factors (Wnt, EGF, R-spondin and Noggin). Furthermore, upon subcutaneous or kidney capsule transplantation of these cells in mice, the organoids grew as invasive carcinomas, however with no metastasis formation in the liver or lung (134, 135). In order to bypass these constraints and to recapitulate disease progression Fumagalli et al., (2017) showed that by injecting the aforementioned edited organoids in the caecum of recipient mice instead, metastasis would form both in the liver and lung (111). This is further supported by the study of de Sousa e Melo et al., (2017). By orthotopically injecting mouse *Lgr5*$^{DTR/eGFP}$, *Apc*^{min}, *Kras*^{G12D}, together with *Trp53* and *Smad4* (by CRISPR/Cas9) mutant organoids (AKVPSL), into the submucosa of recipient animals, the authors concluded that genetic ablation of *Lgr5* in cancer stem cells would halt primary tumour formation and metastases formation (136). Conversely, no metastases were formed when mutant organoids (AKVPS) were injected subcutaneously in recipient mice confirming previous studies (134, 135). Besides showing once again the importance of the native environment for proper tumour development and metastasis formation, this study also showed the central

role of cancer stem cells not only for the initiation but also for disease progression (136). The strength and feasibility of using CRISPR/Cas9 engineered organoids in combination with an orthotopic transplantation model to study CRC is further shown in the seminal study from Roper J., et al (2017). Here the authors tested different approaches for *knocking* out CRC driver genes in different genetically altered mouse and human organoid lines, and showed that the main stages of tumour progression could be recapitulated by their model (2). Additionally, in this study the authors proved once more the strength of the orthotopic injections model from a variety of samples including organoids (both from mouse and human), cell lines and even viruses (2).

The establishment of organoids from transgenic mice harbouring an oncogenic mutation or loss of TSG provides an additional tool from which the effect of a single gene can be addressed in a clean genetic background. For instance, *Cdx2* was found to cooperate with mutant $Braf^{V600E}$ in a mouse model of serrated CRC by analysing organoids derived from wt, $Cdx2^{-/-}$, $Braf^{V600E}$ and $Cdx2^{-/-}$; $Braf^{V600E}$ mice (137). Concomitantly, treating human $BRAF^{V600E}$ mutant organoids with TGFβ drives the progression from adenoma to CRCs from the mesenchymal subtype (CMS4) which have a very poor prognosis (138). Additionally, organoids can be used to study the molecular consequences of mutations found in rare subsets of human tumours and attempt to block them using therapeutic drugs as shown by Yan and colleagues (139).

Although organoids set up a new standard for cancer research and greatly complement previous existing *in vitro* and *in vivo* models, they are however, not devoid of caveats. Besides the high costs involved in organoids maintenance (e.g. Matrigel, culture medium) they lack stroma and other components of tumours and normal tissue such as immune cells and blood vessels. To bypass this limitation, several groups explored the possibility of co-culturing organoids with other cell types including fibroblasts (140), immune cells (141) and even bacteria (142) and have shown promising results. Additionally, it is important to bear in mind that all tumour organoids cultures discussed before are of epithelial origin and therefore derived solely from adenomas or adenocarcinomas. Recent research has focused on trying to establish organoids from other types of cancer (143), however the feasibility of this approach remains elusive. When establishing an organoid culture from a tumour tissue, it is often very difficult to completely exclude that normal cells or parts of the normal tissue are also isolated. The generation of this heterogeneous culture can pose a problem in a long-term setting. In fact, it was reported that when cultured together with cancer organoids, normal colonic organoids will outcompete cancer cells possibly owing to apoptosis of the latter as a result of genomic instability and higher rates of mitotic failure (129). On the other hand, in an elegant study published by Garcia AK et al., (2021) the authors cultured normal tissue organoids together with tumour organoids and

saw that wt cells were actively eliminated by tumour cells. Cell competition within the culture leads to apoptosis of wt cells in a process driven by activation of JNK in distinct genetic CRC models (including *Apc* and *Kras*-driven) (144). Interpretation of these somehow opposing results must include a meticulous look into the models used in the experiments. Some tumours show different proliferation and genomic instability rates which, as reported, will impact their behaviour in culture, as most likely proliferative cancer cells can efficiently eliminate normal cells. Therefore, not only the genetic features of the tumours such as *Apc* loss or *Notch1* overexpression (144) but also the stage of the tumour used in the study have to be carefully assessed. More studies of this kind, including different type of tumours with different driver mutations (e.g. *Braf*) will definitely deepen the information regarding tumour cell dynamics in the presence of other cells, thus giving a more educated hint of the *in vivo* situation. These results combined further highlight the importance of a culture composed of pure tumour material to avoid cell competition, thus maintaining a stable and homogeneous culture.

1.5.4 Tumour progression

The majority of the efforts in the field of cancer research focus on dissecting the mechanisms responsible for the transformation of a normal to a cancer cell and the downstream events that lead to tumour progression. Understanding how tumours evolve and the underlying biological events leading to it has given us the opportunity to develop therapies that greatly increased the lifespan of cancer patients and in some cases even cured disease. The notion that cancer is an evolutionary system was first proposed by Nowell in 1976 (145). Concomitantly, the first genetic progression model was postulated for classical CRC by Fearon and Vogelstein more than 30 years ago (52) and since then the definition of cancer as a genetic disease was established. Following that, several seminal studies focused on understanding how tumours evolved and the identification of the mechanisms that lead to this (146-149). Owing to the technological breakthroughs of sequencing platforms, the knowledge of how tumours are composed and how cancer cell dynamics shift during progression is much greater than ever before. Multi regional sequencing of tumours revealed extensive levels of intra-tumour genetic heterogeneity (ITH) formed by branching evolutionary processes (92, 147). The concepts of “founder” and “progressor” mutations were then suggested. Four main types of clonal evolution were suggested to take place during cancer progression based on the dynamics of changes of clone frequencies (6) (**Figure 6**). Even though these models relate to different events and tumour growth patterns, they can happen at distinct times within the same tumour in response to a stimulus that alters tumour behaviour (e.g. therapy or surgery). By combining

the sequencing results of different regions of the same lesions with mathematical modelling approaches, Uchi et al., (2016) concluded that advanced heterogenous cancers exhibit a set of “founder” mutations that are clonal and are present in all regions of the tumour (that were also driver) while “progressor” mutations, mostly of subclonal nature, contributed to the formation of a heterogenous mutation profile that did not offer any proliferative advantage to the tumour (“neutral evolution”) (147). However, ITH is not an exclusive feature of advanced cancers as it starts already in early stages of CRC (146). By applying the strategy of multiregional sequencing to early lesions Saito T. et al., (2018) report that, in contrast to advanced tumours, several more driver mutations in the “progressor” group were detected which supports the notion of “Darwinian selection”. This model is based on the assumption that mutations are acquired over time in a sequential fashion that give rise to different subclones within a tumour and follows, therefore, the same assumptions as the model of linear evolution. These will then compete for nutrients and metabolites and face all sort of selective pressures (such as physical anatomical barriers or therapy) and only the “fittest” clones will be able to remain as part of the tumour while others will perish (6). Combined, the results of the aforementioned studies concluded that CRC progresses according to 2 temporally distinct events: “Darwinian selection” in early stages and “neutral evolution” in advanced stages. The mechanism responsible for this temporal shift was later identified by looking at the copy number variation (CNV) profile of tumours at different stages of progression (146). Whole-genome doubling, chromosomal chromoplexy and chromothripsis represent examples of single catastrophic events that are more common in advanced lesions and can lead to CNV and chromosomal rearrangement events. These events can occur in a short time window and confer a marked fitness increase on a given subclone which would expand to constitute the tumour mass uniformly (“punctuated evolution”) (150). In sum, according to this model, multiple subclones are generated by mutation acquisition and rounds of clonal expansion (“Darwinian selection”). Over time, one specific clone will acquire a driver CNV event in addition to a sufficient set of driver mutations that will most likely provide the tumour with more aggressive features including invasion and immune escape (“punctuated evolution”). This subclone will eventually grow as the dominant clone which has the capacity to invade or resist to therapy. As the tumour grows, several subclones are then generated by the accumulation of neutral mutations (“neutral evolution”) (151) (**Figure 6**). However, although this proposed model is in line with the classic “Vogelgram” (52) and the recently proposed “Big Bang model” for CRC evolution (148) it seems to be valid only for highly aneuploid MSS tumours (most likely arising through the CIN pathway). Typically, MSI tumours show higher mutation rates and lower CNV events which suggest that other mechanisms of tumour evolution take place in this CRC setting. In the scope of the serrated pathway of CRC (mostly associated with MSI status,

KRAS and *BRAF* mutations) (81) sequencing technologies enabled the genetic characterization of different tumour stages and the identification of the mutations and epigenetic events that are present in those lesions. Nonetheless, the biological events that drive tumour progression remain elusive and no evolutionary model has been suggested for these particular CRC pathways.

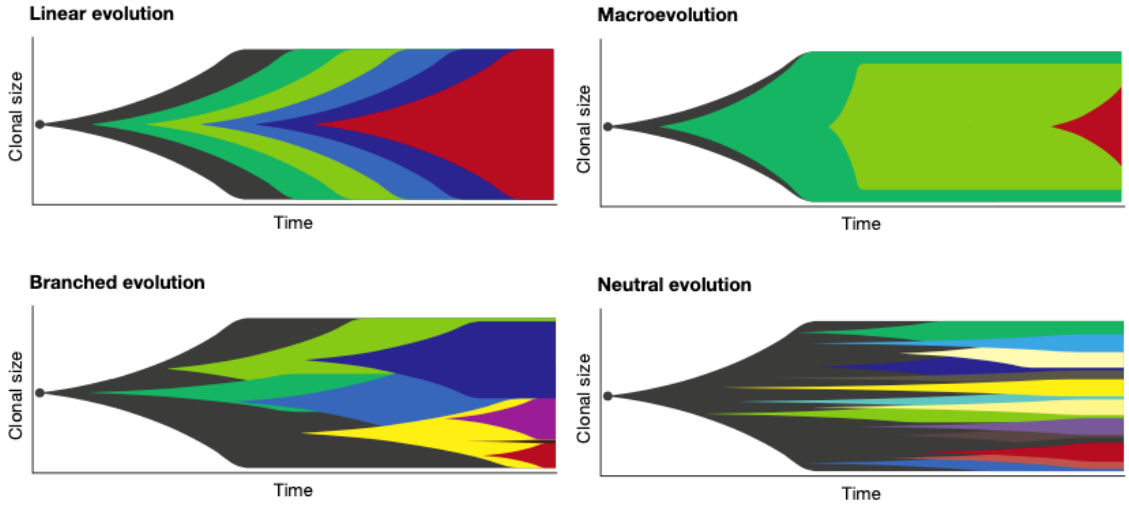


Figure 6 | Proposed models of tumour evolution. *Linear evolution* – The acquisition of new driver mutations provides a strong selective advantage that allows a cell clone to outcompete all other clones within the tumour by a selective sweep. *Branched evolution* – Multiple clonal lineages are generated by acquisition of mutations that increased clonal fitness. These clones diverge from the common ancestor and evolve in parallel within the tumour mass. *Macroevolution* – This model proposes the occurrence of a large number of genomic aberrations in short bursts of time in early stages of tumour progression. ITH can be high at the early stages of tumour evolution but then until a dominant clone is established in the lesion. *Neutral evolution* – It can be seen as an extreme case of branching evolution where no selection is present. Here, intratumour heterogeneity (ITH) is seen due to the accumulation of many random mutations. (Adapted from (6)).

In other cancer entities, great efforts have been put in order to understand how tumours evolve. Of note the work of Mueller S. et al., (2018) from our lab has given a deeper understanding of the molecular evolution of pancreatic ductal adenocarcinoma (PDAC) by showing that an increase in gene dosage (iGD) of mutant *KRAS* drives early disease progression as well as metastasis formation (152). During disease progression tumour cells acquire additional oncogenic gains in *Myc*, *Yap1* and *Nfkb2* that collaborate with mutant *KRAS*. The authors further identified concomitant genetic alterations that are responsible for *KRAS* iGD events including most frequently loss of *Cdkn2a* and other TSGs such as *TP53* and genes from *TGFβ* family. Strikingly, by combining genetic and transcriptomics data together with morphological aspects of tumours and respective cultured cells from *Kras*-driven PDAC mouse models, it was possible to group samples according to its progression profile thereby unveiling how progression takes place in distinct genetic backgrounds of the same disease (152). Interestingly the mechanism of *Kras*-driven tumours seems to extend to other cancer entities including breast cancer (153). By being

able to titrate the expression levels of *Kras in vivo*, a study showed that high levels of *Kras* induction leads to a senescent phenotype driven by p53–Ink4a–Arf dependent senescence checkpoint. On the other hand, low levels of *Kras* drives hyperplasia development in the mammary glands of these animals. Surprisingly, continuous low-level activation of *Kras* led to tumour formation *in vivo* where unexpectedly a strong upregulation of *Kras* expression was reported. Concomitantly, in these tumours p53 was inactivated. Together these results suggest that in order for tumours to form, chronic increased levels of *Kras* expression drive cell transformation only in the absence of functional senescence inducing genes such as *TP53* or *Cdkn2a* (153).

2 Aim of the study

From a molecular standpoint, CRC is defined by inactivation of TSGs such as *APC, TP53* and *SMAD4* and activation of oncogenes including *KRAS, PIK3CA* and *BRAF* (11). The serrated pathway of CRC constitutes a discrete subset of aggressive and heterogenous neoplasias associated with an adverse prognosis (84). Although the main genetic events that initiate these lesions have been identified, the molecular evolution of these cancers is poorly understood. Moreover, targeted therapy for specific initiating genetic mutations (such as $BRAF^{V600E}$ mutation) have been proven efficient in some cancer entities (e.g. melanoma) but not for CRC patients (77). This suggests that the same genetic initiating events might have different effects regarding the cancer context we look at as well as the mechanisms that cooperate with it for tumour progression. As such, the main aim of this study is to deepen the understanding of serrated colorectal tumorigenesis and to identify the molecular events that cooperate with the genetic driver mutations that define central aspects of murine serrated CRC. For that, we defined the following specific aims:

- Generation of GEMMs expressing the main genetic features of both classical and serrated CRC (loss of *Apc,* oncogenic activation of *Kras, Braf* and *Pik3ca*).
- Establishment of a comprehensive murine organoid biobank that recapitulates *in vitro* key properties of normal, hyperplastic and tumour tissue.
- Characterization of the genomes and transcriptomes of the serrated organoids cohort for the identification of mutations that cooperate with other molecular events in a defined disease stage.
- Identify the mechanisms involved in the *in vivo* progression of early to advanced stages of CRC.
    - Establishment of an orthotopic injection mouse model to recapitulate local and distant disease *in vivo*.
    - Identify the genetic events that punctuate different stages of disease progression.

The study of human CRC is often limited to samples obtained at specific times according to the clinical situation of the patients (biopsy, surgery, etc). Although these specimens are of extreme value and have given us crucial information about the genetic

events responsible for malignant transformation, there is a general lack of appropriate models to study these lesions *in vivo* due to obvious constraints. As such, we additionally aim to develop a suitable mouse model for the study of human CRC in an *in vivo* setting. Specifically, we aim to:

- Generation of a human organoid biobank from normal tissue, adenoma and carcinoma-derived tissue samples.
- Establishment of an orthotopic injection mouse model using human CRC-derived organoids.
- Assess the feasibility of this model in terms of genetic and transcriptomic stability of mouse lesions in comparison to the human counterpart.

3 Materials and methods

3.1 Materials

3.1.1 Reagents and consumables

Material	Company
1000bp DNA Ladder	New England Biolabs
100bp DNA Ladder	New England Biolabs
2-Mercaptoethanol, 98%	Sigma-Aldrich
2-Propanol (isopropanol)	Carl Roth
A83-01	Sigma-Aldrich
Adhesive PCR Plate Foils	Thermo Fisher Scientific
Advanced Dulbecco's Modified Eagle Medium (DMEM), HG	Thermo Fisher Scientific
Agarose	Sigma-Aldrich
Ampicillin	Sigma-Aldrich
Antibiotic-Antimycotic	Thermo Fisher Scientific
Antibody p-ERK	Abcam
Antibody Ki67	Abcam
B27	Gibco
Biopsy/tissue embedding cassettes	Simport Scientific
Cell culture flasks (50 mL, 250 mL, 550 mL)	Greiner Bio-One
Cell culture plates (6-well, 12-well, 24-well, 96-well)	Greiner Bio-One
Cell recovery solution	Corning
Collagenase Type IV	Merck
Conical tubes (15 mL, 50 mL)	Greiner Bio-One
Cover slips	Gerhard Menzel B.V.
Cryotubes (1.6 mL)	Sarstedt
Deoxynucleotide (dNTP) Mix (10mM each)	Fermentas
Dimethyl sulfoxide (DMSO)	Carl Roth
DirectPCR Lysis Reagent (Mouse Tail)	Viagen Biotech
DNA LoBind Tubes (1.5 mL)	Eppendorf
EGF recombinant	Peprotech
Eosine	Waldeck
Ethanol absolute	Carl Roth
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
ExonucleaseI	New England Biolabs
Fetal Calf Serum Superior	Biochrom
Forene® isoflurane	Abbott
Formalin	Carl Roth
Gel Loading Dye, Purple (6x)	New England Biolabs
Glass slides SuperFrostTM Plus	Thermo Fisher Scientific
GlutaMAX	Gibco
Glycerol	Sigma-Aldrich
Hard-Shell® 96-Well PCR Plates, high profile, semi skirted	Bio-Rad Laboratories
Hard-Shell® Low-Profile Thin-Wall 96-Well Skirted PCR Plate	Bio-Rad Laboratories
Hematoxylin	Merck
HEPES	Gibco
KAPA Mouse Genotyping Kit	Sigma-Aldrich
LB-Agar (Luria/Miller)	Carl Roth
LB-Medium (Luria/Miller)	Carl Roth
Lipofectamine® 2000	Thermo Fisher Scientific
Matrigel basement membrane phenol red-free	Corning
MicroAmp® optical 96-well reaction plate	Thermo Fisher Scientific
MicroAmp® Optical Adhesive film	Thermo Fisher Scientific
N-acetylcysteine	Sigma-Aldrich
Nicotinamide	Sigma-Aldrich
Noggin recombinant	Peprotech
Normocin	Invivogen
Oral Gavage needle 20G x 25mm	VWR
PCR stripes (8 tubes)	Sarstedt
Penicillin-Streptomycin (5,000 U/ml)	Thermo Fisher Scientific
peqGREEN	VWR International
Petri dishes (100 mm)	Greiner Bio-One
Phosphate buffered saline	Sigma-Aldrich
Pipette tips with filter (10 µL, 100 µL, 200 µL, 300 µL, 1250 µL)	Biozym Scientific
Polybrene	Sigma-Aldrich
Potassium chloride (KCl)	Carl Roth
Proteinase K	Sigma-Aldrich

Puromycin	Sigma-Aldrich
Q5® High-Fidelity DNA Polymerase	Bio-Rad Laboratories
Reaction tubes safe-seal (0.5 mL, 1.5 mL, 2 mL) Reservoirs	Sarstedt
RNAlater	Sigma-Aldrich
Roti®-Histofix 4 %	Carl Roth
SB202190	Sigma-Aldrich
Scalpels	B. Braun Melsungen
Serological pipettes (5 mL, 10 mL, 25 mL, 50 mL)	Greiner Bio-One
Sodium hydroxide solution (NaOH)	Merck
Sucrose	Sigma-Aldrich
Syringes (1 mL, 30 mL)	B. Braun Melsungen
TaqManTM Fast Advanced Master Mix	Thermo Fisher Scientific
TransIT®-LT1 Transfection Reagent	Mirus Bio LCC
TRIS PUFFERAN	Carl Roth
Trypsin-EDTA (0.5%)	Thermo Fisher Scientific
Tween® 20	Carl Roth
Xylene	Carl Roth
Y-27632	Sigma-Aldrich

3.1.2 Mouse strains

Mouse strain	Source
Tg(Vil1-cre)1000Gum	In-house mouse facility of translaTUM
Apc^{tm1tyj}	In-house mouse facility of translaTUM
$Kras^{tm4tyj}$	In-house mouse facility of translaTUM
$Braf^{tm1.1Brd}$	In-house mouse facility of translaTUM
$Gt(ROSA)26Sor^{tm1(Pik3ca*H1047R)Egan}$	In-house mouse facility of translaTUM
NSG ($Prkdc^{scid}$ $Il2rg^{tm1Wjl}$)	In-house mouse facility of translaTUM

3.1.3 Library preparation and sequencing

Material	Company
Agencourt AMPure XP magnetic beads	Beckman Coulter
Agilent High Sensitivity DNA Kit	Agilent Technologies
EB buffer	Qiagen
Ion 520 Chip kit	Thermo Fisher Scientific
Ion Ampliseq library kit 2.0	Thermo Fisher Scientific
Ion library TaqMan quantitation kit	Thermo Fisher Scientific
KAPA DNA Standards and Primers for Illumina	Kapa Biosystems
KAPA HiFi HotStart ReadyMix (2x)	Kapa Biosystems
KAPA Library Quantification Kit	Kapa Biosystems
KAPA SYBR Fast qPCR ABI Mix (2x)	Kapa Biosystems
MiSeq Reagent Nano Kit v2 (300-cycles)	Illumina
NEBNext® Ultra DNA Library Prep Kit for Illumina®	New England Biolabs
Nextera XT Kit	Illumina
Sodium hydroxide (NaOH)	Carl Roth
TG NextSeq® 500/550 High Output Kit v2 (300 cycles)	Illumina

3.1.4 Primers

Primer	Sequence
AmpliconSeq_Apc-gDNA_F	5'-AAGACCAGGAAGCCTTGTGG-3'
AmpliconSeq_Apc-gDNA_R	5'-GCTTGTGTCTCTGCTTACTCC-3'
AmpliconSeq_mKras-gDNA_F	5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG AAGGCCTGCTGAAAATGACTGA-3'
AmpliconSeq_mKras-gDNA_R	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACA GACACCCAGTTTAAAGCCTTGGA-3'
AmpliconSeq_mBraf-gDNA_F	5'-TCGTCGGCAGCGTCAGATGTGTATAAGAG ACAGCCTGTGAGTAGTGGGAACTGT-3'
AmpliconSeq_mBraf-gDNA_R	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAG ACAGCCTTTACTTACTGCACCTCAGA-3'
Genotyping_Apctm1Tyj_1	5'-GATCACTCATCCGATAAGTGC-3'
Genotyping_Apctm1Tyj_2	5'-TTGGTTAAGGTGGTCTTGCAG-3'
Genotyping_Braftm1.1Brd_1	5'-TTTATCATAGTAGGGCTTGCTGTCTTGCTT-3'
Genotyping_Braftm1.1Brd_2	5'-CAAATATGTTTTGAGCAAGACCTTTGTTCT-3'
Genotyping_Braftm1.1Brd_3	5'-CCACTGACCAGAAGGAAAGTGGT-3'
Genotyping_Gt(ROSA)26Sortm1(Pik3ca*H1047R)Egan_1	5'-TGAATAGTTAATTGGAGCGGCCGCAATA-3'

Genotyping_Gt(ROSA)26Sortm1(Pik3ca*H1047R)Egan_2	5'-AAATAGCCGCAGGTCACAAAGTCTCCG-3'
Genotyping_Krastm4Tyj_1	5'-CACCAGCTTCGGCTTCCTATT-3'
Genotyping_Krastm4Tyj _2	5'-AGCTAATGGCTCTCAAAGGAATGTA-3'
Genotyping_Krastm4Tyj _3	5'-GCGAAGAGTTTGTCCTCAACC-3'
Genotyping_Tg(Vil1-cre)1000Gum_1	5'-CAAATGTTGCTTGTCTGGTG-3'
Genotyping_Tg(Vil1-cre)1000Gum_2	5'-GTCAGTCGAGTGCACAGTTT-3'
Genotyping_Tg(Vil1-cre)1000Gum_3	5'-GTGTGGGACAGAGAACAAACC-3'
Genotyping_Tg(Vil1-cre)1000Gum_4	5'-ACATCTTCAGGTTCTGCGGG-3'
Genotyping_Il2rgtm1Wjl_1	5'-GTGGGTAGCCAGCTGCTCTTCAG-3'
Genotyping_Il2rgtm1Wjl_2	5'-CCTGGAGCTGGACAACAAAT-3'
Genotyping_Il2rgtm1Wjl_3	5'-GCCAGAGGCCACTTGTGTAG-3'
IlluminaLibraryQuantification_F	5'-ATGATACGGCGACCACCGAG-3'
IlluminaLibraryQuantification_R	5'-CAAGCAGAAGACGGCATACGAG-3'
Mycoplasma PCR	5´-CACCTGAGTAGTATGCTCGC-3´
Mycoplasma PCR	5´-CGCCTGAGTAGTACGTACGC-3´
Mycoplasma PCR	5´-CGCCTGAGTAGTACGTTCGC-3´
Mycoplasma PCR	5´-CGCCTGAGTAGTATGCTCGC-3´
Mycoplasma PCR	5´-CGCCTGGGTAGTACATTCGC-3´
Mycoplasma PCR	5´-TGCCTGAGTAGTACATTCGC-3´
Mycoplasma PCR	5´-TGCCTGGGTAGTACATTCGC-3´
Nextera® adapter sequence_f	5´-TCGTCGGCAGCGTCAGATGTGTATAAGAGAC AG-3´
Nextera® adapter sequence_r	5´-GTCTCGTGGGCTCGGAGATGTGTATAAGAGA CAG-3´
NexteraIndexPrimer_NGS_i5_S502	5'AATGATACGGCGACCACCGAGATCTACACCTCTC TATTCGTCGGCAGCGTC-3
NexteraIndexPrimer_NGS_i5_S503	5'-AATGATACGGCGACCACCGAGATCTACACTATCC TCTTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S505	5'-AATGATACGGCGACCACCGAGATCTACACGTAA GGAGTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S506	5'-AATGATACGGCGACCACCGAGATCTACACACTG CATATCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S507	5'-AATGATACGGCGACCACCGAGATCTACACAAGG AGTATCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S508	5'-AATGATACGGCGACCACCGAGATCTACACCTAA GCCTTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S510	5'-AATGATACGGCGACCACCGAGATCTACACCGTC TAATTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S511	5'-AATGATACGGCGACCACCGAGATCTACACTCTCT CCGTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i7_N701	5'-CAAGCAGAAGACGGCATACGAGATTCGCCTTAG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N702	5'-CAAGCAGAAGACGGCATACGAGATCTAGTACGG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N703	5'-CAAGCAGAAGACGGCATACGAGATTTCTGCCTG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N704	5'-CAAGCAGAAGACGGCATACGAGATGCTCAGGAG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N705	5'-CAAGCAGAAGACGGCATACGAGATAGGAGTCCG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N706	5'-CAAGCAGAAGACGGCATACGAGATCATGCCTAG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N707	5'-CAAGCAGAAGACGGCATACGAGATGTAGAGAGG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N710	5'-CAAGCAGAAGACGGCATACGAGATCAGCCTCGG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N711	5'-CAAGCAGAAGACGGCATACGAGATTGCCTCTTG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N712	5'-CAAGCAGAAGACGGCATACGAGATTCCTCTACG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N714	5'-CAAGCAGAAGACGGCATACGAGATTCATGAGCG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N715	5'-CAAGCAGAAGACGGCATACGAGATCCTGAGATG TCTCGTGGGCTCGG-3'
RecombinationPCR_mBraf_1	5'-CTGATCCACCTGCAGTCAGCTAT-3'
RecombinationPCR_mBraf_2	5'-GGCCCAGGCTCTTTATGAGAATA-3'
3´pA RNA-Seq_PCR-1_SINGV6	5'-/5Biosg/ACACTCTTTCCCTACACGACGC-3'
3´pA RNA-Seq_PCR-2_i7.1-96	5'-CAAGCAGAAGACGGCATACGAGATNNNNNNNNG TCTCGTGGGCTCGG-3'
3´pA RNA-Seq_PCR-2_P5NEXTPT5	5'-AATGATACGGCGACCACCGAGATCTACACTCTTT CCCTACACGACGCTCTTCCG*A*T*C*T*-3'
3´pA RNA-Seq_RT-TemplateSwitch_E3V6NEXT	5'-/5Biosg/ACACTCTTTCCCTACACGACGCTCTTCCG ATCT[BC6]N10T30VN-3'
3´pA RNA-Seq_RT-TemplateSwitch_E5V6NEXT	5'-iCiGiCACACTCTTTCCCTACACGACGCrGrGrG-3'

3.1.5 Equipment

Equipment	Company
Agilent Bioanalyzer 2100 Axio Imager.A1	Agilent Technologies
Analytical balance A 120 S	Sartorius
AxioCam ICc5	Carl Zeiss
AxioCam MRc	Carl Zeiss
Camera control unit	Karl STORZ
Centrifuge 5424	Eppendorf
Centrifuge 5810 R	Eppendorf
Class II Biological Safety Cabinet	Thermo Fisher Scientific
CO2-incubator HeracellTM VIOS 250i	Thermo Fisher Scientific
Cold light fontain	Karl STORZ
DFC9000 GT camera	Leica Microsystems
DynaMagTM-96 Side Skirted Magnet	Thermo Fisher Scientific
Endoscope 1.9 mm, 10 cm	Karl STORZ
Electrophoresis power supply Power Pac 200	Bio-Rad Laboratories
Fiber optic light cable	Karl STORZ
Heated paraffin embedding module EG1150 H	Leica Microsystems
HiSeq 1500 System	Illumina
Homogenisator Precellys® 24	Bertin Instruments
Incubator NCU-Line® IL 23	VWR International
Ion Chef Instrument	Thermo Fisher Scientific
Ion S5 XL system	Thermo Fisher Scientific
M205FCA microscope	Leica Microsystems
Magnetic stirrer D-6010	neoLab Migge
Microwave	Thermo Fisher Scientific
MiSeq System	Illumina
Mouse heating pad	Conduct science
NanoDrop spectrophotometer	Thermo Fisher Scientific
Neubauer hemocytometer, improved	LO-Laboroptik
NextSeq 550 System	Illumina
Leica Bond RXm	Leica Microsystems
pH meter 521	WTW
Pipettes Reference®, Research®	Eppendorf
Primovert Microscope	Carl Zeiss
Qubit® 2.0 Fluorometer	Thermo Fisher Scientific
StepOne Plus Real-Time PCR System	Applied Biosystems
Thermocycler Tpersonal 48	Biometra
Thermocycler TProfessional Basic 96	Biometra
Thermocycler TProfessional Basic Gradient 96	Biometra
ThermoMixer® comfort 5355	Eppendorf
Tissue processor ASP300	Leica Microsystems
Ultra Low-Temperature Freezer Innova® U725	Eppendorf
UVsolo 2 Gel Documentation System	Analytik Jena
Vortex-Genie 2	Scientific Industries
Weighing Scale A120S	Sartorius

3.1.6 Kits

Kit	Company
DNeasy Blood & Tissue Kit	Qiagen
MinElute Reaction Cleanup Kit	Qiagen
Monarch Plasmid Miniprep kit	New England Biolabs
NucleoBond® Xtra Midi EF	Macherey-Nagel
QIAquick Gel Extraction Kit	Qiagen
BOND Polymer Refine Detection Kit	Leica Microsystems
QIAquick PCR Purification	Qiagen
QIAshredder	Qiagen
Qubit® dsDNA BR Assay Kit	Thermo Fisher Scientific
RNase-free DNase set	Qiagen
RNeasy Mini Kit	Qiagen

3.1.7 Plasmids and bacteria

Plasmid	Reference
plentiCRISPRv2	Addgene, #52961
pMD2.G	Addgene, #12259
psPAX	Addgene, #12260

Bacteria	Company
One Shot® Stbl3™ chemically competent E. coli	Thermo Fisher Scientific

3.1.8 Softwares

Software	Company
Agilent Genomic Workbench v7.0	Agilent Technologies
Aperio ImageScope v12.3.3	Leica Microsystems
AxioVision v4.8	Carl Zeiss
EndNote X20	Thomson Reuters
GraphPad Prism v8.2.1	GraphPad Software
Integrative Genomics Viewer (IGV) v2.3	Broad Institute
Illumina Sequence Analysis Viewer v2.4.7	Illumina
ImageJ (v1.50i)	Open source
Microsoft Office 365	Microsoft
LAS X software	Leica Microsystems
R v4.0.5	The R Project, The R Foundation
Snapgene 3.1	GSL Biotech
StepOneTM v2.3	Applied Biosystems
Torrent Suite software (v5.10.1)	Thermo Fisher Scientific

3.1.9 Databases

Databases	Source
Cancer Cell Line Encyclopedia	https://portals.broadinstitute.org/ccle
cBioPortal	http://www.cbioportal.org/
EnrichR	https://maayanlab.cloud/Enrichr/
Ensembl	http://www.ensembl.org/index.html
Exome Variant Server	http://evs.gs.washington.edu/EVS
Gene Expression Omnibus	https://www.ncbi.nlm.nih.gov/geo/
International Cancer Genome Consortium	https://icgc.org/
Molecular Signatures Database	https://cancer.sanger.ac.uk/signatures/downloads/
PubMed	https://www.ncbi.nlm.nih.gov/pubmed

3.1.10 Manufacturers

Company	Headquarters
Abbott	Ludwigshafen, Germany
Abcam	Cambridge, UK
Addgene	Cambridge, Massachusetts, USA
Agilent Technologies	Santa Clara, CA, USA
Analytik Jena	Jena, Germany
Applied Biosystems	Carlsbad, CA, USA
ATCC	Manassas, VA, USA
B. Braun Melsungen	Melsungen, Germany
Beckman Coulter	Munich, Germany
Bertin Instruments	Montigny-le-Bretonneux, France
Biochrom	Berlin, Germany
Biometra	Göttingen, Germany
Bio-Rad Laboratories	Hercules, CA, USA
Biozym Scientific	Hessisch Oldendorf, Germany
Brand	Wertheim, Germany
Carl Roth	Karlsruhe, Germany
Carl Zeiss	Oberkochen, Germany
Corning	Corning, NY, USA
Covaris	Woburn, MA, USA
Eppendorf	Hamburg, Germany
Eurofins Genomics	Ebersberg, Germany
Fermentas	St. Leon-Rot, Germany
Gerhard Menzel B.V.	Braunschweig, Germany
Gibco	Carlsbad, CA, USA
GraphPad Software	San Diego, CA, USA
Greiner Bio-One	Kremsmünster, Austria
GSL Biotech	Chicago, IL, USA
Illumina	San Diego, CA, USA
Integra Biosciences	Biebertal, Germany
Kapa Biosystems	Wilmington, MA, USA
Leica Microsystems	Wetzlar, Germany
LO-Laboroptik	Friedrichsdorf, Germany

Macherey-Nagel	Düren, Germany
Merck	Darmstadt, Germany
Microsoft	Redmond, Washington, USA
Mirus Bio LCC	Madison, WI, USA
neoLab Migge	Heidelberg, Germany
New England Biolabs	Ipswich, MA, USA
Qiagen	Hilden, Germany
Sarstedt	Nümbrecht, Germany
Sartorius	Göttingen, Germany
Scientific Industries	Bohemia, NY, USA
Seidel Medipool	Gauting-Buchendor, Germany
Sigma-Aldrich	St. Louis, MO, USA
Simport Scientific	Beloeil, QC, Canada
The R Project, The R Foundation	Vienna, Austria
Thermo Fisher Scientific	Waltham, MA, USA
Thomson Reuters	Carlsbad, CA, USA
Viagen Biotech	Los Angeles, CA, USA
VWR International	Darmstadt, Germany
Waldeck	Münster, Germany
Worthington Biochemical	Lakewood, NJ, USA
WTW	Weilheim, Germany
ZytoVisio	Bremerhaven, Germany

3.2 Methods

3.2.1 Animal experiments

Intestine-specific Cre driver lines (99) were crossed with *Apc*$^{fle1-15}$ (5), *Kras*^{G12D} (154), *Braf*^{V637E} (9) and *Pik3ca*^{H1047R} (107) mice for the generation of *Vil-Cre; Apc*$^{fle1-15}$, *Vil-Cre; Kras*^{G12D}, *Vil-Cre; Braf*^{V637E} and *Vil-Cre; Pik3ca*^{H1047R} cohorts. All animals with signs of sickness were sacrificed in compliance with the European guidelines for the care and use of laboratory animals. Prism (GraphPad Software v8.2.1) was used for the generation of Kaplan-Meier survival curves. All mice were maintained on C57Bl/6 background and housed under specific-pathogen-free conditions. Female and male mice were randomly submitted to tumour/mouse model cohorts and monitored weekly for signs of disease. For necropsy of tumour-bearing mice, the small intestine and colon was macroscopically checked for the presence of primary tumours and metastases at the main metastatic routes (liver, lung, lymph nodes). Animal studies were approved by the Institutional Animal Care and Use Committees (IACUC) of Technische Universität München (Regierung von Oberbayern, Munich, Germany).

3.2.2 Genotyping

Ear punches from 3 weeks old pups were used for genotyping. The tissue was lysed in DirectPCR Lysis Reagent (Mouse Tail) (Viagen Biotech) supplemented with 20 µg/ml proteinase K (Sigma-Aldrich) at 55°C overnight. Proteinase K was inactivated at 95°C for 15 min. DNA was then diluted 1:5 for genotyping and 1 µl of the diluted lysis solution was added to 4 µl of H2O (including mouse allele-specific genotyping primers) and 5 µl of 2x genotyping master mix. The 2x concentrated genotyping mix was prepared from peqGOLD

Taq-DNA-Polymerase kit (VWR), supplemented with dNTPs (Fermentas GmbH), sucrose (Sigma-Aldrich) and SucRot buffer, was used for genotyping. PCR program used was similar for all genotyping reactions. The annealing temperatures were specific for each allele for each PCR reaction:

Temperature	Time	Cycles
95 °C	180 sec	1x
95 °C	45 sec	40x
XX °C	60 sec	40x
72 °C	90 sec	40x
72 °C	300 sec	1x
10 °C	Pause	---

Allele	Annealing Temperature
Tg(Vil1-cre)1000Gum	62 °C
Il2rgtm1Wjl	60 °C
Braftm1.1Brd	55 °C
Krastm4Tyj	55 °C
Apctm1Tyj	63 °C
Gt(ROSA)26Sortm1(Pik3ca*H1047R)Egan	64 °C

3.2.3 Murine organoids isolation

Normal tissue

Mouse Intestinal crypts were isolated and cultured in Matrigel (Corning) as previously described (29). Briefly, upon mouse necropsy the intestine is rinsed with ice cold PBS and cut into ~5mm pieces. These are then incubated in a 5 mM EDTA (Sigma-Aldrich) solution in PBS for small intestine and 30 mM EDTA for colon organoids isolation for 10 minutes with gentle shaking in order to remove villi and other debris. These incubation steps are repeated twice to a total of 3 times. After the final incubation, tissue pieces are resuspended in ice cold PBS and shaken vigorously for 3 cycles of 30 seconds, to detach crypts from the intestine. Cell suspension is then filtered twice using first a 70 µm and then a 100 µm Cell Strainer (Greiner Bio-One) to select for crypt units only. Solution is then centrifuged for 300x g for 5 minutes after which crypts are washed in ice cold PBS and counted using a Neubauer chamber (LO-Laboroptik). Around 50-100 crypts were resuspended in 50 µl of Matrigel and plated in a well of a 24-well plate (Corning) supplemented with 500 µl of 50% L-WRN medium. Medium was changed every 2-3 days and the culture passaged every 5-6 days.

Important: All centrifugation steps were performed at 4°C in all protocols involving organoids isolation and maintenance for all tissues from both species (mouse and human).

Tumour tissue

Mouse tumours were minced thoroughly with a scalpel and incubated with 1% PenStrep (Thermo Fisher Scientific) in PBS for 15 minutes on ice to remove possible contaminants from the intestine. After a washing step with PBS and centrifugation for 300x g for 5 minutes the tissue is digested with 230U of collagenase type IV (Merck) at 37°C for 30 minutes. Tissue suspension is then filtered using a 100 µm strainer allowing only small tissue pieces and cell clumps to go through and washed once with PBS. Inactivation of collagenase is then achieved by incubation with a solution containing 10% FCS for 10 minutes. are then washed once with PBS and counted using a Neubauer chamber. After centrifugation at 300x *g* for 5 minutes ~100 cell clumps are resuspended in 50 µl of Matrigel and plated in a well of a 24-well plate (Corning) supplemented with 500 µl of 50% L-WRN medium supplemented with 10 µM of Y27632 (Sigma-Aldrich). Medium was changed every 2-3 days and the culture passaged every 4-5 days.

3.2.4 Human organoids isolation

Normal tissue

After tissue resection, the mucosa piece was washed 3 times with cold PBS and excess muscle layer and sub-mucosa carefully removed by using forceps and a scalpel. The tissue was cut into smaller pieces (~0.5 cm) and incubated with a mixture of Normocin 1:250, (Invivogen) and Antibiotic-Antimycotic 1:100 (Thermo Fisher Scientific) for 15 minutes at room temperature. Then tissue was washed with PBS and incubated twice in 10 mM DTT (Sigma-Aldrich) in 30 ml PBS for 5 minutes at room temperature. After the supernatant was discarded, samples were transferred to 30 ml 8 mM EDTA in PBS on ice and slowly rotated for 60 minutes at 4ºC. Supernatant was replaced with fresh, ice-cold PBS, and shook vigorously to yield a solution enriched in colonic crypts. After centrifugation at 40x *g* for 5 minutes, crypts are washed in PBS, counted and centrifuged at 80x *g* for 5 minutes. A total of 100-150 crypts are resuspended in Matrigel as previously described. Normal tissue organoid cultures are then supplemented with 50% L-WRN, 10mM HEPES (Gibco), 2mM Gutamax (Gibco), 1x B27 (Gibco), 1.25 mM N-Acetylcysteine (Sigma-Aldrich), 50ng/ml EGF (Peprotech), 500 nM A83-01 (Tocris Biosciences), 7.5 uM SB202190 (Sigma-Aldrich) and 10 µM of Y27632 (Sigma-Aldrich). Medium was changed every 2-3 days and the culture passaged every 4 days.

Tumour tissue

Resected tumours were incubated in a solution containing Normocin 1:250, and Antibiotic-Antimycotic 1:100 for 15 minutes at room temperature and then minced with scalpels into smaller pieces. After 2 washing steps with PBS the tumour was digested in a solution containing 5 ml of Dispase II solution (ready to use, Millipore) add 3.33U Collagenase IV and 2U of DNaseI solution (Sigma-Aldrich) and incubated at 37 ºC for 60 minutes. Tissue suspension is then filtered using a 70 µm strainer and washed twice with PBS. After a centrifugation step at 300x *g* for 5 minutes cells are resuspended in PBS and counted as described before. A total of 100-150 cell clumps are resuspended in Matrigel and plated in a well of a 24-well plate supplemented with a medium containing advanced DMEM 10mM HEPES, 2mM Gutamax, 1x B27, 1.25 mM N-Acetylcysteine, 50ng/ml EGF, 500 nM A83-01, 7.5 uM SB202190 and 10 µM of Y27632, 100 ng/mL human Noggin (Peprotech), 10mM Nicotinamide (Sigma-Aldrich), Normocin diluted 1:500, and Antibiotic-Antimycotic diluted 1:100.

3.2.5 Passaging of organoid cultures

Passaging of organoid cultures was done by incubating the culture with Cell Recovery Solution (Corning) for 30 minutes on ice to dissolve the Matrigel and retrieve the cellular part of the solution. After a centrifugation step of 300x *g* for 5 minutes the murine cell pellet was washed once with PBS, and following its removal by centrifugation, cells were chemically digested by incubation with TypLE Express Enzyme (Thermo Fisher Scientific) diluted 1:2 for normal tissue and 1:1 for tumour tissue organoids in PBS and incubated at 37°C for 3 minutes. Cells were then washed once with PBS, centrifuged at 300x *g*, resuspended in Matrigel and plated in 24-well plate (50 µl/well). For passaging of human organoids (normal and tumour tissue) only mechanical dissociation is used and no incubation with TrypLE Express Enzyme is performed. All organoid lines were routinely imaged using a M205FCA microscope equipped with a DFC9000 GT camera (Leica Microsystems).

3.2.6 L-WRN medium production

As culture medium for mouse and human normal human organoids, a conditioned medium containing Wnt3A R-spondin1, and Noggin was prepared using L-WRN cells (CRL-3276, ATCC) according to the manufacturer's protocol. Briefly L-WRN cells were expanded and cultured until a state of overconfluency supplemented with Advanced DMEM/F12 (Thermo Fisher Scientific), 1% PenStrep, 10% fetal calf serum (FCS Biochrom), 1% HEPES and 1% Glutamax and grown overnight. Next day, the supernatant was collected and

centrifuged for 2000x *g* for 5 minutes to remove dead cells and debris, and fresh medium added to the cells. The harvesting step was repeated 3 more times for a total of 4 harvests. On the fourth day the collected medium was combined and diluted 1:1 in the same medium as used for culturing the cells (hence 50% L-WRN medium).

3.2.7 Mycoplasma PCR

For testing mycoplasma contamination, organoids were grown for 3 days in medium lacking PenStrep. From these cultures, 1 ml of cell culture supernatant was harvested in a 1.5 ml Eppendorf tube and stored at -20°C. 5.5 µl of 2x KAPA Genotyping Mix (Sigma-Aldrich), 0.22 µl of primer mix (containing 5 µM of each primer; 7x forward plus 3x reverse primer) and 4.28 µl ddH_2O were prepared for each PCR reaction resulting in a total volume of 10 µl MasterMix per sample. Harvested cell culture supernatant was defrosted on ice and 1 µl of added to the MasterMix. PCR reaction was performed with the following conditions:

Temperature	Time	Cycles
98 °C	180 sec	1x
65 °C	60 sec	1x
72 °C	60 sec	1x
98 °C	15 sec	35x
65 °C	30 sec	35x
72 °C	20 sec	35x
72 °C	300 sec	1x
10 °C	Pause	---

3.2.8 Isolation of genomic DNA and RNA from organoid cultures

Genomic DNA (gDNA) was isolated using the DNeasy Blood & Tissue Kit (Qiagen) from murine and human primary organoid cultures. Briefly, organoids were grown for 3-4 days in Matrigel in a 24-well plate. On the day of cell harvest, cultures were incubated with Cell Recovery Solution for Matrigel digestion for 30 minutes on ice. Then, the content of at least 3 full wells of organoids was pooled and washed once with ice cold PBS. After a centrifugation step at 300x g for 5 minutes, cell pellet was wither snapfrozen in liquid nitrogen for downstream applications or used for DNA isolation according to manufacturer's instructions of the DNeasy Blood & Tissue Kit (Qiagen).

For RNA isolation, organoids were grown under the same conditions as the mentioned above. After Matrigel removal cell pellets were either snapfrozen in liquid nitrogen or resuspended in 300 µl of RLT buffer (Qiagen) supplemented with 1:100 ß-mercaptoethanol (Sigma-Aldrich) and the RNA isolated with the RNeasy Mini kit (Qiagen) according to manufacturer's instructions.

3.2.9 Histological characterization of mouse and human tumour lesions

Formalin-fixed paraffin embedded (FFPE) tissue blocks were cut in 2 µm thick sections and H&E stained at the pathological core facility of the clinic. Analysis of both murine and human tumour histology was performed by an experienced pathologist in comparative gastrointestinal pathology at the Institute of Pathology, TU Munich. Histopathologic grading of primary tumours was performed according to the most recent consensus system for classification of gastrointestinal tumours by the World Health Organization (WHO) (155). The presence of metastatic disease was additionally assessed in mouse liver, lung, pancreas and spleen. Immunohistochemistry (IHC) of tumour tissue and organoids was also performed at the pathological core facility of the clinic. Briefly, tissue sections were pre-treated using heat mediated antigen retrieval with sodium citrate buffer (pH6) for 20 mins. Sections were then incubated with ab15580 (anti-Ki67, 1:500) or ab214362 (anti-ERK, 1:200) for 15 min at room temperature. Antibody detection was performed on the Leica Bond RXm (Leica Microsystems) platform using the BOND Polymer Refine Detection Kit (Leica Microsystems).

3.2.10 Amplicon-based deep sequencing at the *Kras* and *Braf* locus

For amplicon-based deep sequencing of the *Kras* and *Braf* locus, 50 ng of high-quality genomic DNA (gDNA) was used. Q5® High- Fidelity DNA Polymerase (New England Biolabs) and primers with Nextera adapter overhangs were used for amplification of the *Kras* and *Braf* locus using the following protocol:

Temperature	Time	Cycles
98 °C	30 sec	1x
98 °C	10 sec	40x
60 °C	20 sec	40x
72 °C	15 sec	40x
72 °C	120 sec	1x
10 °C	Pause	---

PCR products were then purified by solid phase reversible immobilization. For this, 0.8x volume (20 µl) of Agencourt® AMPure® XP beads (Beckman Coulter GmbH) were added to the total PCR reaction (25 µl) and the cleanup performed according to manufacturer's instructions. After cleanup of the first PCR reaction, Nextera index primers (Illumina) were added to the PCR amplicon in a second Q5® PCR step (15 cycles) for barcoding of up to 96 samples using the following PCR protocol:

Temperature	Time	Cycles
98 °C	30 sec	1x
98 °C	10 sec	15x
65 °C	30 sec	15x
72 °C	60 sec	15x
72 °C	120 sec	1x
10 °C	Pause	---

Cleanup of PCR products was performed using the Agencourt® AMPure® XP kit (Beckman Coulter GmbH) once more, and these were applied on a 1.5% agarose gel. The relative quantity of each PCR product was estimated from the agarose gel by using ImageJ (v1.50i). Equal sample amounts were pooled according to the quantification after normalization of the results. The pooled library was then quantified by KAPA SYBR® Fast qPCR ABI Mix and the KAPA Library Quantification Kit (Kapa Biosystems) by using the following qPCR protocol:

Temperature	Time	Cycles
95 °C	300 sec	1x
95 °C	15 sec	35x
60°C	45 sec	35x

The concentration of the pooled library was adjusted to 4 nM and 20-50% of PhiX DNA was spiked into the pooled library depending on the heterogeneity of amplicons present in the final library (e.g. spike in of 50% PhiX if only one type of PCR amplicon is present in the pooled library). After denaturation, the spiked library was further diluted for sequencing according to manufacturer's instructions. Sequencing of the final library was performed in 300bp paired end mode on a MiSeq system (Illumina). Sequencing raw reads were analysed by Thomas Engleitner, a bioinformatician in our group and were mapped to the mouse reference genome (Ensemble release GRCm38p6, Genome Reference Consortium). For the calculation of *Kras*^{G12D} over *Kras*^{WT} ratios, variant allele calling was performed at the *Kras* locus on chr6 at position 145246771. For the calculation of *Braf*^{V637E} over *Braf*^{WT} ratios, variant allele calling was performed at the *Braf* locus on chr6 at position 39627783.

3.2.11 Whole-exome sequencing (WES) analysis of CRC organoids

Library preparation and sequencing of samples was performed using SureSelectXT Mouse All Exon kit from Agilent according to manufacturer's instructions and conducted by the "High Throughput Sequencing" unit of the Genomics & Proteomics Core Facility, German Cancer Research Center (DKFZ). Bionformatic analysis of output data was

conducted in the lab by Niklas Krätzig. Before mapping, raw sequencing reads were trimmed using Trimmomatic v0.39. Leading and trailing bases with Phred scores below 25 and reads shorter than 50 nucleotides were removed. In addition, the average base quality within a sliding window of 10 nucleotides should be above 25 to keep the read for further downstream analysis. Reads were aligned to the GRCm38.p6 reference genome using BWA-MEM 0.7.17 with default settings. PCR duplicates were marked with sambamba version 0.7.0 together with Picard tools version 2.20.0 and realignment around indels was performed with GATK toolkit version 4.1.3.0. Mutect2 was used for calling somatic mutations with default settings. Potential somatic events were filtered for SNPs by excluding single nucleotide variants (SNVs) which were listed in version 5 of the Mouse Genome Project SNP database. Somatic point mutations were included in the final list, if the read coverage for each position was at least 5 in both control and tumour, variant frequency was at least 10% and read count supporting the variant nucleotide was at least 2 in the tumour sample and equal to 1 or 0 in the control. Further, SNVs marked as strand or PCR bias artefacts by 'LearnReadOrientationModel' were excluded from further analysis. Annotation of somatic events was conducted with SNPeff v4.3. SNVs causing variation in splice sites or upstream/downstream of genes and mutations with low predicted impact in general were excluded from further analysis.

3.2.12 Analysis of mutational signatures

Mutational signatures were analysed with the help of Niklas Krätzig from our group. A list of 67 single base substitutions (SBS) mutational signatures previously identified by Alexandrov et al., (10) was used as the reference dataset of known mutational signatures. For the extraction of COSMIC signatures, the command "extractSignatures" (uses non-negative matrix factorization to decompose the matrix into n signatures) and "compareSignatures" (extracted signatures from previous step can be compared to known signatures from the COSMIC database) from the Maftools package were used and cosine similarity is calculated to identify best match(156).

3.2.13 Amplicon sequencing analysis of *Apc*$^{\Delta Ex16}$ organoids

Amplification of the *Apc* locus containing indels as result of CRISPR/Cas9 editing was done by PCR using specific primers containing the sgRNA target sites. Nextera adapter sequences were added to the 5' tail of each primer using the following conditions:

Temperature	Time	Cycles
98 °C	90 sec	1x
98 °C	10 sec	37x
66 °C	20 sec	37x
72 °C	20 sec	37x
72 °C	300 sec	1x
10 °C	Pause	---

After a cleanup step with AMPure XP magnetic beads, amplicons were quantified using the Quant-iTTM PicoGreenTM dsDNA reagent according to manufacturer´s instructions. Ten nanograms were used subsequently for the indexing step with index i7 and i5 primers from the Nextera XT Index Kit v2 (Illumina):

Temperature	Time	Cycles
98 °C	30 sec	1x
98 °C	10 sec	6x
65 °C	30 sec	6x
72 °C	60 sec	6x
72 °C	120 sec	1x
10 °C	Pause	---

Following another cleanup step with AMPure XP magnetic beads, the DNA library was quantified using the Agilent 2100 Bioanalyzer and the respective Agilent DNA 7500 reagents according to manufacturer's instructions. Library was denatured and 10 pM loaded with >10% PhiX onto the MiSeq Sequencer according to the MiSeq Reagent Kit V2 Nano (500 cycles PE) manual. Sequence was performed in 2 x 250 bp paired-end mode to achieve ~1000-10000x coverage according to the MiSeq system guide.Sequencing reads were aligned to the mouse genome and analyzed using the software CRISPResso2 (157).

3.2.14 lcWGS analysis

Purified DNA was used as input for library preparation with NEBNext Ultra II FS DNA Library Prep Kit for Illumina and processed according to manufacturer´s instructions. Libraries were quality controlled and quantified by Qubit measurement and Agilent DNA Bioanalyzer analysis. The library was sequenced on a NextSeq 500 (Illumina) with 67 cycles for DNA in read1. CNVs were analyzed by Niklas Krätzig from our group according to the MoCaSeq pipeline (158).

3.2.15 RNA-Sequencing analysis

Library preparation for bulk-sequencing of poly(A)-RNA was done as described previously (159). Briefly, barcoded cDNA of each sample was generated with a Maxima RT

polymerase (Thermo Fisher) using oligo-dT primer containing barcodes, unique molecular identifiers (UMIs) and an adaptor. Ends of the cDNAs were extended by a template switch oligo (TSO) and full-length cDNA was amplified with primers binding to the TSO-site and the adaptor. NEB UltraII FS kit was used to fragment the cDNA. After end repair and A-tailing a TruSeq adapter was ligated and 3'-end-fragments were amplified using primers with Illumina P5 and P7 overhangs. In comparison to (159), the P5 and P7 sites were exchanged to allow sequencing of the cDNA in read1 and barcodes and UMIs in read2 to achieve a better cluster recognition. The library was sequenced on a NextSeq 500 (Illumina) with 63 cycles for the cDNA in read1 and 16 cycles for the barcodes and UMIs in read2.

For the analysis part, Gencode gene annotations v35 and the human reference genome GRCh38 were derived from the Gencode homepage (EMBL-EBI). Drop-Seq tools v1.12 (160) was used for mapping raw sequencing data to the respective reference genome. The resulting UMI filtered count matrix was imported into R v4.0.5 and low expressed genes were subsequently filtered out. Prior differential expression with DESeq2 (161) dispersion of the data with a parametric fit using the phenotype as explanatory variable to model the data. The Wald test was used for testing differential expression between genotypes. Shrunken log2 foldchanges were calculated with the apeglm method. If not otherwise stated a gene was determined to be significantly regulated if the p-value was below 0.01 and the absolute, apeglm corrected, log2 foldchange was above 0.5. GSEA (3) was conducted with genesets provided in MsigDB v7.4 (162) in the PreRanked mode with the apeglm corrected log2 foldchanges. A pathway was to be considered to be significantly associated with a phenotype if the FDR was below 0.05. For visualization purposes the data was rlog (as implemented in the rlog function from DESeq2) transformed. Heatmaps show z-scaled rlog transformed data.

3.2.16 Lentiviral transduction of murine organoids for CRISPR/Cas9 genetic *knockout*

The plentiCRISPRv2 vector system (163) was used for the genetic *knockout* of Apc in wt mouse organoids. We validated a previously used single-guide RNA (sgRNA) in another study (target sequence GTCTGCCATCCCTTCACGTTAGG, PAM sequence underlined) (164). The sgRNA was cloned into the plentiCRISPRv2 vector by Golden Gate cloning and cloned vectors were verified using Sanger sequencing. Testing of the CRISPR/Cas9 system was previously tested before proceeding to lentivirus production. For this, we transfected a primary pancreatic cancer cell (P3109) commonly used in our group for guides validation by using Lipofectamine® 2000 according to manufacturer's instructions. Next day, cells were selected with 2 µg/ml Puromycin for 48 hours.

Confirmation of the *knockout* in 2D cells was done by Sanger Sequencing. For lentivirus production HEK293FT cells were transfected with psPAX and pMD2.G virus packaging plasmids and the plentiCRISPRv2::sg*Apc* vector by using TransIT®-LT1. Supernatant-containing viruses of HEK293FT cells was harvested 48h and 72h post transfection, and pooled. Lentivirus was used immediately for cell transduction or frozen in 1 mL aliquots at -80°C. For the organoids transduction we isolated intestinal organoids of wt mice as previously described. The content of 3 dense organoid wells (~300-500 organoids) were transduced by spinoculation. Briefly, organoids were separated from the Matrigel by incubation with Cell Recovery Solution, chemically digested with 1:2 TrypLE/PBS solution for 3 minutes at 37°C. Then organoids pellets were resuspended in lentivurs in the presence of 8 µg/ml polybrene and centrifuged for at 600x *g* for 1 hour at 32°C in a 24-well plate. After spinoculation the plate was incubated for 6 hours at 37°C. Then organoids were centrifuged, resuspended in Matrigel and plated in a 24-well plate (50 µl/well) and supplemented with 50% L-WRN containing 10 µM of Y27632. Next day selection was started by replacing the culture medium with Wnt-depleted medium (-Wnt3a, Rspondin). *Apc*-deficient organoids were cultured under these conditions from here on while wt organoids did not survive.

3.2.17 Human mutations analysis

For the identification of mutations in both human FFPE tissue and organoid lines we performed targeted amplicon sequencing based on a multiplex PCR-based Ion Torrent AmpliSeq technology (Life Technologies) approach using a customized panel of recurrently mutated genes in CRC (CRCv2 panel) containing 379 amplicons spanning the following genes and exons:

Gene	**Exons**
AKT1	(3,4,5,11)
ERBB2	(11,12,13,17,19-21)
MSH6	(3,4,5,8,9)
PMS2	(8,9,10,11)
SMAD4	(2-6,8-12)
APC	(2-16)
KRAS	(2,3,4)
MYC	(2,3)
POLE	(9-14,26,33,34,36)
TGFBR2	(3-7)
BRAF	(8,11,14,15,16)
MED12	(2)
NRAS	(2,3,4)

PTEN	(2-10)
RB1	(2,3,6,13,16,17,18,20,21,22, 23)
CTNNB1	(3,5,6,7,8,15)
MLH1	(2,3,4,8,10,12,13,15)
PIK3CA	(2,3,5,7-10,12,14,19-21)
RB1	(2-27)
RET	(10,11,13,15,16)
EGFR	(2,3,8,18,19,20,21)
MSH2	(1,7,11,13,15)
PIK3R1	(2,9,10,11,13,14)
SMAD2	(4,5,8,11)
TP53	(2,4-10)

Two pools of primers were generated for the library preparation step. Five ng of genomic DNA was used per each pool and mixed with the AmpliSeq HiFi Master Mix (Life Technologies) according to manufacturer's recommendations. Final library was purified using Agencourt® AMPure® XP beads and quantified using qPCR (Ion Library TaqMan Quantitation Kit, Life Technologies) on a StepOnePlus qPCR machine (Life Technologies). Individual libraries were diluted to a final concentration of 25pM, and 6 libraries were pooled, processed for sequencing, and automatically loaded on a 520 Chip using the Ion Chef system (Thermo Fisher). Sequencing was performed on an Ion S5XL instrument (Thermo Fisher scientific). Processing of raw sequencing reads and alignment against the human genome (version hg19) was performed with the Torrent Suite Software v5.10.1 (Thermo Fisher Scientific) using the TMAP algorithm. Mutation analysis was obtained using the build-in plugin "variantCaller" v5.8.0.19. Annotation of the variants was performed using a custom-built variant annotation pipeline using ANNOVAR (165). Visualization of sequencing reads was done using the Integrative Genomics Viewer Browser (IGV, http://www.broadinstitute.org/igv/) and variants were checked for germline or somatic origin using the COSMIC (catalogue of somatic mutations in cancer) database, dbSNP, and Exome Variant Server (http://evs.gs.washington.edu/EVS). Coverage data summary for each sample and amplicons generated by the Torrent Suite software was used for the identification of copy number variations (CNVs, amplifications and deletions) using a four-step algorithm as previously described (165, 166).

3.2.18 Orthotopic organoid transplantation

The orthotopic transplantations of organoids were performed always in collaboration with Dr. Markus Tschurtschenthaler, a PostDoc from the group of Professor Dieter Saur. as

previously described (164). Briefly, organoids were mechanically dissociated into 5-10 cell clusters and resuspended in a minimal medium (Advanced DMEM/F12 containing 1x B27, 1x N2, L-Glutamine (all from Gibco, Thermo Fisher Scientific), 10% Matrigel (Corning), and 10µM Y-27632 (STEMCELL Technologies)). For every injection (2-3 per mouse) 50-70 dissociated organoids in a volume of 80µl were prepared. Subsequently, the colon of the anesthetized mice was gently rinsed with PBS using a syringe and a straight oral gavage needle. Colonoscopy of mice was performed using a rigid endoscope from Karl STORZ (1.9 mm in diameter) with linear Hopkins lens optics (ColoView System). For injections of organoids into the submucosa of the colon a flexible fine needle (Hamilton; 33 gauge, custom length of 16 inches, custom point style of 4 at 45°) was used. Injections that were correctly applied into the submucosa led to the formation of a bubble that closes the intestinal lumen. A score was used to correlate the quality of injections with the outcome:

Score	Description
A	Injection filled the lumen of the intestine
a	Injection filled the lumen of the intestine but is quickly absorbed
B	Injection filled half the lumen of the intestine
b	Injection filled half the lumen of the intestine but is quickly absorbed
C	Small injection filling less than half of the intestine
c	Small injection filling less than half of the intestine quickly abosrbed
d	Injection does not result in bubble formation

4 Results

4.1 GEMMs of intestinal cancer recapitulate key features of human CRC

CRC is a heterogenous disease that can develop through different pathways with underlying distinct initiating events (56). In order to model the behaviour of tumours arising from different cancer-initiating pathways we bred tumour mice expressing the loss of the TSG *Apc* (*Vil-Cre*; *Apc$^{fle1-15}$*) (5) or the oncogenic activation of *Kras* (*Vil-Cre;Kras^{G12D})* (8), *Braf* (*Vil-Cre; Braf^{V637E})* (9) or *Pi3k (Vil-Cre; Pik3ca^{H1047R})* (unpublished model) (**Figure 7A and B**). As previously stated, the loss of *Apc* in the intestine leads to CIN-positive tumours which develop via the classical adenoma-carcinoma sequence (167) whilst the activation of *Kras^{G12D}*, *Braf^{V637E}* and *Pik3ca^{H1047R}* leads to tumours arising from the serrated and mucinous route of intestinal cancer. *Vil-Cre*; *Apc$^{fle1-15}$* mice exhibited a median survival time of 40.3 weeks (**Figure 7C**) and developed tumours with a villotubular morphology mostly in the SI, and rarely in the colon. *Vil-Cre; Kras^{G12D}, Vil-Cre; Braf^{V637E}* and *Vil-Cre; Pik3ca^{H1047R}* mice displayed elongated and dilated SIs due to IEC hyperproliferation which subsequently progressed to hyperplasia and then to serrated lesions (adenoma and carcinoma) (**Figure 7A**). Of note, most of the tumours in this cohort were evaluated as TSAs which are extremely rare in humans (<1%) (82). Conversely, hyperplasia is attenuated in the colon of these mice which was devoid of tumours in 95-100% of animals. The number of tumours per mouse differed between the classical and serrated mouse models (**Figure 7D**). *Apc*-deficient mice developed significantly more tumours per mouse than in the serrated models ($\bar{x}$=12.8 tumours/mouse) and shorter survival (*Vil-Cre; Kras^{G12D}* 89 weeks, *Vil-Cre; Braf^{V637E}* 78 weeks, *Vil-Cre; Pik3ca^{H1047R}* 89 weeks). Moreover, in the latter, not only the number of tumours was similar between animals from distinct serrated models (*Vil-Cre; Kras^{G12D}* $\bar{x}$=1.4 tumours/mouse*, Vil-Cre; Braf^{V637E}* $\bar{x}$=2.1 tumours/mouse*, Vil-Cre; Pik3ca^{H1047R}* $\bar{x}$=2.1 tumours/mouse) but also, overall mice showed a long latency for tumour development, usually for more than a year (**Figure 7C**). Nevertheless, although these mouse models faithfully recapitulate human CRC subtypes, mice predominantly develop tumour lesions in the small intestine which is a rare location for sporadic tumours in humans (**Figure 7E**). In humans, SI tumours account for less than 5% of all gastrointestinal cancers and are usually associated with hereditary predisposition syndromes (168). Also, metastases formation is a very infrequent finding in these mouse models which hampers the study of tumour progression. In our mouse cohort only 3 mice developed distant metastases in the liver (2 *Vil-Cre; Kras^{G12D}* mice and 1 *Vil-Cre; Braf^{V637E}* mouse).

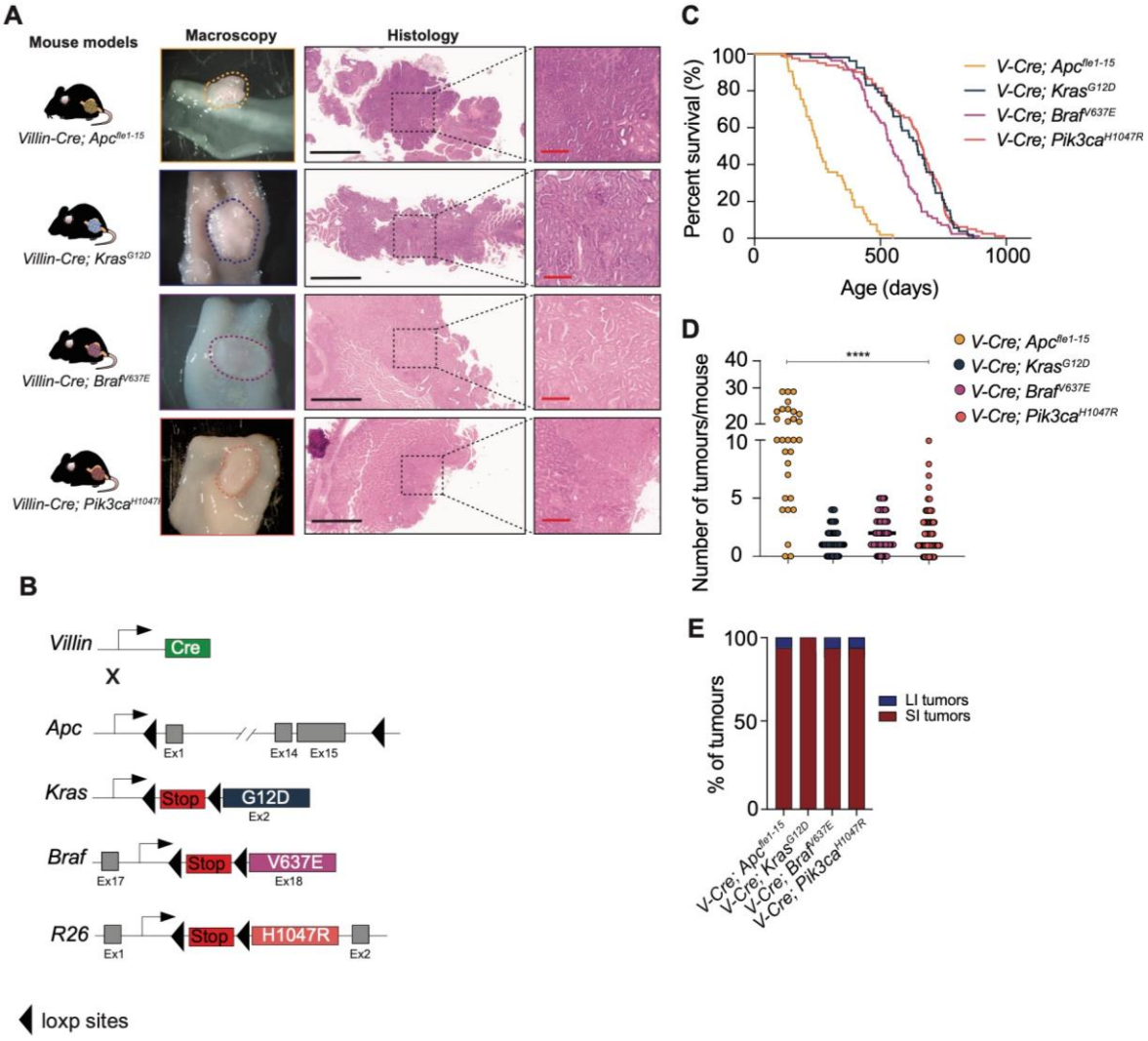


Figure 7 | Overview of GEMMs included in the study. All mouse alleles shown in this overview can be activated through the expression of Cre recombinase. *Vil-Cre* lines can be used to express Cre recombinase specifically in intestinal cells. **[A]** In the intestine of *Vil-Cre*; $Apc^{fle1-15}$ mice, exons 1-15 are deleted from the *Apc* gene leading to a complete absence of protein function (5). In this cohort, mice developed tumours with a polypoid appearance at the macroscopic level. Histologically these tumours are classified as classical tubulovillous adenomas shown by haematoxylin and eosin (H&E)-stained sections. The serrated CRC models were generated by the expression of mutant $Kras^{G12D}$ (8), $Braf^{V637E}$ (9) and $Pik3ca^{H1047R}$ (unpublished) in the intestine. Macroscopically these lesions present a serrated morphology characterized by a "saw-toothed" infolding of the colonic crypt epithelium which was confirmed histologically by H&E staining (scale bars, black 2 mm, red 300 µm). **[B]** Alleles used in this study. Upon crossing *Vil-Cre*-expressing mice with mice harbouring the desired genetic alteration, Cre recombinase recognizes the *loxP* sequences and, because they have the same orientation, excises the genetic material between them. Deletion of *Apc* and oncogenic expression of $Kras^{G12D}$and $Braf^{V637E}$ is driven from the endogenous locus upon deletion of exons 1-15 ($Apc^{fle1-15}$) or of the STOP cassette by the Cre recombinase ($Kras^{G12D}$and $Braf^{V637E}$). The transgene containing the mutation in *Vil-Cre; Pik3ca*^{H1047R} mice is located in the ubiquitously expressed *Rosa26* locus. **[C]** Kaplan-Meyer survival curves for the mice used in this study. Tumourigenesis is accelerated in *Vil-Cre*; $Apc^{fle1-15}$ mice (median survival of 40.3 weeks) compared to *Vil-Cre; Kras*^{G12D} (median survival of 89 weeks), *Vil-Cre; Braf*^{V637E} (median survival of 78 weeks) and *Vil-Cre; Pik3ca*^{H1047R} (median survival of 89 weeks). **[D]** Number of tumours per mouse upon necropsy. All mice included in the analysis were "endpoint" mice meaning they were sacrificed when termination criteria as a result of symptomatology caused by the tumour were met. *Vil-Cre*; $Apc^{fle1-15}$ mice developed an average of 12.8 tumours per mouse in comparison to 1.4 for *Vil-Cre; Kras*^{G12D} and 2.1 for both *Vil-Cre; Braf*^{V637E} and *Vil-Cre; Pik3ca*^{H1047R} (Kruskal-Wallis test $p<0.001$). **[E]** Percentage of tumours in the small intestine (SI) and large intestine/colon (LI), in all mice from the distinct cohorts. All the mouse models were generated and maintained together with Dr. Markus Tschurtschenthaler, Antonio Zaurito and Valentina Brunner from the group of Prof. Dieter Saur.

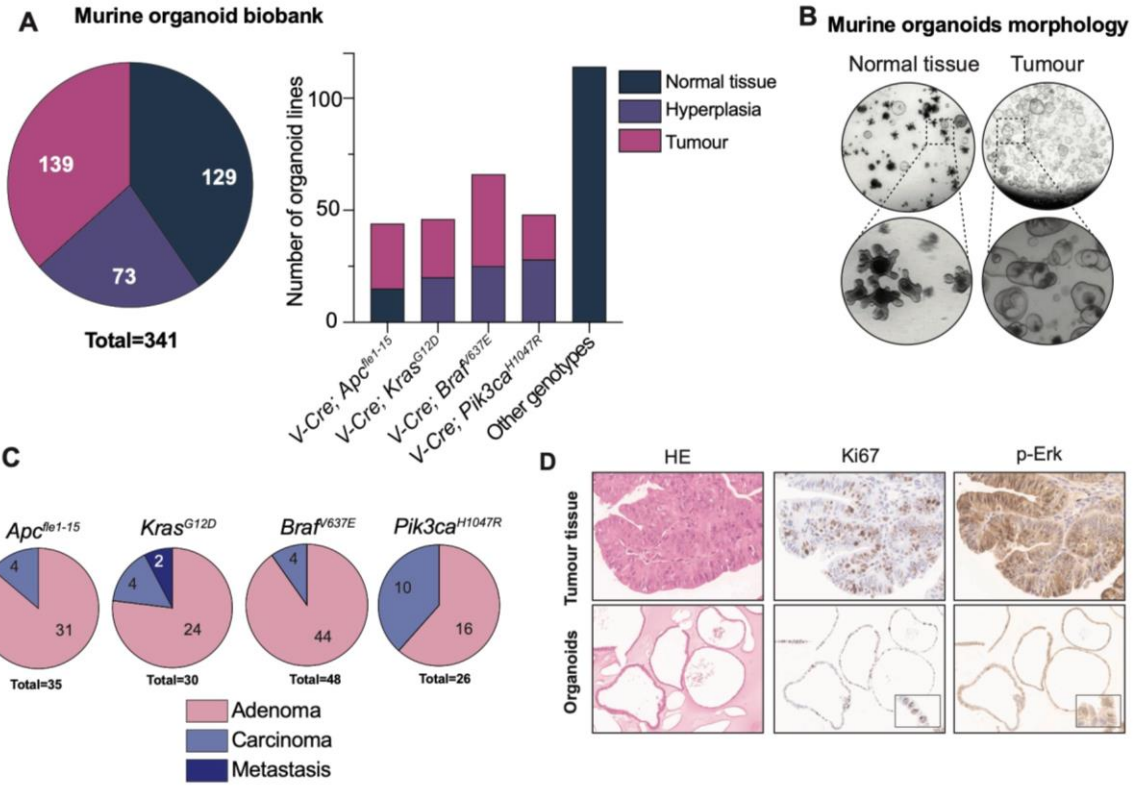
4.2 Generation of a murine and human organoid biobank that recapitulate features of *in vivo* tumours

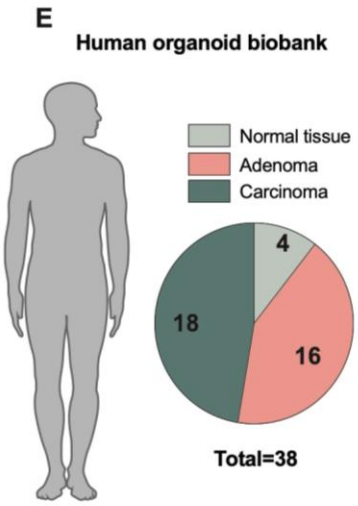
Since the establishment of a protocol enabling the use of organoids as an *in vitro* model (120), this cellular system provided novel opportunities to study tumours in a systematic and comparable manner, not possible until now. Because organoids can be also derived from healthy tissue, they provide the proper control setting to study neoplastic tissue from the same organism (119). In order to capture the *in vitro* biological features of both normal and tumour cells from the endogenous mouse models used in this study, we established a total of 129 normal tissue (65 from SI, 64 from LI), 73 hyperplasia (53 from SI, 20 from LI) and 139 tumour murine organoid lines derived from the intestine of *Vil-Cre*; $Apc^{fle1-15}$, *Vil-Cre;* $Kras^{G12D}$*, Vil-Cre;* $Braf^{V637E}$ and *Vil-Cre;* $Pik3ca^{H1047R}$ mice (**Figure 8A**). For the sake of simplicity, from here on the names of the genotypes will be referred only to as $Apc^{fle1-15}$, $Kras^{G12D}$, $Braf^{V637E}$ and $Pik3ca^{H1047R}$ implying that recombination has taken place leading to the activation of the oncogene or loss of the TSG in the intestine. Regarding normal tissue and hyperplasia samples, SI and colon organoids were isolated from the same animal given the reported differences between both anatomical locations (7). Morphological features of the different organoid lines were also captured in culture. Organoids derived from normal tissue differentiate after 3-5 days in culture and form the characteristic “budding” structures corresponding to the intestinal crypts also found *in vivo* in the intestine (120) (**Figure 8B**). Additionally, normal tissue organoids show increased levels of cell death which is in line with the fact that upon differentiation, the stem cells at the crypt base move towards the lumen of the organoid where eventually they undergo apoptosis. Dead cells accumulate in the lumen (center) of the organoid which can lead to the disruption of the whole structure if the culture is not maintained properly (e.g. passaging at the right time). On the other hand, organoids derived from tumour tissue typically do not differentiate but rather remain in a cystic form, possibly reflecting the increased proliferation status of the tumour from where they were derived (**Figure 8B**). We were also able to capture *in vitro* the different stages of tumour progression since organoids were established from normal tissue, hyperplasia, adenoma, carcinoma and metastasis samples from the different genotypes (**Figure 8C**). A total of 35 $Apc^{fle1-15}$ (31 adenomas, 4 carcinomas), 30 $Kras^{G12D}$ (24 adenomas, 4 carcinomas, 2 metastases), 48 $Braf^{V637E}$ (44 adenomas, 4 carcinomas) and 26 $Pik3ca^{H1047R}$ (16 adenomas, 10 carcinomas) tumour organoid lines were generated. Furthermore, we show that the isolated organoids were able to recapitulate key histological and expression markers features of the mouse tissue of origin (**Figure 8D**). Because organoids from normal and hyperplastic tissue are grown using the same culture

conditions as for tumour-derived organoids, it is imperative that upon mouse necropsy proper care is taken so that only tissue from the tumour lesion is harvested.

A total of 38 human organoid lines were additionally generated from tumour and normal tissue of CRC patients (4 from adjacent normal tissue, 16 from adenoma lesions and 18 from carcinomas) (**Figure 8E**). Adenomas in the cohort were graded as either tubular (31%) or tubulovillous (69%) regarding their histologic type and half of them are low grade (LG) and the other half were classified as high grade (HG) adenomas (**Figure 8F**). In addition, TNM staging as well as the pathological grading of the carcinomas are shown in **Figure 8F** bottom table. Sixty-one percent of the carcinomas in the cohort (11/18) are advanced lesions which showed involvement of lymph nodes or distant organs such as the liver, corresponding to stage III-IV disease. The remaining carcinoma samples were resected from localized stage I-II disease. In accordance to what was observed for murine organoid lines, also human organoids display different morphologies in culture according to the tissue from where they were derived. In **Figure 8G** 3 different morphologies are shown which correspond to organoids derived from normal tissue, adenoma and carcinoma respectively. Upon isolation of the intestinal crypts from the human intestine organoids grow as 3D structures from single crypts. After 3-4 days in culture, organoids differentiate giving rise to “mini-guts” containing a proliferative stem cell compartment at the bottom of the crypts. In resemblance to the murine context, also in human organoids stem cells move up from the bottom of the crypt and are actively shed to the lumen of the organoid after differentiation occurs.

Figure 8 | Organoid biobank from the endogenous mouse models and human patient tumour samples established in this work. [A] Overview of the murine organoid cohort (n=341) established from different tissue origins (normal tissue, hyperplasia and tumour) of the GEMMs included in this study. **[B]** Microscopic pictures depicting the differences in the morphology of normal tissue and tumour organoids. Organoids are cultured under the same conditions and supplemented with 50% L-WRN medium. Normal tissue organoids differentiate *in vitro* as can be seen by the typical “budding” appearance that resemble intestinal crypts, and continuously shed dead cells to its lumen. Tumour organoids present a cystic shape and do not differentiate. **[C]** Overview of the number of murine organoid lines established from different tumour lesions (adenoma, carcinoma, metastases) from the different genotypes included in this study. **[D]** H&E and immunohistochemistry (IHC) stainings of primary tissue and the respective derived organoid lines. Besides morphological recapitulation of the *in vivo* setting, organoids also retain protein marker expression from the respective tissue of origin. **[E]** Human tumour organoid biobank established in the scope of the work for this thesis (n=38). Organoids were derived from different tissues (normal mucosa, adenoma and carcinoma) from cancer patients at the Klinikum rechts der Isar. The clinical aspects of both carcinomas and adenomas are shown in the tables from **[F]**. Carcinomas were derived from different disease stages (I-IV) (4) covering most histological subtypes. Adenomas included in the cohort were either low (LG) or high grade (HG) and presented a tubular or tubulovillous architecture (NOS, not otherwise specified). **[G]** Human organoids morphology *in vitro*. Individual crypts are isolated from a normal (healthy) intestine, embedded in Matrigel and used for subsequent organoid culture. Normal tissue organoids start to differentiate under the right culture conditions while adenoma-derived organoids remain cystic. Carcinoma derived organoid lines show distinct morphological aspects that can be associated with different histological subtypes (optical zoom of 12.5x for upper pictures and 50x for lower pictures). Human tumour samples were obtained from Dr. Moritz Jesinghaus from Klinikum Rechts der Isar. Murine organoid samples were generated with the help of Dr. Markus Tschurtschenthaler, Antonio Zaurito and Valentina Brunner from the group of Prof. Dieter Saur.





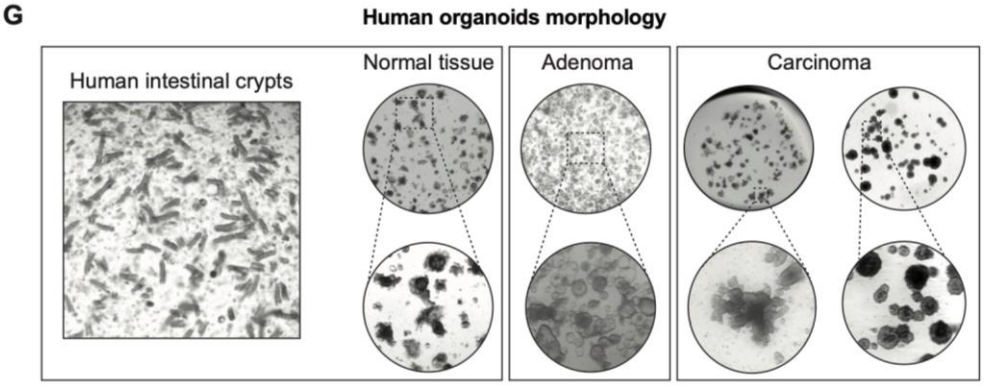
F

Carcinomas

Line ID	Pathological grading	Histological subtype
CRC-R11	UICC Stage I, G1	NOS
CRC-R22	UICC Stage III, G2, MSS	Serrated adenocarcinoma
CRC-R14	UICC Stage III, G3, MSS	Micropapillary adenocarcinoma
CRC-R15	UICC Stage III, G3	NOS
CRC-R17	UICC Stage III, G3, MSS	Serrated adenocarcinoma
CRC-R12	UICC Stage IV, G3, MSS	NOS
CRC-R25A	UICC Stage I, G1	Adenoma-like adenocarcinoma
CRC-R16	UICC Stage IV, G3, MSS	Micropapillary adenocarcinoma
CRC-R25B	UICC Stage IV, G2	NOS
CRC-R13	UICC Stage III, G2	NOS
CRC-R19	UICC Stage II, G2, MSI	NOS
CRC-R30	UICC Stage I, G2	NOS
CRC-R32	UICC Stage I, G2	NOS
CRC-R20	UICC Stage II, G2, MSI	Micropapillary adenocarcinoma
CRC-R23	UICC Stage III, G3	NOS
CRC-R21	UICC Stage III, G3, MSS	Medullary adenocarcinoma
CRC-R18	UICC Stage IV, G3, MSS	Mucinous adenocarcinoma
CRC-R7	UICC Stage IV, G2, MSS	Serrated adenocarcinoma

Adenomas

Line ID	Histological type
A8	Tubular, LG
A10	Tubulovillous, LG
A11	Tubulovillous, LG
A14	Tubulovillous, LG
A15	Tubular, LG
A17	Tubulovillous, LG
A18	Tubulovillous, HG with imCA
A19	Tubulovillous, up to HG
A20	Tubular, up to HG
A21	Tubular, LG
A23	Tubulovillous, HG
A24	Tubulovillous, HG
CRC-R27 A2-2	Tubulovillous, up to HG
CRC-A26	Tubulovillous, up to HG
A27	Tubulovillous, up to HG
A28	Tubular, LG



Organoids derived from adenoma lesions show a stark different morphology which consists mostly of cystic structures as seen in the mouse counterpart (**Figure 8G**). Adenomas are characterized by increased proliferation usually through disruption of Wnt pathway (56). Carcinomas, however, are more progressed lesions which as discussed before have acquired additional mutations in genes belonging to pathways such as MAPK and TGFβ. Therefore, the latter are not necessarily more proliferative than adenomas, but are rather provided with features that facilitates their invasion into the next intestinal layers. Carcinoma-derived organoids lose the proliferative morphology seen in adenomas. Instead, they present distinct morphologies in culture which can reflect the histological features of the tumours from where they were isolated from. Additionally, both adenoma and carcinoma organoid cultures can be maintained for longer periods (5-7 days) before passaging in comparison to normal tissue organoids (3-4 days).

4.3 Progression of murine serrated tumours show discrete chromosomal aberrations and require Wnt pathway activation

In order to genetically characterize the murine serrated organoids from our cohort we performed whole exome sequencing (WES) in 12 *Kras*^{G12D} (6 adenomas, 4 carcinomas and 2 liver metastases), 16 *Braf*^{V637E} (15 adenomas, 1 carcinoma) and 2 *Pik3ca*^{H1047R} (1 adenoma, 1 carcinoma) organoid lines (**Figure 9A**). The oncoplot from **Figure 9A** shows the recurrent mutated genes (mutated in at least 2 independent samples) in the cohort. In contrast to other sample types (e.g. FFPE and whole tissue piece) the use of organoids allows the assessment of genetic features in pure epithelial cancer cells where no other cells are present that would affect tumour purity. In addition to the organoids, tail DNA from each corresponding mouse was used as a reference control for the identification of germline variants in each animal. In the overall cohort we detected a total of 1102 mutations where the majority (90%) resulted in the generation of a codon that specifies a different amino acid (*missense*). Seventy percent (21/30) of samples showed disruption of Wnt pathway either by *Ctnnb1* or *Apc* mutations. All mutations in *Ctnnb1* were missense consistent with a gain of function effect. In fact, *CTNNB1* mutations in human cancers are more prevalent in MSI-H tumours (169) which is in line with the serrated nature of the samples included in this study. In contrast, 2 *nonsense Apc* mutations were detected which most likely render the final protein product not functional. Mutations in *Apc* and *Ctnnb1* were mutually exclusive. These findings are in line with the work of Rad et al., (2013) where *Braf*^{V637E}-driven tumours were found to activate Wnt signalling in later stages of tumour progression (9). In contrast, mutations in Wnt pathway genes were not previously described in *Kras*^{G12D}-driven mouse tumours (8). However, in human CRC cells derived from advanced cancers, oncogenic

activation of *KRAS* is necessary for β-catenin nuclear translocation and upregulation of Wnt signalling (170). Mutations in Wnt-related genes were not detected in the 2 *Pik3ca*^{H1047R} samples. Interestingly, excluding the 2 *Pik3ca*^{H1047R} organoid lines, wt samples for Wnt signalling genes and *Trp53* ("TM7513_T1-2", TM7510_T2-1", "TM2118_T4", "TM1541_T1" and "TM1179_T1-3") showed much lower number of mutations and lacked alterations in other recurrent genes. Interestingly, although sample "TM1918_T1" showed no mutations in Wnt-related genes, it had a *missense* mutation in *Fat4* (not shown). Loss of function mutations in *FAT4* are associated with Wnt signalling activation in gastric cancer (171). Therefore, we speculate this mechanism to be the one responsible for Wnt pathway upregulation in this particular sample. Overall, mutations in *Trp53* were present in 4/8 (50%) carcinomas and in 3/22 (14%) of adenomas. *TP53* mutations are found in about 60% of human CRC (55) and this is usually regarded as a late stage genetic event. Nonetheless, inactivating mutations in *TP53* can also occur in adenomas albeit to a lesser extent (172) as seen in our adenoma cohort. Concomitantly, mutations in *Ttn* were detected in 30% (9/30) of samples in the cohort including both adenoma and carcinoma samples. Mutations in *TTN* were recently associated with increased TMB in human CRC (173). In the organoid cohort used in this study we did not report this correlation (Welch Two Sample t-test, p-value = 0.1114). Besides the aforementioned genes (Wnt-related genes *Trp53* and *Ttn*) there were no recurrent mutated genes in the organoid cohort. Additional single nucleotide variants (SNVs) were found only in 7-17% of samples in genes involved in different biological processes such as cell motility and ECM degradation (*Col11a1, Lama4, Eml4*), transcription regulation (*Junb, Aff3*) and cell transport (*Gria4, Slc44a5*).

Of note, samples "MT577 T1-2", "MT577 LivMet1" and "MT577 LivMet2" (Series 1) correspond to organoids derived from the primary tumour and 2 liver metastases respectively from the same mouse. Even though they correspond to different samples, they could be barely distinguished genetically both at the single nucleotide variant (SNV) and copy number variation (CNV) level. Another comparison of the same sort was performed in 1 carcinoma and 1 lung metastasis-derived organoid line harvested from the same recipient mouse implanted with a *Pik3ca*^{H1047R} organoid line (Series 2). Also here, samples shared the same alterations and similar CNV profiles (**Figure 9G**).

Using the output from WES data we further looked at the pattern of single base substitutions (SBSs) in order to understand if differences were found between different stages of tumour progression (**Figure 9A**). No major differences can be seen across samples between different genotypes and tissue origin. Interestingly however, samples "TM7513 T1-2", "TM7510 T2-1", "TM2118 T4", "TM1541 T1"and "TM1179 T1-3" show very few to none T>C and T>A transitions which are present in a much higher percentage in all other samples. Possible explanations for these findings will be explored in the **Discussion**

chapter. Furthermore, as an exploratory analysis we crossed the SBSs pattern detected in the organoid cohort to the reference signatures generated by Alexandrov and colleagues (10). Based on the number and pattern of the six possible base pair substitutions in cancer genomes (C>A, C>G, C>T, T>A, T>C, and T>G), the authors defined 67 signatures (SBSs) that could be attributed to several cancer types. Specifically for CRC, SBS1, SBS5 and SBS40 were reported to be the most common mutational signatures. However, performing the same analysis with our dataset retrieved SBS6 and SBS17 as the most common signatures (**Figure 9H**). Although no aetiology is attributed to SBS17, SBS6 is caused by defective DNA mismatch repair (10). Our cohort is mainly composed of organoids derived from serrated lesions. Given the strong association between the MSI phenotype (which is caused by errors in the DNA mismatch repair machinery) and the serrated pathway of CRC (84), it is not surprising to find SBS6 to be the strongest associated signature with our dataset.

To further characterize genetically the organoids included in this study, we looked at chromosomal aberrations present in these tumours. For that, we performed low-coverage whole genome sequencing (lcWGS) of 26 *Kras*^{G12D} (20 adenomas, 4 carcinomas and 2 liver metastases), 39 *Braf*^{V637E} (35 adenomas, 4 carcinomas) and 11 *Pik3ca*^{H1047R} (4 adenomas, 7 carcinomas) organoid samples. The graphs in **Figure 9B** depict the frequency of CNVs found in each group of samples from the different genotypes. Each graph shows the percentage of samples with a given chromosomal alteration (amplification or deletion) from a defined histological grade and genotype. Also at this level, samples showed overall, high levels of heterogeneity which hampered the attribution of alterations to a given group. Overall, adenomas showed low levels of aneuploidy regardless of the genotype which is in line with the literature (174). Focal deletions in chromosome 3 and 4 and focal amplifications of chromosome of 7, 8 and 19 were seen in half (2/4) of the *Pik3ca*^{H1047R} -driven adenomas while in carcinomas most alterations were focal deletions in chromosome 4 (harbouring the *Cdkn2a* locus) and focal amplifications in chromosome 14 in 3/7 samples. In *Kras*^{G12D}-driven lesions, we found focal amplifications in chromosome 4, 5, 6, 12 and 18 to be present in 20% (4/20) of adenomas whereas focal deletions in chromosome 1, 7, 9 and 16 were present in 3/6 carcinoma samples and deletions in chromosome 4 (containing the *Cdkn2a* locus) were detected in 4/6 carcinomas. Additionally, amplifications of chromosome 3, 6 and 12 were detected in 3/6 samples and amplification in chromosome 14 detected in 4/6 carcinomas of this cohort.

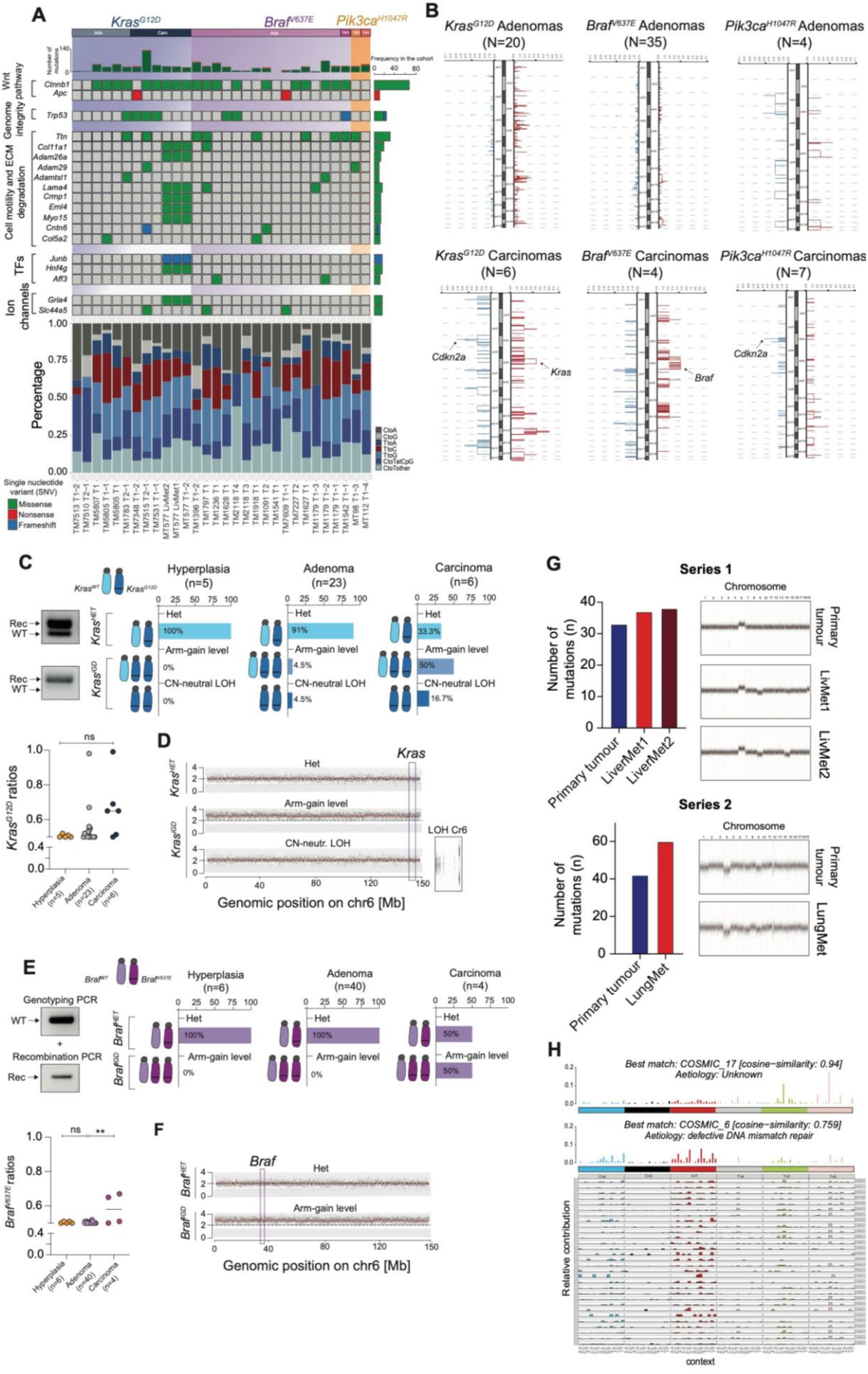

A
KrasG12D
BrafV637E
Pik3caH1047R
Wnt pathway
Genome integrity
Cell motility and ECM degradation
TFs
Ion channels
Ctnnb1
Apc
Trp53
Ttn
Col11a1
Adam26a
Adam29
Adamtsl1
Lama4
Crmp1
Eml4
Myo15
Cntn6
Col5a2
Junb
Hnf4g
Aff3
Gria4
Slc44a5
Percentage
Single nucleotide variant (SNV)
Missense
Nonsense
Frameshift
B
KrasG12D Adenomas (N=20)
BrafV637E Adenomas (N=35)
Pik3caH1047R Adenomas (N=4)
KrasG12D Carcinomas (N=6)
BrafV637E Carcinomas (N=4)
Pik3caH1047R Carcinomas (N=7)
Cdkn2a
Kras
Braf
C
Hyperplasia (n=5)
Adenoma (n=23)
Carcinoma (n=6)
Het
Arm-gain level
CN-neutral LOH
KrasG12D ratios
ns
D
Kras
Het
Arm-gain level
CN-neutr. LOH
LOH Cr6
Genomic position on chr6 [Mb]
E
Hyperplasia (n=6)
Adenoma (n=40)
Carcinoma (n=4)
Genotyping PCR
Recombination PCR
BrafV637E ratios
F
Braf
G
Series 1
Series 2
Number of mutations (n)
Primary tumour
LiverMet1
LiverMet2
LungMet
Chromosome
H
Best match: COSMIC_17 [cosine-similarity: 0.94]
Aetiology: Unknown
Best match: COSMIC_6 [cosine-similarity: 0.759]
Aetiology: defective DNA mismatch repair
Relative contribution
context

Figure 9 | Genetic characterization of murine serrated tumour organoids. All analyses were performed in organoid lines derived from endogenous GEMMs. **[A]** Oncoplot depicting the most common single nucleotide variants (SNVs) across samples in the different genotypes (*Kras*^{G12D} n=12, *Braf*^{V637E} n=16, *Pik3ca*^{H1047R} n=2) as a result of whole exome sequencing (WES) analysis. SNVs were filtered for variant allele frequency (VAF) of at least 10% and read coverage of at least 20 in both the control and the tumour sample. Only genes recurrently mutated (>2 independent samples) are shown. Mutations in Wnt signalling (*Ctnnb1* and *Apc*) were the most common events followed by *Trp53* and *Ttn* mutations (upper part). Single base substitution profiles for each sample in the cohort are shown in the lower part of the figure. **[B]** Copy number variation (CNV) profiles of tumour organoids from different disease stages from the genotypes in the cohort as a result of low coverage whole genome sequencing (lcWGS) analysis. Each graph shows the frequency of a given chromosomal alteration in the indicated cohort subset (adenoma, carcinoma). Deletions in chromosome 4 harbouring the *Cdkn2a* locus are common events in *Kras*^{G12D} and *Pik3ca*^{H1047R} carcinomas while amplification of chromosome 6 was found predominantly in *Kras*^{G12D} and *Braf*^{V637E}-driven carcinomas. **[C]** Mechanisms of oncogenic amplification in *Kras*^{G12D}-driven samples. Two distinct oncogenic states were found in these samples: *Kras*^{HET}, where no alteration at the *Kras* locus occurred, or *Kras*^{iGD} where oncogenic imbalances are observed. Oncogenic amplification can be qualitatively assessed by PCR where the absence of a wt band at the *Kras* locus hints that only the recombined allele is present and therefore iGD is occurring. Two biological mechanisms were found to lead to iGD (chromosome 6 arm-gain level and copy number (CN) neutral loss of heterozygosity (LOH)) happening less frequently in adenomas and often in carcinomas (p=0.109, Mann-Whitney U test). By amplicon deep sequencing of the *Kras locus* in *Kras*-mutant samples its possible quantitatively define the amount of recombined and wt reads (*Kras*^{G12D} ratios). **[D]** Genetic mechanisms leading to alterations in *Kras* gene dosage were analysed on the basis of lcWGS and WES. Whole chromosome amplification of chromosome 6 (trisomy) involving the *Kras*^{G12D} locus happens most likely through mitotic missegregation. LOH of only a distal region of chromosome 6 indicates that this alteration probably results from mitotic recombination. **[E]** Mechanisms of oncogenic amplification in *Braf*^{V637E}-driven samples. Also, in this setting two distinct oncogenic states were detected: *Braf*^{HET}, where no alteration at the *Braf* locus occurred, or *Braf*^{iGD} where oncogenic imbalances were specifically found in advanced stages of disease (**p≤ 0.01, Mann Whitney test). Qualitative assessment of recombination can be performed in 2 separate PCR reactions, however, information regarding the amplification state of the oncogene cannot extracted by this method. By amplicon deep sequencing of the *Braf locus* it is possible to quantitatively identify the number of recombined and wt reads (*Braf*^{V637E} ratios). Arm-gain level of chromosome 6 is the biological mechanism responsible for iGD and it is found exclusively in carcinoma samples. **[F]** Genetic mechanisms leading to alterations in *Braf* gene dosage were analysed on the basis of lcWGS and WES. Whole chromosome amplification of chromosome 6 (trisomy) involving the *Braf*^{V637E} locus happens most likely through mitotic missegregation. **[G]** Genetic comparison of organoids derived from primary tumour and metastases from the same mouse. In series 1, the number of mutations in metastases (n=37, n=38) was slightly increased compared to the primary tumour (n=33) and CNV profiles showed that features of primary tumour were recapitulated in metastases. In series 2, referring to a different mouse, the same is shown (primary tumour n=42 mutations, LungMet n=60 mutations) and CNV profiles are very similar in both samples **[H]** Comparison of the single base substitution (SBS) profile from our organoid cohort with the published cancer signatures dataset by Alexandrov et al., (10). Our dataset was strongly associated with mutational signature SBS17 (co-sine similarity=0.94) and SBS6 (co-sine similarity=0.76). WES and lcWGS raw data were analysed by Sebastian Lange and Niklas Krätzig.

While no recurrent alterations were found in *Braf*^{V637E}-driven adenomas, we saw that half of the carcinomas (2/4) showed also amplifications of chromosome 6 similarly to what was seen in carcinomas from the *Kras* cohort. Aneuplody is a common feature of advanced human CRC and it was reported that CNVs increase drastically from early to late-stage disease (146). Moreover, adenomas and early stage carcinomas have comparable levels of CNVs (175) while advanced CRC and metastatic lesions show higher levels of similarity (176). In the present dataset conclusions of this nature are difficult to infer given the low number of samples in the carcinoma groups. However, we report more alterations in the carcinomas in comparison with adenomas as well as a selection for specific genetic events (e.g. amplification of chromosome 6) in both *Braf*^{V637E} and *Kras*^{G12D}-driven carcinoma samples.

4.4 Increased gene dosage of *Kras^{G12D}* and *Braf^{V637E}* is observed during serrated CRC progression

The analysis of lcWGS data from serrated organoids derived from adenoma and carcinoma samples revealed that 4.5% of adenomas and 50% of *Kras*-driven carcinomas show amplification of the *Kras* locus (**Figure 9B**). Amplification of the *KRAS* locus has been reported in a small percentage of CRC patients and was shown to confer resistance to anti-EGFR therapy in these patients (177, 178). Additionally, loss of heterozygosity (LOH) analysis based on SNP information retrieved from WES data revealed that chromosome 6 was also affected by LOH in 1 adenoma and 1 carcinoma organoid line. Combining the data from CNV and LOH analysis, we observed that in total 2 adenoma (9%) and 4 carcinoma lines (66.7%) harboured genetic alterations of the *Kras* locus (**Figure 9C**). These alterations reflected 3 different states of the *Kras^{G12D}* gene: "Het"- when no change to the heterozygous state of the allele is made; "Arm-gain level"- when an extra copy of chromosome 6 is gained; and "CN-neutral LOH"- copy number neutral loss of the *Kras wt* allele with no alterations to the number of copies of the mutant allele. Allelic imbalances of the *Kras* locus results in increased gene dosage of the mutant *Kras^{G12D}* gene (referred to as *Kras^{iGD}*) (**Figure 9D**). Two main mechanisms can generate *Kras$^{G12D\text{-}LOH}$* (homozygosity of *Kras^{G12D}*): (i) mitotic recombination of the *Kras* locus or (ii) mitotic error resulting in gain of an additional copy of chromosome 6 with *Kras^{G12D}* and subsequent loss of the wt allele at the same *locus.* Arm-level gain of *Kras* locus through amplification of the whole chromosome 6 is, most likely, the result of chromosome missegregation during mitosis (152). In our group we developed ways for the quantitative and qualitative assessment of *Kras^{G12D}* imbalances without the need for expensive whole genome/exome approaches. Based on the band pattern on the agarose gel following a PCR reaction it is possible to qualitatively assess the extent of the loss of the wt *Kras* allele and therefore predict imbalances of the mutated allele (**Figure 9C** and **Figure 13F**). In parallel, amplification of the *Kras* locus by next generation sequencing (NGS) using a primer pair spanning both the wt and the mutated allele allows the quantification of the number of mutant and wt *Kras* copies in a given sample (see Materials and Methods section). By calculating the ratio between mutant and wt *Kras* copies it is possible to assess the oncogenic dosage status in the sample (hence called *Kras ratios*). Typically, a ratio of 0.5 would indicate the same number of copies of mutated and wt *Kras* alleles are present in the sample (Het), while values >0.5 indicate the presence of iGD. Combining the output from 3 different technical approaches (CNV, WES and *Kras ratios*) assessment of allelic imbalances was performed in a subset of 5 *Kras*-induced hyperplasia, 23 adenoma and 6 carcinoma-organoid lines (**Figure 9C**).

Allelic imbalances at the *Kras* locus typically occurred during adenoma-carcinoma transition thereby suggesting that there is a selective pressure towards a *Kras*^{iGD} state during CRC progression. Although the analysis turned out not to be statistically significant (Hyperplasia vs Adenoma, Mann-Whitney U test, p=0.897, Carcinoma vs Adenoma, Mann-Whitney U test, p=0.109) there is a clear trend towards *Kras*^{iGD} samples in the carcinoma group. Of note, the adenoma-derived sample "TM1783 T2-1" showed *Kras*^{iGD} acquired through CN-neutral LOH (*Kras ratios*=0.99). Histologically this sample presented mostly features of high-grade adenoma but also focal areas of advanced disease. Therefore, it cannot be ruled out that the part harvested for organoids isolation although graded as an adenoma, did not contain any advanced tumour clones (due to sampling bias) that were cultured and used for the genetic analysis.

We then inquired if similar mechanisms would also affect the other 2 genotypes of the serrated cohort. *Pik3ca*^{H1047R} samples, as previously mentioned, showed no recurrent alterations at the CNV level at the *Pik3ca* or *Rosa26* (genetic location of the transgene) locus. However, half of the *Braf*^{V637E}-driven carcinomas showed amplifications of chromosome 6 where *Braf* is located. By developing a similar strategy as the one used for the assessment of the *Kras ratios* we developed a NGS assay that allowed the calculation of the *Braf ratios* (see Materials and Methods section). Qualitative assessment of *Braf ratios* however is not possible with a PCR approach due to the fact that amplification of the wt and recombined alleles have to be performed in separate reactions (**Figure 9E**). Also here, by combining the data from CNV, WES and *Braf ratios* we assessed the presence of allelic imbalances in a cohort consisting of 6 hyperplasia, 40 adenoma and 4 carcinoma-derived *Braf*-driven organoid lines. Allelic imbalances were reported only in the carcinoma cohort (n=2) which suggests that iGD might be necessary for late tumour progression (Hyperplasia vs Adenoma, Mann-Whitney U test, p=0.381, Carcinoma vs Adenoma Mann-Whitney U test, p=0.0067). Interestingly, unlike in the *Kras*^{G12D} setting, we did not report CN-neutral LOH events leading to iGD in the *Braf*-driven samples. Instead, only chromosome 6 arm-gain level events were detected (**Figure 9F**) which might suggest different routes of disease progression between the 2 models. Nonetheless, analysis of more samples would be necessary to support this hypothesis.

In fact, as mentioned several times during this work, the lack of advanced lesions (carcinomas and distant metastases) in these mouse models hampers the identification of the events leading to tumour progression. Even with a long tumour latency time, mice in this study do not develop carcinomas and metastases (**Figure 8C**).

4.5 Generation of an orthotopic tumour mouse model to study tumour progression

To overcome some of the limitations inherent to the use of endogenous tumour mouse models including i) long tumour latency, ii) lack of advanced disease and metastasis formation and iii) incorrect location for tumour development (in comparison to the human counterpart) we established an orthotopic mouse model based on the submucosal injection of a cell suspension in the colon of mice for *in vivo* modulation of CRC (**Figure 10A**). This protocol is based on the previously published study by Roper J. et al., (2017) which showed that genetically edited organoids give rise to adenomas upon orthotopic injection that later progressed to carcinoma and distant disease (2). In order to validate this approach, we isolated organoids from the intestine of wt mice and performed genetic *knockout* (*KO*) of the *Apc* gene by CRISPR/Cas9 technology (**Figure 10B**). We cloned a previously validated sgRNA designed to target exon 16 of *Apc* into a lentiviral vector containing the Cas9 coding sequence (2). Cells were transduced *in vitro* and edited organoids ($Apc^{\Delta Ex16}$) were subsequently selected by the removal of Wnt pathway agonists from the culture medium (Wnt3a and R-spondin) and expanded under these conditions. In contrast, Apc^{wt} organoids stop proliferating and do not survive under these conditions due to lack of Wnt activation necessary for stem cell maintenance. Interestingly, edited organoids changed their morphology *in vitro*, from a differentiated state into a cystic phenotype (**Figure 10B**) which is consistent with increased stemness also reported for other tissue contexts (179). Confirmation of the genetic *KO* was performed by NGS of the genomic sequence flanking the sgRNA target locus to identify the presence of insertions and deletions (*indels*) in $Apc^{\Delta Ex16}$ organoids. As a result of CRISPR/Cas9 editing almost all reads (98.7%) in $Apc^{\Delta Ex16}$ organoids showed *indels* at, or near, the target site (**Figure 10B**). After confirmation of the *KO,* organoids were orthotopically injected in the colonic submucosa of 12 mice and monitored for tumour formation (**Figure 10C**). Mice developed tumours 5-40 weeks after transplantation with 42% of mice (5/12) developing adenocarcinomas. Organoids were further derived from these lesions. The remaining mice from this experiment that did not develop carcinomas, had dysplastic lesions at the site of implantation without clear invasion. In order to understand if also at the genetic level there would be differences between $Apc^{\Delta Ex16}$ organoids from normal tissue and carcinoma-derived organoids from an implanted tumour we performed WES in both states (before and after transplantation) of the same organoid line. We detected a total of 18 SNVs in the parental organoid line while 138 SNVs were present in the organoids derived from the carcinoma lesion (**Figure 10C**). Of note, more than half of the mutations found in the implanted line had a very low variant allele frequency (VAF) (<10%) which suggests that the lesion is composed of many different

subclonal tumour cell populations. We found 13 mutations to be shared between both lines. Interestingly, the effect of *Apc KO* in the parental organoid line led to the acquisition of additional mutations in an otherwise wt background that were mostly retained after orthotopic transplantation.

An illustration of the injection procedure is depicted in the bottom panel of **Figure 10A**. Briefly injection of the organoid suspension (70-100 µL) forms a "bubble" that ideally closes the intestinal lumen and is later absorbed to prevent luminal obstruction of the colon of the animal. To prove that, using this experimental approach, cells engraft at the site of injection and do not leak to other locations in the intestine, we injected GFP-expressing organoids orthotopically into a mouse. Upon animal necropsy we detected GFP signal by fluorescent microscopy only at the site of injection which was then confirmed by histology.

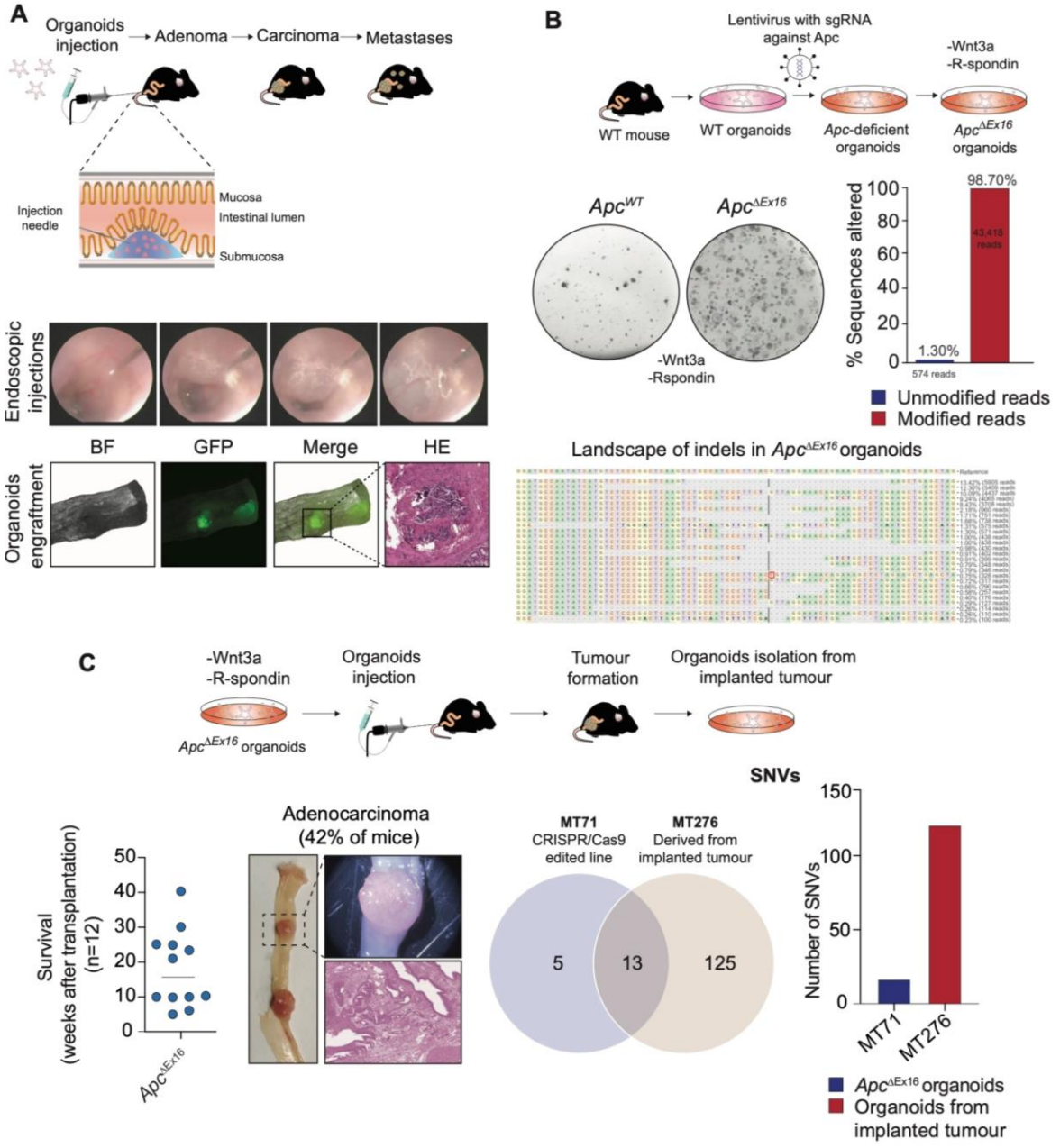


Figure 10 | Establishment of an orthotopic CRC mouse model using organoids. [A] Organoid injection is performed by colonoscopy using a camera-assembled endoscope through which a fine needle is inserted to allow the injection of organoids into the mouse colon submucosa. Injection of 70-100 µl of an organoid suspension in 10% Matrigel leads to the formation of a bubble in the intestinal lumen which is absorbed within minutes. *Ex vivo* analysis of the intestine following orthotopic transplantation of a GFP-expressing organoid line shows that cells engraft in a confined area which was confirmed by immunofluorescence of the whole mouse intestine and histologically by H&E staining. **[B]** Genetic *knockout* of *Apc* in wt organoids by CRISPR/Cas9. Intestinal organoids were isolated from a wt mouse and expanded *in vitro*. We cloned a previously validated sgRNA targeting *Apc* (2) into the plentiCRISPRv2 vector. Edited organoids were then selected in culture by the removal of Wnt pathway agonists (Wnt3a, R-spondin) that are present in standard culture conditions (L-WRN medium). Amplicon-based deep sequencing of the sgRNA target genomic site in $Apc^{\Delta Ex16}$ organoids revealed extensive editing at the *Apc* locus with indels present in almost all reads. The most common alteration found was a 37bp deletion in 13.42% of reads. **[C]** Mouse orthotopic injection of $Apc^{\Delta Ex16}$ organoids. Mice were injected with $Apc^{\Delta Ex16}$ organoids (n=12). Termination criteria were met after 18 weeks on average and 42% of mice developed adenocarcinomas at the site of injection that were visible macroscopically and confirmed histologically. Organoids were derived from these lesions and WES analysis was performed in both $Apc^{\Delta Ex16}$ parental and carcinoma-derived organoids. A total of 18 SNVs were detected in the parental line in comparison to 138 SNVs in the carcinoma-derived line. Thirteen SNVs were shared between the 2 samples possibly indicating that these mutations are important for the progression of this tumour. Establishment of the orthotopic injections model was done together with Dr. Markus Tschurtschenthaler from the group of Prof. Dieter Saur.

4.6 Murine adenomas show different levels of progression *in vivo* following orthotopic transplantation of tumour-derived organoids

4.6.1 Histologic progression

Following the confirmation that organoids can engraft and form tumours subsequent to mouse colonic orthotopic transplantations, we set to explore this model to study progression of early to advanced disease using the different genotypes of our cohort (**Figure 11A**). We orthotopically injected adenoma-derived organoids into either immunocompetent (C57BL/6J) or immunodeficient (NSG) mice. Twelve mice were injected with $Apc^{fle1-15}$, 45 with $Kras^{G12D}$, 17 with $Braf^{V637E}$ and 1 with $Pik3ca^{H1047R}$ adenoma organoid lines using 3 different $Apc^{fle1-15}$, 12 $Kras^{G12D}$, 4 $Braf^{V637E}$ and 1 $Pik3ca^{H1047R}$ adenoma-derived lines respectively (**Figure 11B**). A distinct tumour penetrance was observed depending on the genotype of the implanted line. Organoids´ engraftment and subsequent tumour formation was achieved in a high percentage of animals which was higher for serrated lesions (50% for $Apc^{fle1-15}$ n=6, 89% for $Kras^{G12D}$ n=40, 88% for $Braf^{V637E}$ n=15 and 100% for $Pik3ca^{H1047R}$ n=1). Strikingly, progression from adenoma-carcinoma was achieved in 50% (Apc^{fle1-1} , n=3), 63% ($Kras^{G12D}$, n=25), 40% ($Braf^{V637E}$, n=6) and 100% ($Pik3ca^{H1047R}$, n=1) of mice that developed tumours. Furthermore, the presence of metastasis was reported in a subset of mice from all genotypes ($Apc^{fle1-15}$ n=2; $Kras^{G12D}$ n=13; $Braf^{V637E}$ n=4; $Pik3ca^{H1047R}$ n=1). Every lesion that formed as a result of transplantations, was divided into halves. One half was fixed in formalin and added to our biobank of archived tissue (FFPE) and the other half was used for organoids isolation. Downstream purposes of the isolated organoids included genetic (WES, lcWGS, *Braf* and *Kras ratios* assessment) and transcriptomic (3´pA RNA-seq) analysis. In a small number of cases organoids did not

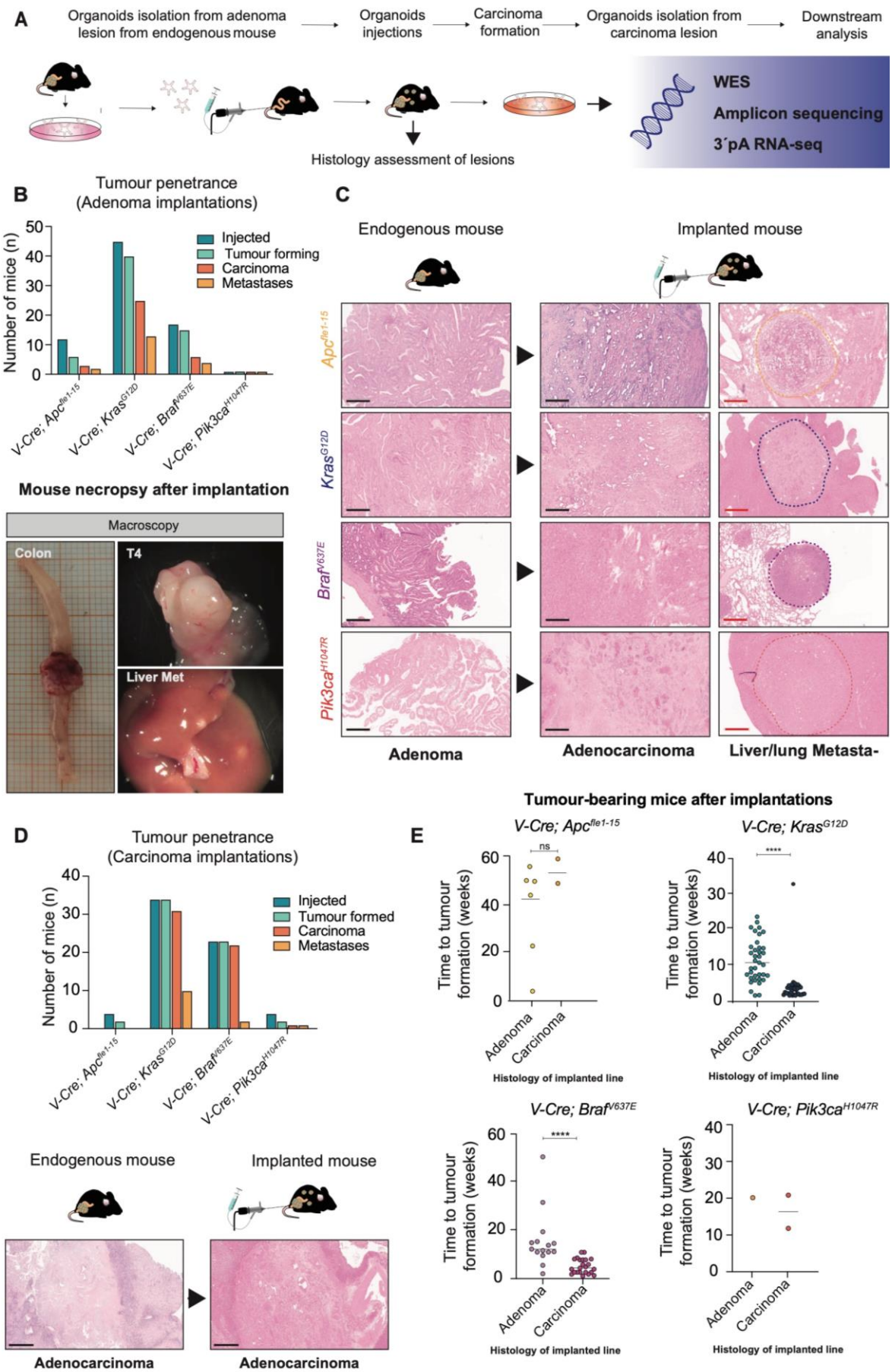
engraft in the colon of the mouse. We believe this might be related to the less aggressive nature of the endogenous adenoma line (**Figure 11B**).

CRC usually metastasizes first to the liver, then to the lungs and from here to the rest of the body (180). Our orthotopic model faithfully recapitulates these features of advanced disease in the mouse. The bottom part of **Figure 11B** depicts a representative example of a picture of a mouse necropsy after implantation with an adenoma-derived organoid line that consequently developed an adenocarcinoma (T4) and several liver metastases (Liver Met). In some instances, tumour formation can cause luminal obstruction (leading to constipation) of the mouse colon preventing normal digestion and faeces evacuation from taking place in the animal. *In vivo* progression of the implanted adenomas into carcinomas was confirmed histologically as well (**Figure 11C**) where both the primary tumour and associated metastases (liver and/or lung) were graded in transplanted mice. The comparison between the adenoma lesion in endogenous mice from where the organoids were derived from, with the carcinoma lesion in implanted mice after organoid transplantation clearly shows an increased histological grade for all genotypes. All adenomas had features of intraepithelial neoplasia that grew along the pre-existing epithelium with no signs of invasion. Adenocarcinomas showed invasion of additional tissue layers including fat and other organs.

An additional set of transplantations were performed using murine carcinoma-derived organoid lines (**Figure 11D**). In this setting, progression would not be expected due to the late stage of the disease already in the endogenous mice. To test this hypothesis, we implanted 4 mice with $Apc^{fle1-15}$, 34 with $Kras^{G12D}$, 23 with $Braf^{V637E}$ and 4 mice with $Pik3ca^{H1047R}$ carcinoma organoid lines. Two different $Apc^{fle1-15}$, 4 $Kras^{G12D}$, 4 $Braf^{V637E}$ and 1 $Pik3ca^{H1047}$ carcinoma-derived organoid lines were used for this experiment. Although only half of the mice (n=2) developed tumours as a result of $Apc^{fle1-15}$ carcinoma organoids injections, none of them were graded as carcinomas, implying perhaps that the implanted line was not a *bonafide* carcinoma and instead contained more adenoma parts. In contrast, in mice injected with serrated carcinoma lines, tumour penetrance was 100% for both $Kras^{G12D}$ and $Braf^{V637E}$ lines and 50% for $Pik3ca^{H1047R}$ from which almost all were carcinomas (31/34 $Kras^{G12D}$, 22/23 $Braf^{V637E}$, 1/2 $Pik3ca^{H1047R}$). A few mice died prematurely in the cage probably due to fast tumour development and/or other complications. Due to extensive autolysis of the tissue it was not possible to histologically assess the grade of the lesions in these cases. This explains why the carcinoma penetrance in transplanted mice was lower than the number of injected mice. Furthermore, metastasis dissemination was seen in 32%, 9% and 100% of carcinoma-developing mice transplanted with $Kras^{G12D}$, $Braf^{V637E}$, $Pik3ca^{H1047R}$ lines respectively. Histologically, carcinoma lesions from implanted mice with carcinoma-derived organoids resemble the tumour from the endogenous animal (**Figure**

11D). The survival of mice implanted with tumour lines that developed tumours can be seen in **Figure 11E.** Mice implanted with *Apc*$^{fle1-15}$ adenoma-derived lines show on average a longer tumour latency period (39 weeks compared to 13 weeks for *Kras*^{G12D}, 16 weeks for *Braf*^{V637E} and 20 weeks for *Pik3ca*^{H1047R} lines). However, for each genotype a certain level of heterogeneity can be seen which reflects intrinsic characteristics and tumorigenic features of different lines within the subcohort. The same trend regarding mouse survival for all genotypes was observed for mice implanted with carcinoma lines. Given that carcinomas are more aggressive lesions than adenomas, we would expect that mice injected with carcinoma-derived organoids would develop tumours faster and show lower survival times than if injected with adenoma-derived organoids. In the *Braf*^{V637E} and *Kras*^{G12D} cohorts mouse survival correlates very well to the grading of the implanted line (adenoma or carcinoma). In the *Apc*$^{fle1-15}$ setting it is possible that, as mentioned above, no carcinoma organoids were harvested upon macroscopic dissection of the tumours from the endogenous models. Here the reported trend seen in the serrated cohorts (*Braf* and *Kras*) could not be confirmed and this result is in line with what is shown in **Figure 11D**. Mice injected with *Apc*$^{fle1-15}$ carcinoma derived organoid lines survived longer compared to mice injected with adenoma organoids (57 vs. 39 weeks, Mann-Whitney U test, p=0.143). In the *Kras*^{G12D} cohort, the survival of mice implanted with carcinoma lines was significantly lower compared to mice implanted with adenoma lines (5 vs 13 weeks, Mann-Whitney U test, p<0.0001). Similar results were observed for the *Braf*^{V637E} cohort (5 vs 16 weeks, Mann-Whitney U test, p<0.0001). For the *Pik3ca*^{H1047R} cohort results are inconclusive due to the low number of mice injected making the comparison between mice implanted with adenoma and carcinoma organoids inconclusive.

Figure 11 | Orthotopic organoid transplantation model for the study of tumour progression. [A] Organoids were derived from endogenous mouse lesions and orthotopically injected in the submucosa of recipient mice. Upon tumour formation, half of the tissue was used for histological grading of the lesion and the other half for organoids isolation. Downstream use of implanted lines involved genetic (WES, *Kras*, *Braf* ratios) and transcriptomic (3´pA RNA-seq) analysis **[B]** Tumour penetrance upon transplantation of adenoma-derived organoids of each genotype **(***Vil-Cre*; *Apc*$^{fle1-15}$, *Vil-Cre; Kras*^{G12D}, *Vil-Cre; Braf*^{V637E} and *Vil-Cre; Pik3ca*^{H1047R}). The bottom panel shows representative pictures of the necropsy of an implanted mouse. **[C]** H&E stainings showing histological progression of endogenous adenomas into carcinomas after orthotopic transplantation of organoids for all 4 genotypes included in this study. Metastasis formation in liver and lung of several animals was also confirmed (scale bars, black 300 µm, red 600 µm). **[D]** Tumour penetrance upon transplantation of carcinoma-derived organoids of each genotype **(***Vil-Cre*; *Apc*$^{fle1-15}$, *Vil-Cre; Kras*^{G12D}, *Vil-Cre; Braf*^{V637E} and *Vil-Cre; Pik3ca*^{H1047R}). Bottom panel shows representative H&E stainings of an endogenous carcinoma lesion and the respective carcinoma formed as a result of organoids transplantation (no progression) (scale bar 300 µm). **[E]** Survival of implanted mice with either adenoma or carcinoma-derived organoid lines (*Vil-Cre*; *Apc*$^{fle1-15}$ p=0.143 Mann-Whitney U test, *Vil-Cre; Kras*^{G12D} p<0,0001 Mann-Whitney U test, *Vil-Cre; Braf*^{V637E} p<0,0001 Mann Whitney test). All experiments using orthotopic injections were done together with Dr. Markus Tschurtschenthaler from the group of Prof. Dieter Saur.


A
Organoids isolation from adenoma lesion from endogenous mouse
Organoids injections
Carcinoma formation
Organoids isolation from carcinoma lesion
Downstream analysis
WES
Amplicon sequencing
3´pA RNA-seq
Histology assessment of lesions
B
Tumour penetrance (Adenoma implantations)
Number of mice (n)
Injected
Tumour forming
Carcinoma
Metastases
V-Cre; Apcfle1-15
V-Cre; KrasG12D
V-Cre; BrafV637E
V-Cre; Pik3caH1047R
Mouse necropsy after implantation
Macroscopy
Colon
T4
Liver Met
C
Endogenous mouse
Implanted mouse
Adenoma
Adenocarcinoma
Liver/lung Metasta-
D
Tumour penetrance (Carcinoma implantations)
Number of mice (n)
Injected
Tumour formed
Carcinoma
Metastases
Endogenous mouse
Implanted mouse
Adenocarcinoma
Adenocarcinoma
E
Tumour-bearing mice after implantations
V-Cre; Apcfle1-15
V-Cre; KrasG12D
V-Cre; BrafV637E
V-Cre; Pik3caH1047R
Time to tumour formation (weeks)
ns
****
Adenoma
Carcinoma
Histology of implanted line

4.6.2 Genetic progression

In addition to the pre-existing biobank of mouse endogenous organoid lines, 46 lines derived from carcinoma lesions from implanted mice, hence called implantation-derived organoids (IDOs) were generated, comprising the 3 serrated genotypes included in this study (**Figure 12A**). After organoids establishment from implanted tumour lesions, all cultures were confirmed to be free of possible contamination from tissue of the recipient animal as exemplified in the gel picture of **Figure 12B**. As previously stated, proper assessment of the tumour lesion upon mouse necropsy is crucial to ensure a pure tumour organoid culture where no normal tissue cells grow. In this case, since the genotype of implanted cells and recipient mice differ, it is possible to assess the purity of IDO cultures. By PCR amplification of the Rag2/*Il2rg* locus using primers specific for these loci we can identify both wt and mutated alleles. Pure tumour organoids derived from implanted NSG mice should therefore only show amplification of the wt allele. Transplantations in the immunocompetent setting was always performed in recipient mice containing different alleles (e.g. reporter) allowing the genetic distinction between the implanted and host cells. Since we were interested in the genetic events that distinguish early from advanced lesions, we proceeded with the genetic characterization of endogenous (adenoma) and IDO (carcinoma) lines. For that, we performed WES in 2 $Kras^{G12D}$, 1 $Braf^{V637E}$ and 1 $Pik3ca^{H1047R}$ adenoma-derived lines that were orthotopically injected into recipient mice and in 4 $Kras^{G12D}$, 4 $Braf^{V637E}$ and 2 $Pik3ca^{H1047R}$ organoid lines from the resulting carcinomas (IDOs) (**Figure 12C**). Overall, all IDOs with the exception of “NSG480 T4-2” exhibited higher mutation load in comparison to the respective injected endogenous line. Additionally, although a significant number of mutations was shared between the pairs, there was an increase in the number of unique mutations in IDO lines (**Figure 12D**).

The endogenous adenoma line “TM5805 T1” was injected into 3 different mice and organoids were derived from the resulting carcinoma lesions (“MT388 T4”, “MT345 T4-2” and “NSG425 T4”) (**Figure 12D** $Kras^{G12D}$). An increased number of mutations was detected in the IDOs from these mice (n=73, n=67, n=68 respectively) compared to the endogenous line (n=57) and, although most of the mutations are shared between these sample pairs, 24-34% of SNVs were uniquely found in IDOs. Similar results were observed for the *Braf*-mutant cohort (**Figure 12D** $Braf^{V637E}$). Line “TM2118 T4” was implanted into 3 different mice and organoids derived from the resulting carcinoma and metastasis lesion (“NSG533 T4”, “NSG532 T4-1”, “NSG532 LivMet”, and “MT1473 T4”). In this setting the number of mutations in the IDO cohort was also higher (n=28, n=28, n=30, n=33 for IDOs respectively compared to n=18 in the endogenous line).


A
Implantation-derived organoid (IDO) lines
2
20
24
Kras G12D
Braf V637E
Pik3ca H1047R
Total=46
B
Ladder
IDO 1
IDO 2
IDO 3
Il2rg -/- (NSG)
600
500
400
300
200
100
Mutant: 349 bp
WT: 269 bp
C
Endogenous mouse
Implanted mouse
Adenoma
WES
Carcinoma
V-Cre; Kras G12D (n=2)
V-Cre; Braf V637E (n=1)
V-Cre; Pik3ca H1047R (n=1)
V-Cre; Kras G12D (n=4)
V-Cre; Braf V637E (n=4)
V-Cre; Pik3ca H1047R (n=2)
D
Endogenous vs IDO lines
Kras G12D
Number of SNVs
TM5805 T1
MT388 T4
TM5805 T1
MT345 T4-2
TM5805 T1
NSG425 T4
TM7515 T2-1
RagII4 T4
Shared
Unique
Endogenous vs IDO lines
Braf V637E
TM2118 T4
NSG533 T4
TM2118 T4
NSG532 T4-1
TM2118 T4
NSG532 Liver Met
TM2118 T4
MT1473 T4
Endogenous vs IDO lines
Pik3ca H1047R
MT98 T1-3
NSG480 T4-2
MT98 T1-3
NSG480 Lung Met
E
Kras G12D
VAF
TM5805 T1
NSG425 T4
(6.6 weeks)
TM5805 T1
MT345 T4-2
(16 weeks)
TM5805 T1
MT388 T4
(18.4 weeks)
TM7515 T2-1
RagII4 T4
(2.9 weeks)
Endogenous
IDO
Braf V637E
TM2118 T4
NSG533 T4
(5.7 weeks)
TM2118 T4
MT1473 T4
(15.9 weeks)
TM2118 T4
NSG532 T4-1
(12 weeks)
TM2118 T4
NSG532 LivMet
(12 weeks)
Pik3ca H1047R
MT98 T1-3
NSG480 T4-2
(20.3 weeks)
MT98 T1-3
NSG480 LungMet
(20.3 weeks)

Figure 12 | Genetic progression after orthotopic implantation of adenoma-derived organoids. [A] Overview of the murine organoid lines generated after orthotopic transplantation (IDOs) of either adenoma or carcinoma-derived organoids. **[B]** Assessment of culture purity after IDO isolation. By performing a PCR on the *Il2rg* locus of IDOs it is possible to assess the presence of DNA from the host NSG mouse. Mutant band (349 bp) should be amplified only in specimens from NSG recipient mice, and should therefore be absent in pure IDO cultures. **[C]** Endogenous and respective IDO lines included in the WES analysis from the different serrated genotypes. **[D]** Number of mutations (SNVs) per sample for each of the endogenous-IDO pair. For each pair the number of mutations present in both samples (shared) and present exclusively in one of the samples (unique) is shown. **[E]** Mutational dynamics of the shared mutations in each endogenous-IDO sample pair in serrated genotypes. For each sample pair, the variant allele frequency (VAF) of each shared mutation as well as the copy number variation (CNV) profile is depicted. Organoid samples used for implantation were either adenoma ("TM5805 T1", "TM2118 T4", "MT98 T1-3") or carcinoma-derived ("TM7515 T1-2") (*p≤ 0.05, **p≤ 0.01, ns p> 0.05, two-tailed *t* test). WES and lcWGS raw data were analysed by Sebastian Lange and Niklas Krätzig.

Furthermore, the percentage of unique mutations in implanted lines was higher in comparison to the *Kras*-mutant cohort, comprising 71-76% of all mutations in these samples. Of note, samples "NSG532 T4-1" and "NSG532 LivMet" were derived from a primary tumour and a liver metastasis respectively from the same implanted mouse. However, no significant differences are seen between the number of mutations when compared to each other. Furthermore, shared mutations between pair "TM2118 T4 - NSG532 T4-1" are exactly the same as the ones shared between "TM2118 T4 - NSG532 LivMet" suggesting that no genetic evolutionary events occur between primary and distant lesions. Lines "NSG480 T4-2" and "NSG480 Lung Met" were derived from a primary lesion and a lung metastasis resulting from implantation of the adenoma line "MT98 T1-3" in the same mouse (**Figure 12D** *Pik3ca*^{H1047R}). In contrast to all the other sample pairs, "NSG480 T4-2" had less mutations compared to the endogenous line. Also here, the number of shared mutations between both IDO and endogenous lines is significantly lower (n=4). Conversely, 92% and 97% of mutations in IDO samples from primary tumour and metastasis, respectively, are unique.

Interestingly, even though the same endogenous organoid line was injected in different animals, the route of genetic tumour progression seems to be very conserved. In order to understand if genetic progression would occur in advanced lesions we also analysed a sample pair from an endogenous carcinoma line and the respective IDO (pair "TM7515 T2-1 - RagIL4") (**Figure 12D** *Kras*^{G12D}). The number of mutations in the endogenous line (n=181) was higher than for the endogenous adenoma lines for all cohorts regardless of the genotype. Both the endogenous and IDO lines shared 141 mutations in total, however the latter acquired in addition 128 mutations (48% total mutations) that were not present in the endogenous line. The high rate of acquisition of additional mutations in the IDO line can be explained by the high levels of aneuploidy and by the homozygous mutation in *Trp53* already present in the endogenous line that together contribute to genomic instability.

Combined, these results are in line to what we observed when carcinomas were formed upon orthotopic transplantation of *Apc*$^{\Delta Ex16}$ organoids (**Figure 10C**) and suggest that although some mutations are necessary for tumour development (driver) and are therefore shared between parental and implanted lines, during disease progression (adenoma → carcinoma) there is accumulation of additional mutations that most likely reflect subclonal populations within the tumour with increased fitness.

Strikingly, for all the adenoma-carcinoma organoid pairs there is an increase in the median variant allele frequency (VAF) of all shared mutations between the two settings (**Figure 12E**). Looking at every single mutation individually we can conclude that the VAF of a given mutation in the IDO line is either the same or is increased in comparison to the endogenous adenoma organoid line. While in some cases this increase is not so pronounced (**Figure 12E** *Kras*^{G12D}) in other instances there was a stark increase in the VAF of all shared mutations between the adenoma and carcinoma organoid lines (**Figure 12E** *Pik3ca*^{H1047R}). In contrast, comparison of IDO and endogenous carcinoma lines show that mutation frequency is similar in both samples. Additionally, LOH at chromosome 6 (*Kras* locus) and at chromosome 11 (*Trp53* locus) was detected in both samples of the pair (not shown) and is selecting for the wt alleles in both loci which explains the high VAF of *Trp53* mutation and the *Kras*^{iGD} state (**Figure 12E** *Kras*^{G12D}). In addition to increased mutation frequency, implanted lines acquire deletions and amplifications at the chromosome level that were not present in the endogenous line. Above every chart in **Figure 12E** the copy number variation (CNV) profile for each line is depicted. All IDO lines show a different CNV profile compared to the endogenous line with the exception of the carcinoma pair. In this case, both profiles are very similar which argues to the fact that tumour progression did not occur. While some IDO lines show single additional alterations like amplification of cr6 in lines “NSG425 T4” and “MT388 T4” and loss of heterozygosity (LOH) in chromosome 6 of line “MT345 T4-2” (not shown), others such as “NSG480 T4-2” present a complex CNV profile that differs greatly from the one from the endogenous line. Additionally, several samples (“NSG533 T4”, “NSG532 T4-1”, “NSG532 LivMet”, “NSG480T4-2” and “NSG480 LungMet”) specifically select for *Trp53* mutant allele and loose the wt allele due to a LOH event in chromosome 11, during progression. The acquisition of mutations and chromosomal aberrations might reflect not only the genotype of the injected line but also the longer survival time of the implanted mouse. Longer mouse survival time gives the tumour more time to progress and to acquire additional genetic events. Because we previously reported that both the *Kras* and *Braf ratios* were increased in later stages of CRC progression we performed the same analysis in the carcinoma IDOs (**Figure 13**). The analysis was done in 3 *Kras*^{G12D} heterozygous adenoma lines used for injections and in 9 IDOs. Strikingly, the vast majority of IDOs resulting from adenoma implantations (8/9)

resulted in *Kras*^{G12D} iGD in comparison to the endogenous line (**Figure 13A**). Of note, one of the IDOs (implanted with “Adenoma 2”) unexpectedly showed a *Kras ratio* below 0.5 and this event was called decreased gene dosage (dGD). Carcinoma lines (iGD) were also implanted which revealed that the *Kras ratios* were not affected in 6/7 of IDOs and decreased in 1 IDO (dGD). CNV analysis of IDO lines showed that the genetic events leading to altered *Kras ratios* were mainly chromosome 6 arm gain level gains (75% of samples) and LOH on the *Kras* locus (18.75% of samples) (**Figure 13B**). Similar results were observed for IDO lines resulting from *Braf*^{V637E} adenoma and carcinoma organoids implantations (**Figure 13C).** *Braf ratios* were assessed in 2 *Braf*^{V637E} heterozygous adenoma lines and 7 respective IDOs. Here, 2/7 IDOs showed altered ratios (1 iGD and 1 dGD) while the remaining samples showed no alteration. Concomitantly, one heterozygous *Braf*^{V637E} and one *Braf*^{iGD} carcinoma lines were also implanted into 1 and 8 mice respectively (**Figure 13C**). Implantation with the heterozygous carcinoma line resulted in an iGD-expressing IDO. At the histological level the endogenous and implanted tumour were classified with the same grade which suggests that the progression of this tumour was only detected on the genetic level. No alterations in gene dosages were observed in IDOs resulting from implantation of the iGD carcinoma line. Overall, in 31.3% of IDOs there was no alteration of the *Braf* mutant allele (Het) whereas the remaining samples (68.7%) showed increased gene dosage by chromosome 6 arm-gain levels (7/16 samples) and 1/16 samples showed dGD (**Figure 13D**). Combined, these results confirm that both *Kras*^{G12D} and *Braf*^{V637E} alleles are amplified by distinct mechanisms during the progression of early to advanced CRC. **Figure 13E** includes all samples from both cohorts (*Kras*^{G12D} and *Braf*^{V637E}) including IDOs and gives a more comprehensive and updated view of **Figure 9C** and **E**. Interestingly, the *Kras* state (Het or iGD) can be predicted by a PCR reaction as well as the kind of molecular event leading to iGD as shown in **Figure 13F**. Gain of one copy of the mutant allele is consistent with a trisomy of chromosome 6 (*Kras ratios*=0.65) whereas LOH selecting for the wt allele shows absence of the wt band after PCR and increased ratios (0.99). In all samples of the study both qualitative and quantitative approaches were used for all samples.

These results are consistent with the data derived from the histologic assessment of implanted lesions together with the genetic analysis of endogenous and IDOs in which progression is observed upon adenoma-derived organoids implantations. We additionally show that, albeit to a lesser extent, progression can also occur upon carcinoma-derived organoids implantations. The use of a colon orthotopic mouse model using adenoma-derived organoids proved to be suitable for the study the progression of early to advanced CRC. Moreover, using this approach we increased the number of carcinoma samples in our

cohort thus overcoming one of the greatest caveats of single-driver gene endogenous cancer mouse models.

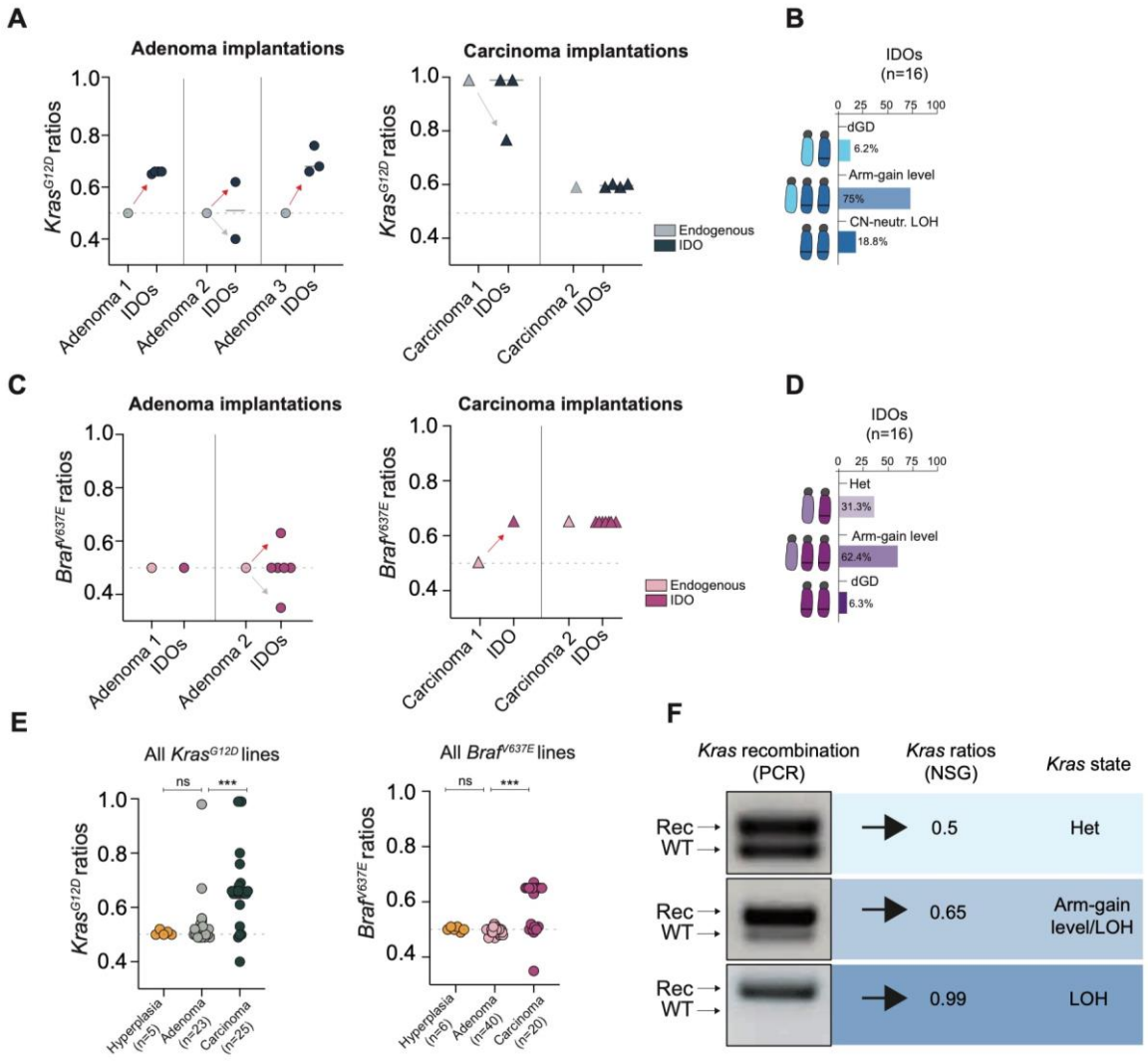


Figure 13 | Oncogenic amplification of *Kras*^{G12D} and *Braf*^{V637E} after orthotopic transplantations of adenoma-derived organoids. [A] Amplicon deep sequencing of the *Kras* locus was used for the assessment of the *Kras*^{G12D} ratios in implantation-derived organoid (IDOs) and endogenous adenoma (circle) or carcinoma (triangle) lines. **[B]** Genetic mechanisms leading to alterations in *Kras* gene dosage were analysed on the basis of lcWGS and WES. Whole chromosome amplification of chromosome 6 (trisomy) involving the *Kras*^{G12D} locus is the most common mechanism of oncogene amplification followed by copy number neutral LOH. **[C]** The same analysis was performed in the same experimental setting for *Braf*^{V637E}-mutant samples. **[D]** Mechanisms of oncogenic amplification in IDOs are driven exclusively by arm-gain level amplifications of chromosome 6. **[E]** Overview of the oncogenic state of *Kras*^{G12D} (left) and *Braf*^{V637E} (right) alleles in all samples from the cohort (endogenous + IDOs). This figure offers an updated version of the dot plots from **Figure 9C** and **E. [F]** Association of qualitative (PCR) and quantitative (NGS) methods for the quantification of the *Kras*^{G12D} ratios in relation to the oncogenic state of the sample (Het or iGD). When both bands are present in the agarose gel with the same intensity after PCR (qualitative assessment) we can predict the ratios not be altered as shown by sequencing of the *Kras* locus (*Kras* ratios=0.5) (quantitative assessment) implying a *Kras*^{G12D} heterozygous state. A faint wt and strong recombined band on the agarose gel is an indicator of chromosome 6 amplification consistent with a trisomy as assessed by NGS (*Kras* ratios=0.65). Absence of wt and detection of a recombined band on the agarose gel hints to a CN-neutral LOH event occurring in that sample consistent with no wt allele detection after NGS (*Kras* ratios=0.99).

4.7 Transcriptomic analysis of serrated organoids reveals similar tumour initiating signatures but require different stimuli for progression

4.7.1 *Kras*^{G12D} gene dosage amplification activates tumour intrinsic immune signalling programs.

In order to dissect the differences at the transcriptional level between lesions and to identify pathways that are activated during the progression of serrated tumours, we used a 3-prime end RNA sequencing technology (3´pA RNA-seq) (159, 160) in a subset of the established organoid cohort comprising the 3 serrated genotypes included in this study. Eleven *Kras*^{G12D} , 11 *Braf*^{V637E} and 4 *Pik3ca*^{H1047R} organoid lines derived from hyperplastic tissue and 11 lines derived from normal intestine of wt mice were used as control. In addition, 30 *Kras*^{G12D} tumour lines (20 adenomas, 10 carcinomas), 41 *Braf*^{V637E} (33 adenomas, 8 carcinomas) and 22 *Pik3ca*^{H1047R} (8 adenomas, 14 carcinomas) were included in the analysis (**Figure 14**).

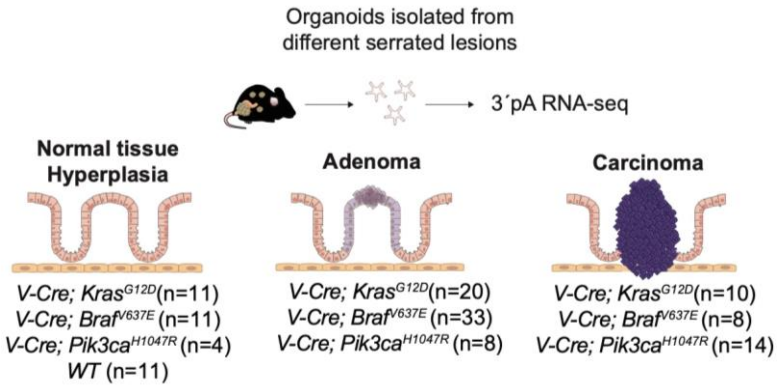


Figure 14 | Overview of the organoid samples included in the transcriptomic analysis. From each genotype, a snapshot of each stage of CRC progression was sampled by including organoid lines derived from hyperplasia, adenoma and carcinoma. Organoids derived from the intestine of pure wt mice were used as control.

For the analysis of the *Kras*^{G12D} cohort, principal component analysis (PCA) was performed in the samples from all three tissue origins including the wt samples (**Figure 15A**) or not (**Figure 15B**). Of note, 8 out of the 10 carcinoma lines included in the analysis and 2 out of 20 adenoma lines showed iGD. A clear separate cluster comprising only wt organoids can be seen when these samples are included which suggests that the oncogenic activation of *Kras*^{G12D} is driving this separation (PC1 32%) (**Figure 15A**). A second cluster containing only *Kras*^{G12D} samples can be further divided into 3 additional subclusters according to tissue origin (hyperplasia in orange, adenoma in dark blue, carcinoma light blue). Interestingly, the clusters are not well defined and resemble instead a continuum that reflects the progression of these lesions (hyperplasia → adenoma → carcinoma) which becomes clearer when no wt samples are included in the analysis (PC1 30%) (**Figure 15B**).

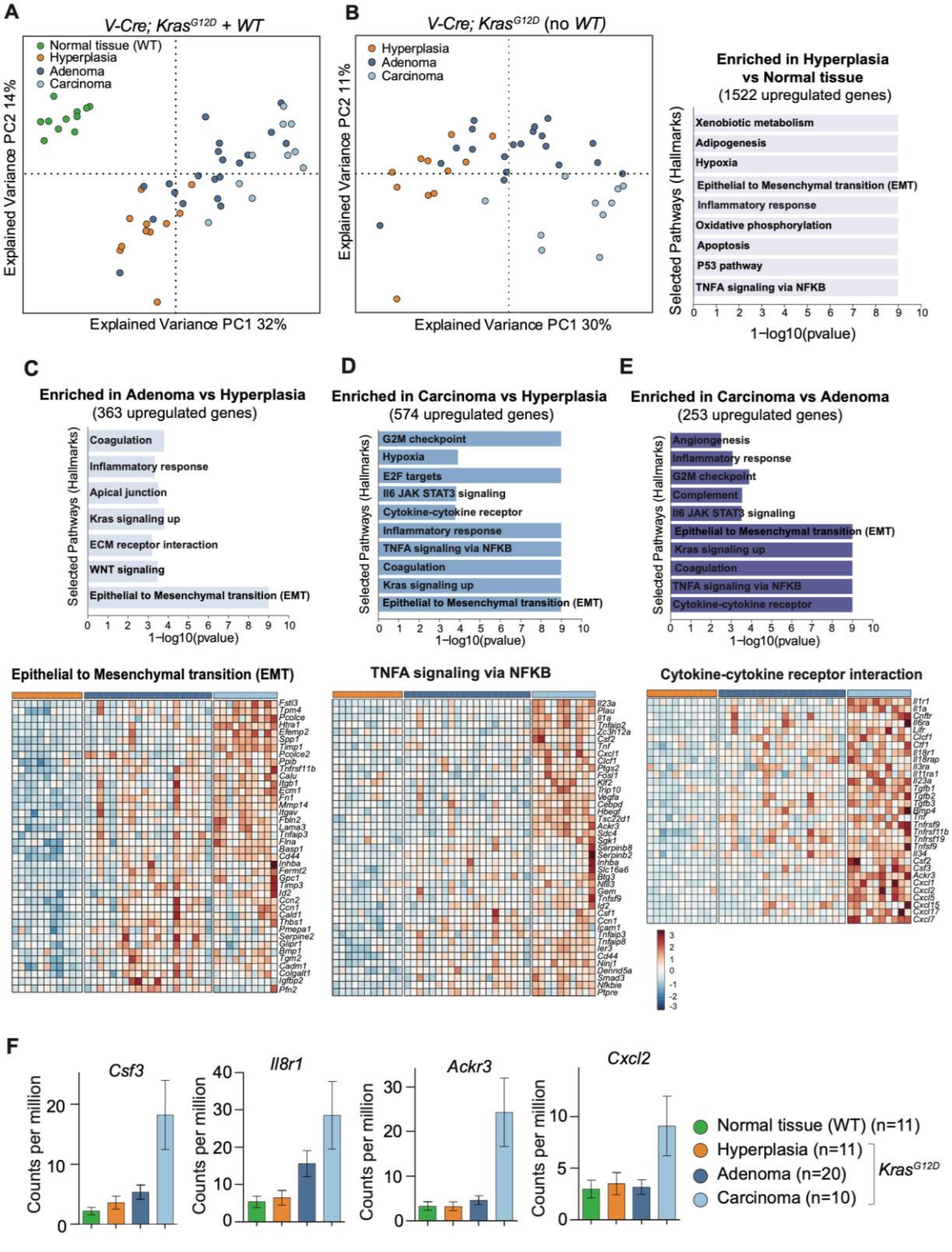
Within each histological grade (hyperplasia, adenoma, carcinoma) the fact that some samples overlap with samples from other groups is a direct evidence of the heterogeneity within each tissue group from where the organoids were generated. A few adenoma samples contain more advanced histological features, which cluster closer to the carcinoma samples. Conversely, low-grade adenomas were closer to hyperplasia samples in the PCA (**Figure 15B**).

We further identified the genes that were differentially expressed between the different groups of samples and that were responsible for the separation of the clusters in the PCA (adjusted P value ≤0.01, fold change ≥0.5). The lists of differentially expressed genes used in this work were subjected to pathway analysis using the "Molecular Signature Database" (MsigDB, H collection: Hallmark gene set) (3). Significantly enriched gene sets were evaluated for pathways that could be condensed into a few predominating molecular categories/processes. First, we compared the wt normal organoids with hyperplasia-derived the *Kras*^{G12D} organoids. In this setting we can assess the effect of the *Kras*^{G12D} mutation in non-transformed intestine and identify the pathways that are regulated in response to oncogenic activation. A total of 1522 genes were differently expressed between wt normal tissue and *Kras*^{G12D} hyperplasia organoids (**Figure 15B,** histogram). Pathway analysis using this *geneset* retrieved the GO terms "p53 pathway", "apoptosis" and "TNF-alpha signalling via NF-κB" to be enriched in mutant compared to the wt organoids. *Kras*^{G12D} oncogenic activation has been linked to OIS (181), and the cooperation with p53 seems to stabilize this genetic program (182). Furthermore, induction of p53 in response to oncogene activation induces apoptosis (183) which is also one of the main activated pathways in *Kras* mutant hyperplasia organoids. TNF-alpha is an inflammatory cytokine that kills tumour cells (184). These results suggest that *Kras*^{G12D} activation in non-transformed tissue induces both an apoptotic response orchestrated by p53, but at the same time induces a very early pro tumorigenic program that results in TNF-alpha induction.

Further analysis focused on the progression of *Kras*^{G12D}- driven tumours, thus differential gene expression analysis was performed between the organoids derived from hyperplasia, adenoma and carcinoma samples to understand which stage-specific pathways were being regulated. Pathway analysis using the upregulated genes (n=363) in the adenoma cohort in comparison to hyperplasia samples revealed "epithelial to mesenchymal transition (EMT)" and "Wnt signalling" to be the strongest activated pathways in this comparison (**Figure 15C**). Additional adenoma-specific pathways include "ECM receptor interaction", upregulation of "*Kras* signalling" and "inflammatory response". In the heatmap from **Figure 15C** we can see specifically the genes that contribute for the activation of EMT pathway in adenomas in comparison to hyperplasia samples since this was the most activated process in the comparison. Genes associated with cell-cell or cell-

ECM interaction including integrins (*Itgb1, Itgav, Fermt2*) matrix metalloproteases (*Mmp14, Timp1. Timp3*), *Cd44*, *Ecm1* among others were seen to be upregulated in adenomas and carcinomas which suggests that these mechanisms are tumour-specific and are activated early in tumorigenesis. Five hundred and seventy-four genes were found to be upregulated in carcinomas in comparison to hyperplasia (**Figure 15D**). Pathway analysis using these genes, retrieved once again EMT as the most regulated process between the groups in line to what was shown in **Figure 15C**. In addition, several immune-related pathways are upregulated in carcinomas, including "TNF-alpha signalling via NF-κB", "inflammatory response" "cytokine-cytokine receptor interaction" and "IL6 JAK STAT3 signalling". Genes from the top regulated immune-related pathway ("TNF-alpha Signalling via NF-κB") include cytokines (*Il23a, Il1a*), chemokines (*Cxcl1, Ackr2*) among other gene families. Immune signatures became more prominent when carcinomas were compared to adenomas. A total of 253 genes were found to be upregulated in carcinomas compared to adenomas, and the pathways "Cytokine-cytokine receptor interaction" and "TNF-alpha Signalling via NF-κB" were found to be the most significant terms after pathway analysis. Additional terms from the latter include "*Kras* signalling", "EMT", "IL6 JAK STAT3 signalling" and "inflammatory response" (**Figure 15E**). Interestingly, less genes were differentially regulated in the carcinoma-adenoma comparison than in the adenoma-hyperplasia and carcinoma-hyperplasia comparisons hinting that at the transcriptional level adenomas are more similar to carcinomas than to hyperplasia. Several genes coding for interleukins, Tnf molecules, Csf, chemokines and chemokine receptors (CXC), responsible for important immunomodulatory processes were seen to be differently regulated between the different stages of tumour progression (**Figure 15E**).

Figure 15 | Differential expression analysis between different stages of *Kras^{G12D}*-driven tumourigenesis. Principal component analysis (PCA) based on the 10% most variable genes, including wt samples **[A]** or only *Kras^{G12D}*-mutant samples **[B]** that characterize the trends exhibited by the expression profiles of different stages of tumour progression. Each dot represents an organoid sample and each colour represents the tissue type from which the line was derived as shown on the right section of **B**. Components 1 and 2 explain 46% and 41% of the total variance in the dataset when wt samples are included or not, respectively. **[C], [D]**, **[E]** and right part of **[B]** Gene set enrichment analysis (GSEA) using the differentially expressed genes (adjusted P value ≤0.01, fold change ≥0.5) between normal tissue, hyperplasia, adenoma and carcinoma samples was performed with "Molecular Signature Database" (MsigDB, H collection: Hallmark gene set) (3). In the right part of **[B]** 1522 genes upregulated in *Kras^{G12D}*-driven hyperplasia in comparison to wt normal tissue were used as input for GSEA. In **[C]** 363 genes upregulated in adenoma in comparison to hyperplasia were used as input for GSEA. Below, unbiased hierarchical clustering of transcriptome profiles of *Kras^{G12D}* organoids (n=41) on the basis of an "EMT" hallmark gene set (3). In **[D]** 574 genes upregulated in carcinoma in comparison to hyperplasia were used as input for GSEA. Below, unbiased hierarchical clustering of transcriptome profiles of *Kras^{G12D}* organoids (n=41) on the basis of a "TNFA signaling via NFKB" hallmark gene set (3). In **[E]** 253 genes upregulated in carcinomas in comparison to adenomas were used as input for GSEA. Below, unbiased hierarchical clustering of transcriptome profiles of *Kras^{G12D}* organoids (n=41) on the basis of "cytokine-cytokine receptor interaction". **[F]** Expression levels (counts per million - cpm) of immune-related genes during progression of *Kras^{G12D}*-tumours. Some genes (*Csf3* and *Il8r1*) increase their expression in early stages of tumour development while others (*Ackr3* and *Cxcl2*) are overexpressed specifically in carcinomas. Analysis of transcriptome raw data and clustering of samples was performed by Thomas Engleitner.



Mounting evidence highlights the crucial role of the crosstalk of immune and cancer cells. Particularly, several molecules involved in regulation of central immune processes such as CXCR2 and its ligand IL8 together with IL1, CCL5 and GM-CSF have been

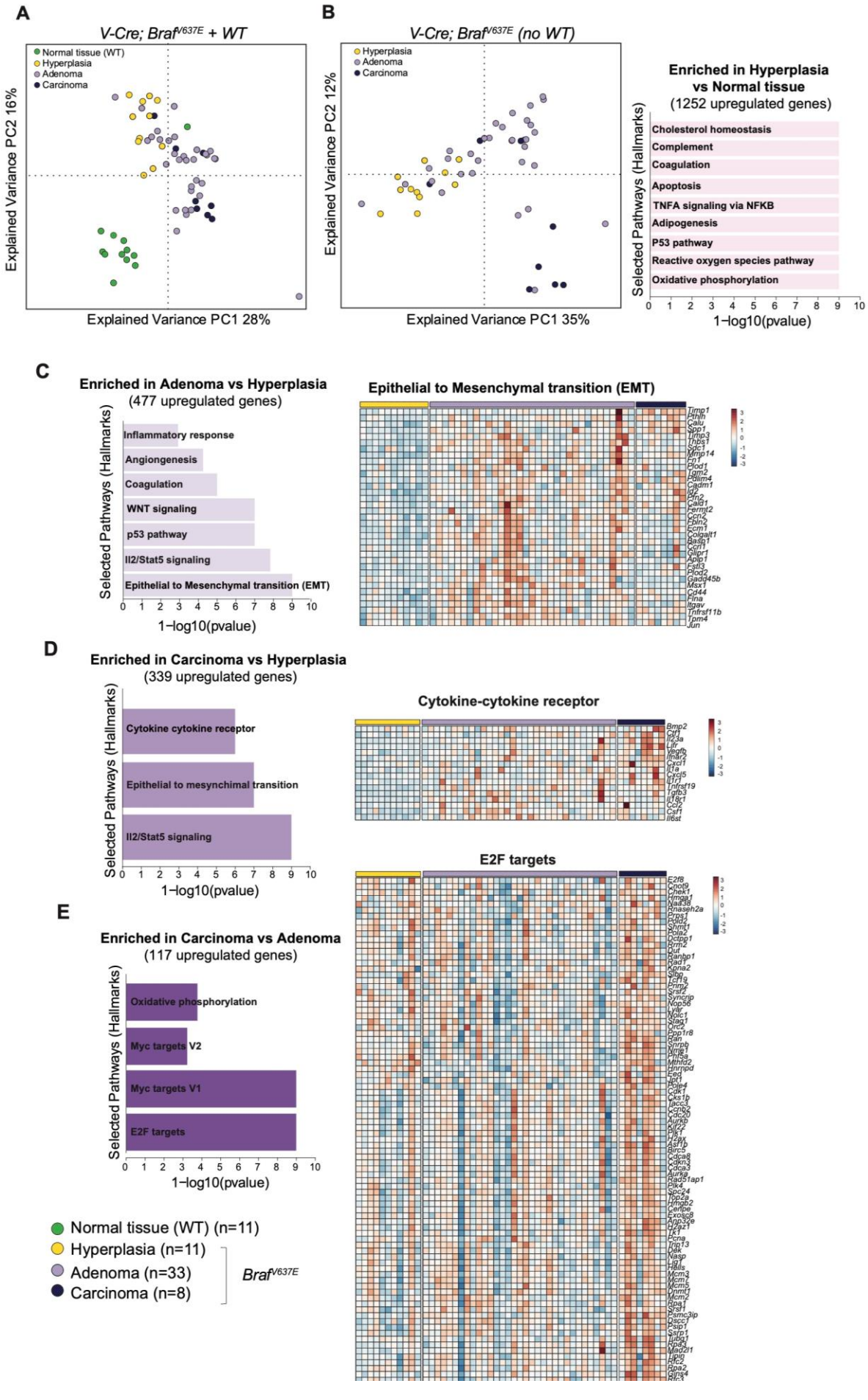
associated with lung and pancreas tumorigenesis in a *KRAS*-mutant background (185-187). Furthermore, we show that different immune-related genes are differentially regulated according to the tissue context. Specifically in the intestine, we observed the expression of interleukins such as *Il3a*, *Il11ra1* and *Il8r1*, *Tnf*, and the cytokines *Clcf1* and *Ctf1* to be gradually increased starting already at the hyperplasia-adenoma transition. In contrast, *Csf2, Csf3* and several CXC chemokines (*Cxcl1, Cxcl5, Cxcl15, Cxcl17*) were found to be induced specifically in the transition from adenoma to carcinoma (**Figure 15E**). Granulocyte/neutrophil chemokines comprise mostly the ELR+ CXC chemokines CXCL1, CXCL2, CXCL3, CXCL5, CXCL6, and CXCL7 (PPBP), as well as interleukin-8 (CXCL8). These chemokines are ligands for CXCR2 and are known powerful neutrophil chemoattractants (188). In our dataset, *Cxcl1*, *Cxcl2*, *Cxcl5*, *Cxcl7*, *Cxcl15* and *Cxcl17* are mostly upregulated in carcinomas when compared to adenomas and hyperplasia samples. In fact, by analysing the absolute expression values (counts per million - cpm) of the immune-related genes across samples we concluded that some (*Csf3* and *Il8r1*) are probably important for transformation and initial stages of disease development given their increasing expression from adenomas to carcinomas, while others such as *Ackr3* and *Cxcl2* are most likely responsible for progression of advanced CRC (**Figure 15F**). Combined these results suggest that increased gene dosage activate cell intrinsic programs that rely on the communication of cancer cells with the tumour microenvironment particularly with recruiting immune cells (e.g. neutrophils) thereby contributing to immune evasion.

4.7.2 *Braf*^{V637E} late-stage tumorigenesis is driven by *Myc* activation

To further dissect the pathways that are upregulated during progression of *Braf*^{V637E}-driven tumours we analysed the transcriptomes of 11 hyperplasia, 33 adenomas and 8 carcinoma organoid lines from this genotype. Of note, only 2/8 carcinoma samples used for this analysis are *Braf*^{iGD}. The same normal tissue samples used for the analysis of *Kras*-driven samples were also used in this setting (**Figure 16**). PCA of the top most variant genes in this cohort showed allowed the distinction between wt and *Braf*^{V637E}-mutant samples (**Figure 16A**). In a similar way as what was seen in the *Kras* experiment, also here a separation between hyperplasia, adenomas and carcinomas was achieved which becomes clearer when wt normal samples are excluded (**Figure 16B**). We did not observed separation of samples into defined clusters instead, we saw a similar pattern that reflects the order of progression from hyperplasia to adenoma and then to carcinoma, as also seen in the *Kras* mutant setting. However, the degree of overlap of samples from different histological groups is considerably higher here where a few carcinoma samples overlap with adenoma and even hyperplasia samples. We then performed differential gene

expression analysis between samples from the different histological subtypes within the cohort. First, we compared wt normal organoids with hyperplasia-derived *Braf*^{V637E} organoids. A total of 1252 genes were differently expressed between wt normal tissue and *Braf*^{V637E} hyperplasia organoids (**Figure 16B,** histogram). The top 2 pathways retrieved after GSEA were the metabolism-related "oxidative phosphorylation" and "reactive oxygen species pathway". Both pathways have been associated with tumour progression and have been suggested as potential therapeutic targets (189, 190). Other pathways include "p53 pathway" and "TNFA signalling via NFκB" as seen in the *Kras*^{G12D} experiment. A total of 477 genes were upregulated in adenomas compared to hyperplasia and pathway analysis revealed several processes like "EMT", "p53 pathway", "Wnt signalling" and "IL2/SAT5 signalling" to be regulated between these 2 groups (**Figure 16C**). Interestingly, EMT signature was also one of the most regulated pathways in the comparison between carcinoma and hyperplasia (**Figure 16D**), suggesting that this is a tumour specific mechanism that mediates tissue transformation. Here, 339 genes were found to be differentially expressed between the 2 conditions and besides "EMT" we observed immune-related pathways such as "IL2/STAT5 signalling", and "cytokine-cytokine receptor". The comparison of early stage (adenoma) to advanced disease stage (carcinoma) showed 117 genes to be upregulated in the latter. Pathway analysis using these genes retrieved upregulation of "E2F targets" and Myc signalling pathways ("Myc targets V1" and "Myc targets V2") (**Figure 16E**). Comparison of regulated pathways in the progression of *Kras*^{G12D} and *Braf*^{V637E} tumours suggests that the mechanisms driving transformation and early tumour progression do partially overlap between both serrated cohorts (EMT, Wnt signalling, immune-related pathways). However, adenoma → carcinoma transition in the latter seems to evolve distinct pathways more associated with cell cycle regulation (E2F, Myc signalling).

Figure 16 | Differential expression analysis between different stages of *Braf*^{V637E}-driven tumourigenesis. Principal component analysis (PCA) based on the 10% most variable genes, including wt samples **[A]** or only *Braf*^{V637E}-mutant samples **[B]** that characterize the trends exhibited by the expression profiles of different stages of tumour progression. Each dot represents an organoid sample and each colour represents the tissue type from which the line was derived as shown on the right section of **B**. Components 1 and 2 explain 44% and 47% of the total variance in the dataset when wt samples are included or not, respectively. **[C], [D]**, **[E]** and right part of **[B]** Gene set enrichment analysis using the differentially expressed genes (adjusted P value ≤0.01, fold change ≥0.5) between normal tissue, hyperplasia, adenoma and carcinoma samples was performed with "Molecular Signature Database" (MsigDB, H collection: Hallmark gene set) (3). In the right part of **[B]** 1252 genes upregulated in *Braf*^{V637E}-driven hyperplasia in comparison to wt normal tissue were used as input for GSEA. In **[C]** 477 genes upregulated in adenoma in comparison to hyperplasia were used as input for GSEA. Below, unbiased hierarchical clustering of transcriptome profiles of *Braf*^{V637E} organoids (n=52) on the basis of an "EMT" hallmark gene set (3). In **[D]** 339 genes upregulated in carcinoma in comparison to hyperplasia were used as input for GSEA. Below, unbiased hierarchical clustering of transcriptome profiles of *Braf*^{V637E} organoids (n=52) on the basis of a "cytokine-cytokine receptor interaction" hallmark gene set (3). In **[E]** 117 genes upregulated in carcinomas in comparison to adenomas were used as input for GSEA. Below, unbiased hierarchical clustering of transcriptome profiles of *Braf*^{V637E} organoids (n=52) on the basis of "E2F targets". Analysis of transcriptome raw data and clustering of samples was performed by Thomas Engleitner.


A
V-Cre; Braf V637E + WT
Normal tissue (WT)
Hyperplasia
Adenoma
Carcinoma
Explained Variance PC2 16%
Explained Variance PC1 28%
B
V-Cre; Braf V637E (no WT)
Hyperplasia
Adenoma
Carcinoma
Explained Variance PC2 12%
Explained Variance PC1 35%
Enriched in Hyperplasia vs Normal tissue
(1252 upregulated genes)
Cholesterol homeostasis
Complement
Coagulation
Apoptosis
TNFA signaling via NFKB
Adipogenesis
P53 pathway
Reactive oxygen species pathway
Oxidative phosphorylation
Selected Pathways (Hallmarks)
1−log10(pvalue)
C
Enriched in Adenoma vs Hyperplasia
(477 upregulated genes)
Inflammatory response
Angiongenesis
Coagulation
WNT signaling
p53 pathway
Il2/Stat5 signaling
Epithelial to Mesenchymal transition (EMT)
Selected Pathways (Hallmarks)
1−log10(pvalue)
Epithelial to Mesenchymal transition (EMT)
D
Enriched in Carcinoma vs Hyperplasia
(339 upregulated genes)
Cytokine cytokine receptor
Epithelial to mesynchimal transition
Il2/Stat5 signaling
Selected Pathways (Hallmarks)
1−log10(pvalue)
Cytokine-cytokine receptor
E
Enriched in Carcinoma vs Adenoma
(117 upregulated genes)
Oxidative phosphorylation
Myc targets V2
Myc targets V1
E2F targets
Selected Pathways (Hallmarks)
1−log10(pvalue)
E2F targets
Normal tissue (WT) (n=11)
Hyperplasia (n=11)
Adenoma (n=33)
Carcinoma (n=8)
Braf V637E

Analysis of the heatmaps for the different comparisons show a high level of heterogeneity within each histological group particularly for adenomas and carcinomas. Three out of 8 carcinomas overlap completely with the adenoma samples in the PCA plots, and a few adenoma samples cluster together with hyperplasia samples. Also in this cohort, high and low-grade adenomas were combined (adenomas). Therefore, it would be reasonable to expect that low grade adenomas share more similarities with hyperplasia samples while high-grade adenomas would be closer to carcinomas. For some cases (outliers) we cannot rule out the possibility of a sampling bias.

4.7.3 Neoplastic transformation of *Pik3ca*^{H1047} organoids is driven by inflammation

The transcriptomes of 4 hyperplasia, 8 adenoma and 14 P*ik3ca*^{H1047R}-driven carcinoma samples were sequenced by 3´pA RNA-seq. PCA revealed that, unlike what was observed for the other genotypes, there was a significant overlap between wt, hyperplasia and carcinoma organoids and no clear clusters could be seen (**Figure 17A**). In this cohort, the oncogenic activation of P*ik3ca*^{H1047R} is not driving differences between groups. Furthermore, in the case of the *Kras* and *Braf* cohorts the PCA analysis revealed the sequential clustering of samples that reflected disease progression (hyperplasia → adenoma → carcinoma). In the present cohort however, there is considerably heterogeneity in the carcinoma subgroup and although some samples are closer to adenomas, they also partially overlap with hyperplasia samples. Additionally, differences between adenomas and hyperplasia seem greater than between carcinomas and hyperplasia which is not expected (**Figure 17B**). Given the considerable overlap of carcinoma samples with hyperplasia and wt samples combined with the fact that adenoma organoids do seem to differ significantly from hyperplasia, prompt us to focus the analysis of this dataset solely on the comparison between adenoma and hyperplasia. A total of 677 genes were found to be upregulated in adenoma-derived organoids in comparison to hyperplasia (**Figure 17C**). Pathway analysis using this *geneset* retrieved “Interferon Alpha (IFNα) response” and “Interferon Gamma (IFN*γ*) response” as the top regulated pathways between both groups. IFN responses are usually regarded as enhancers of immune responses by either killing tumour cells or improving ICB therapy, however they can have opposing effects in anti-tumour responses (191). In fact, IFN can contribute to tumour immune evasion by inducing the expression of PD-L1, indoleamine 2,3-dioxygenase (IDO), and arginase in the tumour microenvironment that, all together, shield the tumour from immune cells (192). Several genes including *Irf7*, *Irf9* and members of *Ifitm* family of genes (*Ifitm1*, *Ifitm2*, *Ifitm3*) were seen to be overexpressed in adenomas in comparison to hyperplasia. Deregulation of IFN and *Ifitm* genes has been reported in cancer, however the mechanisms or reasons for this finding

are still elusive (193, 194). Strikingly all other significant pathways were also found to be present in the comparison between adenoma and hyperplasia samples from the other serrated genotypes. Of note, “EMT” was the third most regulated pathway, followed by several immune-related similar pathways including “IL6 JAK STAT3 signalling”, “inflammatory response” and “TNFA signalling via NFκB”. These observations suggest that a strong inflammatory signature is mediating, or at least greatly contributing, for the initial neoplastic transformation in *Pik3ca*^{H1047R}-mutant organoids. Even though *Pik3ca* mutations have not been associated with Wnt pathway activation and no mutations in this pathway were reported in our cohort, the top upregulated gene in adenomas (6.7 fold) was *Wnt5a*. Wnt5a is a nontransforming member of the Wnt family of secreted and cysteine-rich proteins that exert its cellular effects via autocrine or paracrine routes which shows a promiscuous functional behaviour in the development of different cancer entities (195). Although we did not assess Wnt pathway activation in these samples we consider a possibility that Wnt might be activated in this context.

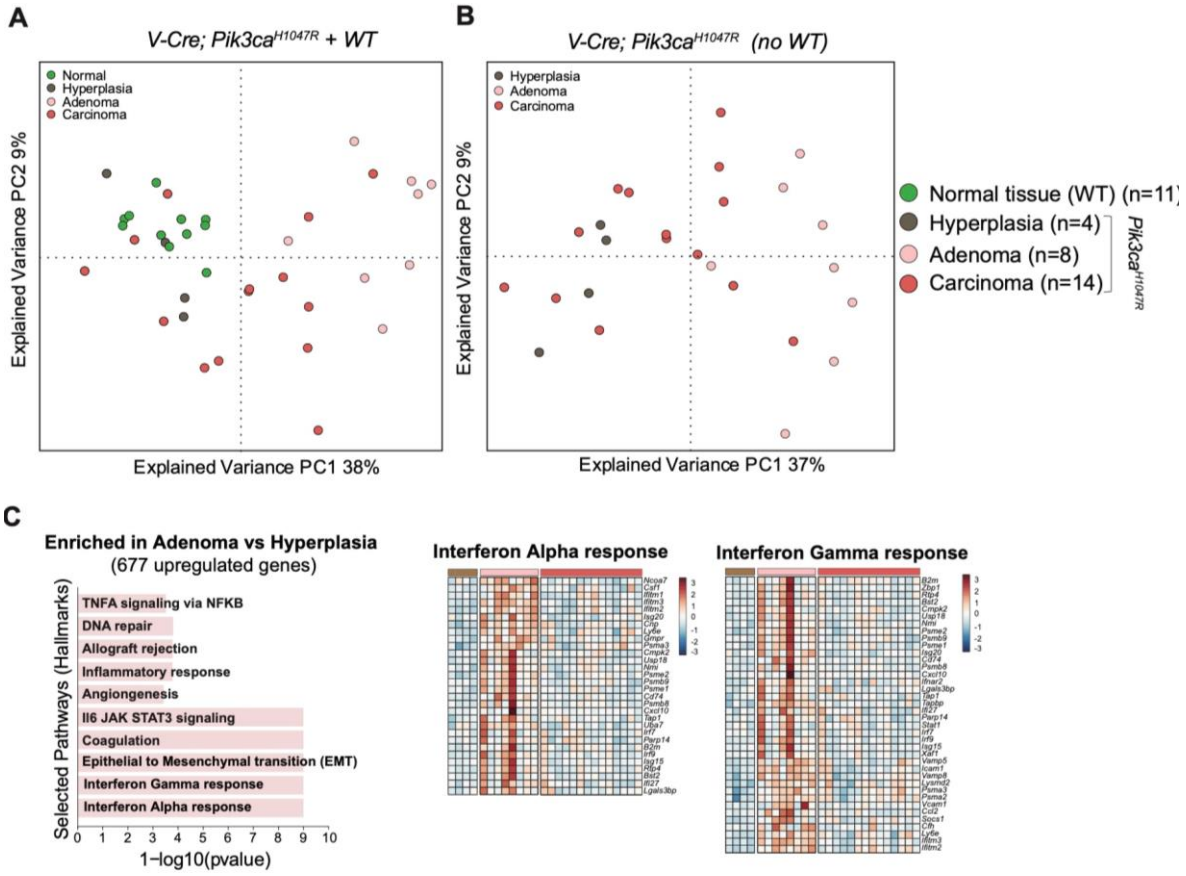


Figure 17 | Differential expression analysis between *Pik3ca^{H1047R}*-driven hyperplasia and early-stage disease. Principal component analysis (PCA) based on the 10% most variable genes, including wt samples **[A]** or only *Pik3ca^{H1047R}* -mutant samples **[B]** that characterize the trends exhibited by the expression profiles of different stages of tumour progression. Each dot represents an organoid sample and each colour represents the tissue type from which the line was derived as shown on the right section of **B**. Components 1 and 2 explain 47% and 46% of the total variance in the dataset when wt samples are included or not, respectively. **[C]** Gene set enrichment analysis using the differentially expressed genes (adjusted P value ≤0.01, fold change ≥0.5) between hyperplasia and adenoma samples was performed with "Molecular Signature Database" (MsigDB, H collection: Hallmark gene set) (3). A total of 677 genes upregulated in adenomas in comparison to hyperplasia samples were used as input for GSEA. On the right, unbiased hierarchical clustering of transcriptome profiles of *Pik3ca^{H1047R}* organoids (n=26) on the basis of an "Interferon Alpha response" and "Interferon Gamma response" hallmark gene sets (3). Analysis of transcriptome raw data and clustering of samples was performed by Thomas Engleitner.

The gene expression dynamics in disease progression in the *Pik3ca* cohort differed greatly as already suggested by the PCA (**Figure 17A** and **B**). The pathways identified to be upregulated in adenomas, where absent in the carcinoma cohort which might suggest that they are necessary for initial transformation (hyperplasia → adenoma) but not for tumour progression (adenoma → carcinoma). Due to the high level of heterogeneity within the carcinomas and the overlap with hyperplasia and wt samples, differential gene expression between carcinoma and hyperplasia samples retrieved no significant results (data not shown).

4.8 Establishment of an orthotopic PDOX mouse model to study human tumours *in vivo*

In a similar approach to what was established with murine organoids we set to explore the orthotopic CRC injections model in the human setting (**Figure 18A**). Organoids were isolated from resected human carcinomas and subsequent to *in vitro* establishment and expansion of the line, we transplanted them into immunodeficient mice (NSG). After organoids engraftment in the mouse was achieved (PDOX), tumours were formed in the colon of these mice as well as metastasis in distant organs (e.g. liver and lung). A few mice (n=3) also had ascites at the time of necropsy. Subsequently, from the implanted tumours (PDOX), organoids were derived and the culture established from the primary tumours in the mouse and, whenever possible, from the metastatic lesions as well. Nine primary CRC organoid lines were used for transplantation into at least 3 NSG mice per line. A total of 30 mice developed tumours thus reaching the criteria for "endpoint" experiment (**Figure 18B**). With the exception of mice transplanted with CRC R16 line, 1 mouse injected with the line CRC R19 and 1 mouse transplanted with line CRC R12, survival of implanted mice did not vary considerably and was rather consistent among injected lines. Thirty-two IDO cultures were established as a result of human CRC implantations where 22 were derived from primary tumour lesions and 10 from metastatic sites (1 from a metastasis in the diaphragm,

9 from metastases in the liver). Human organoids were naturally selected *in vitro* since mouse organoids cannot grow in the medium used for human cultures thereby ensuring the purity of the human culture. Interestingly, the histology of the primary resected patient tumour was recapitulated in the mice upon orthotopic transplantation of the respective organoid line, regardless of the histological type. **Figure 18C** shows the histology pictures from 4 representative patient primary tumours in the panel above, and below the histology of the tumour generated in the mouse after organoids transplantation.

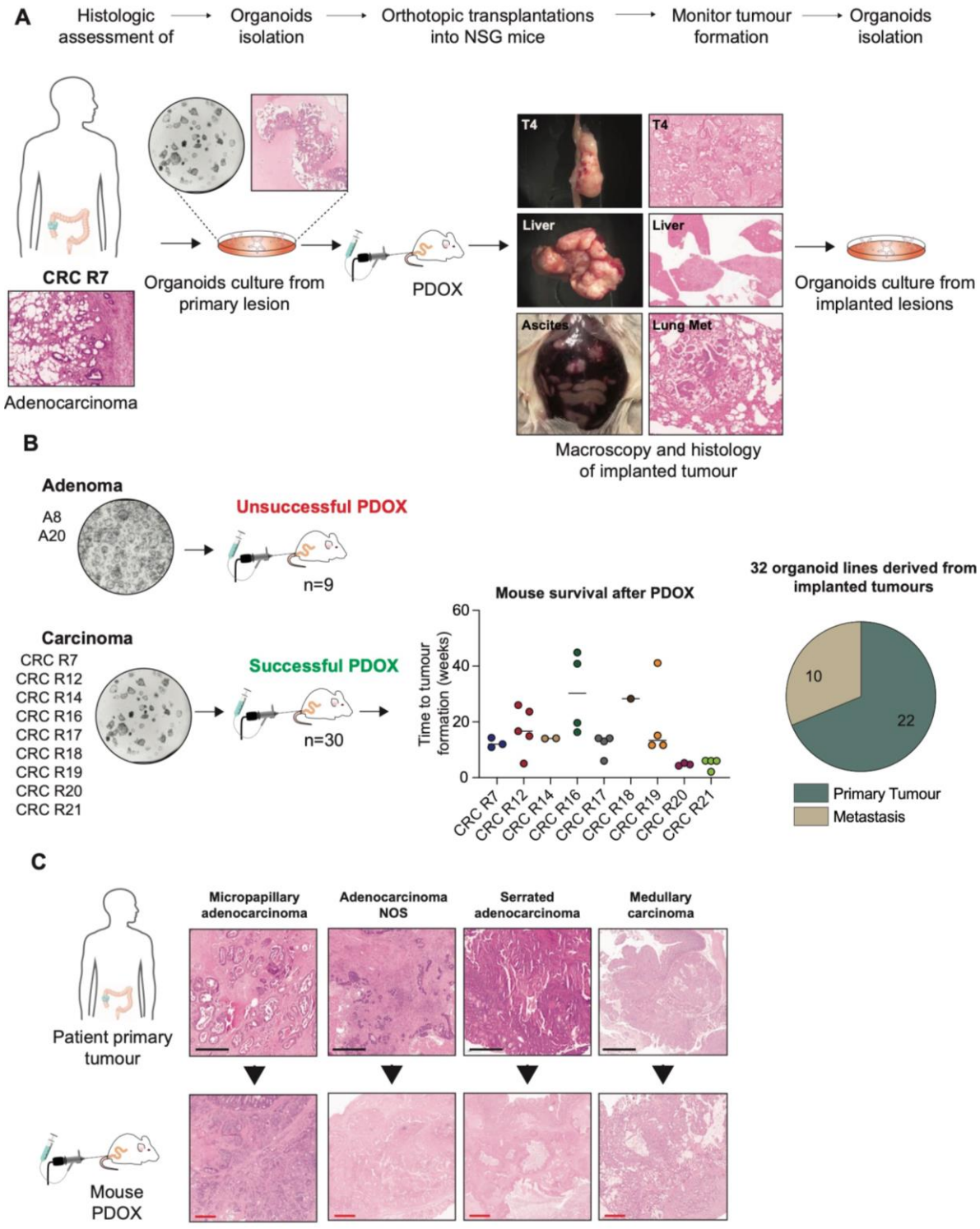


Figure 18 | Establishment of an orthotopic CRC mouse model for the *in vivo* study of human CRC. [A] Organoid lines are established from the patient tumour upon surgical resection. H&E stainings were performed for both the primary tissue and respective organoid line for histological tumour subtyping. Orthotopic injection of carcinoma-derived organoids into the colon submucosa of NSG mice recapitulates local and distant disease states found in human patients shown by the representative pictures of mouse necropsy and H&E staining of mouse lesions. Organoids from implanted lesion (PDOX) were subsequently isolated. (PDOX - patient derived organoids based xenograft; T4 - primary tumour in the colon). **[B]** Implantation with human adenoma-derived organoid lines led to unsuccessful PDOXs as cells did not engraft in the mice (n=9). Engraftment is only possible when carcinoma organoids are injected. Nine carcinoma organoid lines were injected into 30 NSG mice from which 32 lines were subsequently isolated. Survival of mice injected with carcinoma lines is shown in the graph in the right part of the figure. **[C]** H&E stainings of human primary tumours (above) and of mouse tumours after organoids transplantation (below). The histological subtype of the human primary tumour is recapitulated in the mouse tumour after injection of the respective carcinoma-derived organoid line (NOS, not otherwise specified; scale bars, black 600 µm, red 300 µm). All experiments using orthotopic injections were done together with Dr. Markus Tschurtschenthaler from the group of Prof. Dieter Saur.

The 4 different histological subtypes present in the human tumours (micropapillary, NOS, serrated and medullary) were recapitulated in the mice after transplantations. In order to assess the feasibility of this system to study tumour progression we additionally attempted to generate PDOXs using adenoma-derived organoid lines from the human cohort. Surprisingly however, adenoma lines did not engraft in all of the 9 mice tested (**Figure 18B**).

4.9 PDOX-derived organoids reflect main genetic features of the primary tumour of the patient

Figure 19 summarizes all the PDOX experimental “trees” performed including information regarding the type of lesions the mice developed and the respective tissues harvested from the implanted animals. On a histological level we confirmed that human tumours do not progress in mice after organoids transplantation and rather recapitulate key morphological features from the original tumour in the patient. In order to understand if this model can also recapitulate the genetic makeup of patient tumours we selected 4 experimental “trees” from **Figure 19** (CRC R7, CRC R12, CRC R14 and CRC R17) and performed targeted amplicon sequencing using a predefined panel of known cancer genes (see Materials and Methods section) on 3 different tissues: i) primary tumour of the patient (FFPE tissue), ii) the organoid line established from the primary tumour of the patient, and iii) organoid lines established from mouse PDOX after organoids transplantation (**Figure 20A**). For every experiment, each PDOX corresponds to a different animal that was injected with the primary tumour organoid line at the same timepoint. Overall, driver mutations found in the FFPE tissue sample from the primary tumour were also found in the respective organoid line (**Figure 20B**).

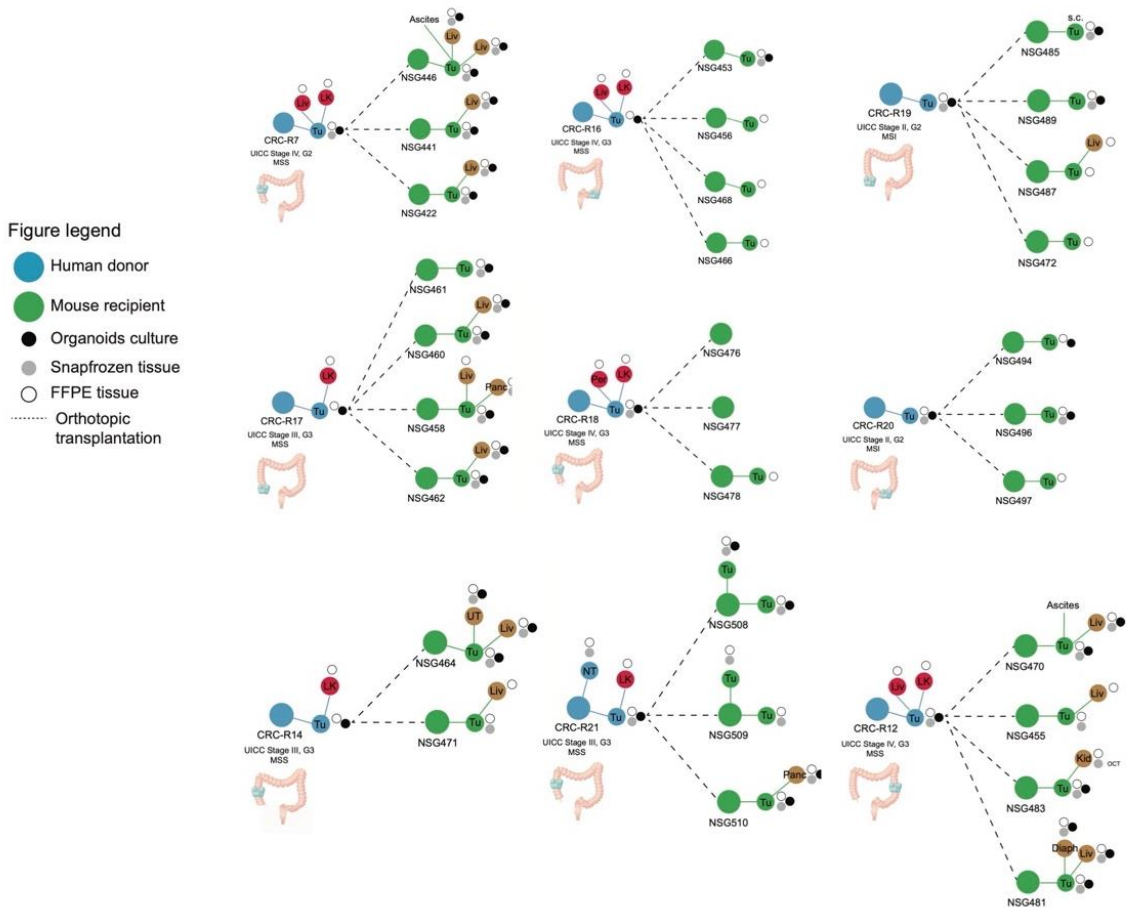


Figure 19 | Human experimental "trees" summarizing the type and number of samples generated as a result of carcinoma-derived organoids transplantations. Four NSG mice were implanted with each of the 9 organoid lines generated from human primary tumours, however a few mice had to be removed from the study due to premature death unrelated to tumour formation. When possible, from tumour-bearing mice we generated a biobank of FFPE, whole tissue pieces and organoid lines from all lesions resulting from transplantations. Mouse sample collection was done together with Dr. Markus Tschurtschenthaler from the group of Prof. Dieter Saur and Dr. Moritz Jesinghaus from the Klinikum rechts der Isar.

Two exceptions can be made for a *SMAD4* mutation in experiment CRC R14 and a *CTNNB1* mutation in experiment CRC R17 where both mutations were not detected in the paraffin tissue but were present in the organoid line. Interestingly, in this experiment, PDOX A and PDOX B had a deletion of exon 3 of *CTNNB1*. Exon 3 of *CTNNB1* encodes serine threonine phosphorylation sites for GSK-3β which in turn leads to β-catenin degradation (196). Single point mutations and deletions at these sites are frequent events in CRC and were shown to induce Wnt pathway by stabilizing β-catenin activity (197). Moreover, considering all 4 experiments, the VAF of the shared driver mutations between FFPE tissue and organoids was found to be starkly increased in the latter. In FFPE samples, not only epithelial tumour cells compose the tissue section. Several other cell types of stromal, immune and vascular origins are also sequenced when DNA is isolated from these specimens, thus reducing tumour purity and contributing with an increased number of wt

fraction. On the other hand, organoids are derived solely from tumour epithelial cells and therefore constitute a much “purer” setting for the assessment of somatic mutations. This becomes clear when the comparison between samples from primary tumour (FFPE and organoids) is made in all experiments (**Figure 20B**). Increases in VAF of certain mutations can be subtle (e.g. CRC R12 experiment, *APC* mutation: 25% in FFPE to 44% in organoids) or very marked (e.g. CRC R7 experiment, *PIK3CA (A)* mutation: 33% in FFPE to 100% in organoids). Additionally, almost all mutations detected in the organoid lines from the primary patient tumour were also present in the animals after orthotopic transplantation of the line (PDOX).

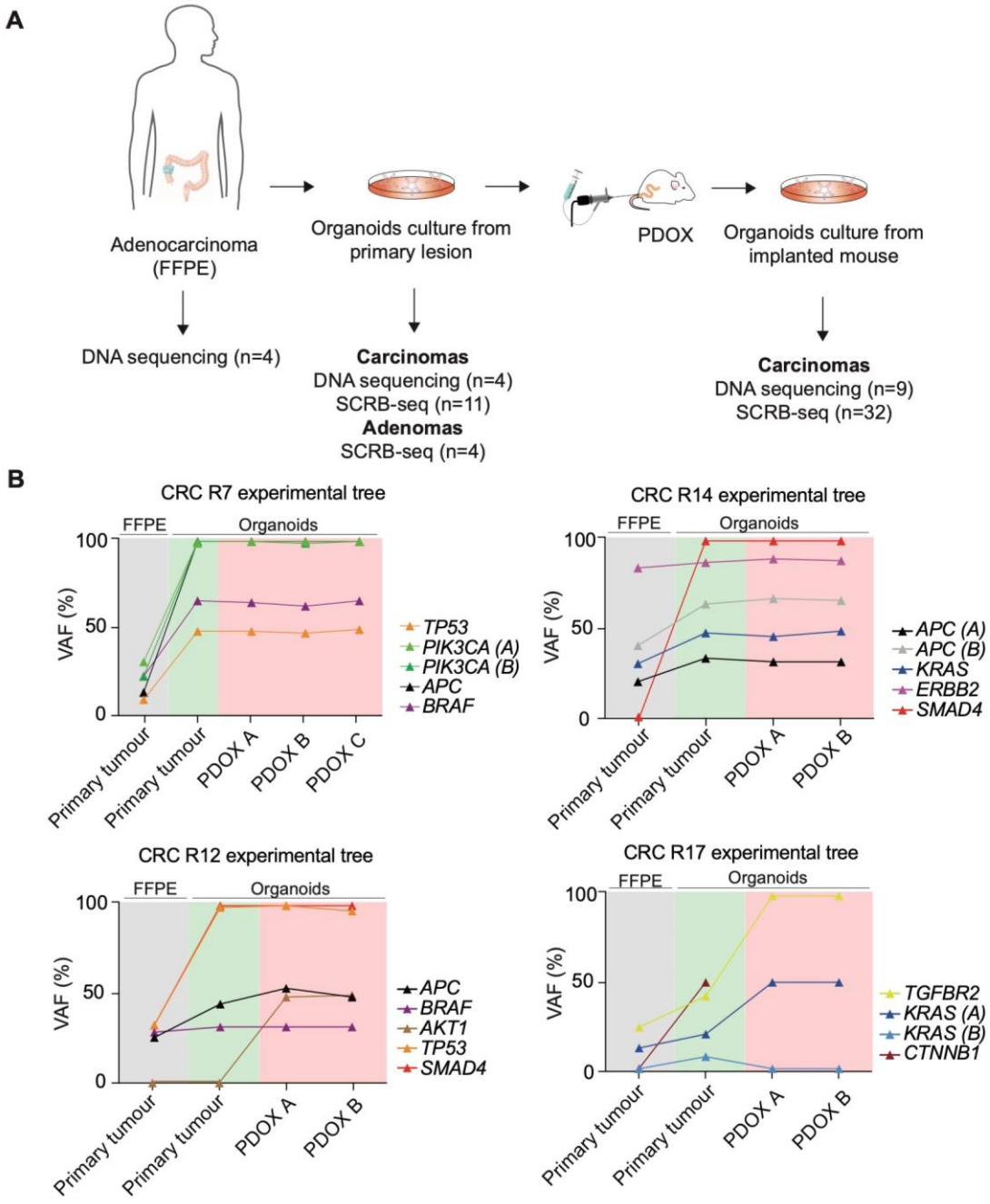


Figure 20 | Organoids retain the genetic features of patient tumours. [A] Workflow of the sequencing strategy for the comparison of tumour tissue (FFPE) and organoids from both the primary human tumour and PDOX. DNA amplicon deep sequencing using a panel of 379 amplicons covering mutational hotspots of 25 recurrent mutated genes in CRC was performed in 4 FFPE and respective organoid samples, and in 9 PDOX-derived lines. Transcriptomic analysis (3´pA RNA-seq) was performed in 11 primary carcinoma and 4 adenoma-derived organoid lines, and 32 lines derived from the tumours of NSG mice after orthotopic transplantations. **[B]** Genetic dynamics of driver mutations in the patient primary tumour (FFPE and organoids) and after orthotopic transplantation (PDOX). For each experimental tree, the letters (A, B, C in the x-axis) correspond to different organoid lines from mice that were transplanted with the same primary tumour organoid line. DNA sequencing and analysis of the human data was performed by Nicole Pfarr from the Pathology department at the Klinikum rechts der Isar.

A mutation in *AKT1* in experiment CRC R12 that was not present in the primary tumour organoid line but was detected in the transplanted mice (PDOX A and PDOX B) and a *KRAS* mutation present in the primary tumour organoids in experiment CRC R17 that was then lost in the PDOX samples make up the only two exceptions. The VAF of the shared mutations did not alter significantly between organoids from the primary tumour and implanted animals in 3 of the 4 experiments. In experiment CRC R17 the VAF of the *TGFBR2* and *KRAS* mutations was higher in the PDOX samples in comparison to the primary tumour organoids which might suggest that some cell clones within the tumour were selected after transplantation. Orthotopic transplantations for each experiment were performed in at least 2 animals (PDOX) at the same time. Strikingly, the VAF of all mutations was not altered between mice injected with the same primary tumour line and from a genetic standpoint they could not be distinguished. These results combined suggest that upon implantation into NSG mice, organoids do not change their mutational landscape.

4.10 Different CRC histologic subtypes can be captured at the transcriptional level

The human organoid cohort used in this study includes tumours from distinct histologic types which are punctuated by different morphologies at the cellular level as well as prognostic values. At the genetic and histological level, we demonstrated that tumour organoid lines after transplantation into mice remain, for the majority of cases, unchanged. We inquired if this could also be seen at the transcriptomic level. In order to understand if there are differences in terms of transcriptomic signatures between the different CRC subtypes included in the study, we performed 3´pA RNA-seq on 11 organoid lines derived from primary human carcinomas and 4 derived from adenomas. In addition, 32 organoid lines resulting from carcinoma organoids transplantations (PDOX) were included (**Figure 20A**).

Figure 21A shows the PCA graph containing these samples after sequencing. Each individual sample is represented by 2 points which are technical replicates. Samples

labelled as “CRC R7”, “CRC R12”, “CRC R14”, “CRC R16”, “CRC R17”, “CRC R19”, “CRC R20” and “CRC R21” correspond to samples from both the primary patient carcinoma and the respective PDOX-derived organoids. Samples labelled with only “Carcinoma” correspond to 3 additional organoid lines derived from human primary CRC which were not implanted. Analysis of the PCA reveals a clear separation between all adenomas and carcinomas and the former tend to cluster together away from any carcinoma lines. Interestingly, 2/3 of the non-implanted carcinomas correspond to early-stage CRC (UICCI Stage I) and were therefore closer to the adenoma samples. In the carcinoma subgroup of samples that have been implanted, 3 main clusters can be seen which mostly refer to distinct histologic subtypes of CRC (Medullary, Serrated, NOS). Micropapillary adenocarcinoma organoid lines (n=3) showed a more heterogenous expression pattern since one of them clustered together with the serrated samples and the other 2 with the NOS subtype. Our results suggest that even though NOS tumours are usually regarded as a separate and single histological entity, there is extensive heterogeneity at the transcriptional level. CRC is generally classified as adenocarcinoma NOS because it resembles normal intestinal crypts or conventional colonic adenoma where no extra histological findings are reported. However additional or different morphologic and histologic patterns can be present in tumour lesions which prompted the identification of additional CRC subtypes (155). Micropapillary carcinomas represent a recently identified CRC subtype shown to be associated with *BRAF* and *KRAS* mutations, which are also common in serrated carcinomas (198). Moreover, CRC with pure micropapillary patterns are extremely rare and most micropapillary lesions coexist with another histological subtypes (199). Thus, this explains the clustering at the transcriptional level of these samples with other CRC subtypes (serrated and NOS). Serrated adenocarcinomas constitute a subset of lesions characterized by neoplastic glands with prominent epithelial serrations, low nucleus-to-cytoplasm ratio, eosinophilic and abundant cytoplasm, and vesicular nuclei (155). The medullary CRC subtype and respective PDOX samples formed another cluster in the PCA. Medullary carcinoma is another rare histologic type usually diagnosed when cancer cells appear as solid or sheet-like structures where lymphocytic infiltration is prominent with tumour-infiltrating lymphocytes and neutrophils. Although these cancers are often microsatellite unstable, their microenvironment features sets them apart from other subtypes as suggested by the PCA plot (200). The heatmap in **Figure 21B** shows the top 2000 regulated genes across the human organoid cohort. Of note, the adenoma samples included in this analysis belong to the same histological type (tubullovillous) which reflects the high level of similarity at the transcriptional level within this group. Remarkably, all implanted carcinoma and the respective PDOX-derived organoid lines cluster together and are markedly different from the other experiments. Furthermore, gene expression patterns

allow the distinction of different experiments reflecting what is observed in the PCA plot. These results combined suggest that no major alterations at the transcriptional level can be assessed between the implanted human primary and mouse PDOX-derived organoid lines.

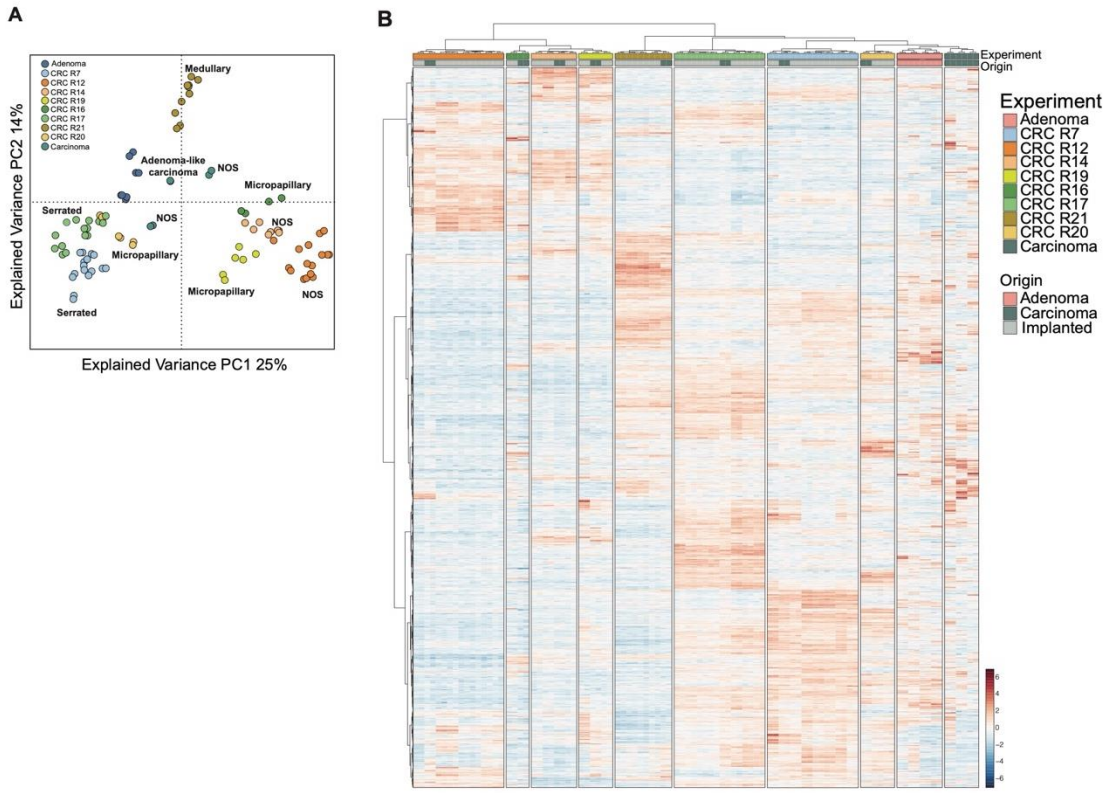


Figure 21 | Transcriptomic profiling of human organoids before and after mouse orthotopic transplantations. [A] Principal component analysis (PCA) based on the 10% most variable genes patient primary carcinoma (n=11), adenoma (n=4) and mouse PDOX (n=32) organoid samples. For each sample a technical replicate was used to enhance statistical power in the analysis. Components 1 and 2 explain 39% of the total variance in the dataset. Each colour corresponds to a group of organoid samples ("Adenoma" refer to primary patient adenomas; "Carcinoma" refer to primary patient carcinomas that were not transplanted into mice; "CRC Rx" refer to primary patient carcinomas that were implanted and samples derived from the respective PDOXs). The histological subtypes of each human CRC are annotated. **[B]** Unbiased hierarchical clustering of transcriptome profiles of human organoid samples before and after transplantations using the 10% most variable genes in the dataset. Analysis of transcriptome raw data and clustering of samples was performed by Thomas Engleitner.

5 Discussion

5.1 Organoids and GEMMs as complementary approaches for CRC research

CRC is currently the fourth most commonly diagnosed cancer (6.1% of all cancer cases) and the second most deadly (9.2% of all cancer-related deaths) after lung cancer (12). Even though cases are steadily decreasing in developed countries, CRC is still considered a major health issue with a total economic burden of 19.1€ billion in Europe alone (data from the year of 2015) (201). As it is the case for many other cancer entities, one of the biggest problems of the clinical management of CRC comes from late-stage diagnosis which usually reflects the lack of pathognomonic symptoms in early stages of tumour development. At the time of diagnosis, in a significant percentage of cases, patients present already signs of advanced disease including distant metastases mostly in the liver or lungs. In fact, the 5-year survival rate in patients with metastatic CRC is only 12-14% (202). Therefore, implementation of screening techniques such as colonoscopy could prevent an early lesion to progress to cancer (15). The normal intestinal epithelium architecture is maintained by a combination of finely controlled signalling cues orchestrated by distinct cell types (7). Proliferative signals typically act in the stem cell compartment at the bottom of intestinal crypts while differentiation occurs towards the villus typ. CRC typically develops through the stepwise accumulation of genetic and epigenetic insults which affects this equilibrium and leads to cell transformation. Classic CRC cases initiate with Wnt pathway disruption, that through serial acquisition of mutations in other genes (*KRAS*, *SMAD4* and *TP53*) and chromosomal gains and losses prompts tumour progression (11). However, 15-30% of CRC cases develop through the “serrated pathway” which start in most cases with a *KRAS* or *BRAF* mutation. The clinical presentation of these tumours is very heterogenous and prognosis is often dependent on the presence of additional genetic features like MSI or CIMP (203). Likewise, although the main alterations have been identified for serrated tumours, the routes of tumour progression remain poorly understood.

In the scope of this thesis, we employed tools and resources in order to deepen the knowledge of the features of serrated tumours and to identify the events responsible for the progression of these lesions. As such we used previously established mouse models of classical (*Vil-Cre; Apc*$^{fle1-15}$) and serrated CRC (*Vil-Cre; Kras*^{G12D}, *Vil-Cre; Braf*^{V637E} and *Vil-Cre*; *Pik3ca*^{H1047R}). The possibility to study disease under a defined clean genetic background which allows the interrogation of single genetic alterations in cancer aetiology is largely provided by GEMMs. Using these 4 mouse models we can assess in a

comprehensive manner different pathways of CRC formation. Since our main interest is the study of serrated CRC, the use of a model reflecting aspects of classical CRC was deemed essential for proper comparison of these 2 entities. In our setting we showed that the loss of *Apc* has more drastic effects in terms of tumour formation as seen by mouse survival, and number of tumours per mouse in comparison to the serrated models (**Figure 7**). Inactivating mutations in *APC* lead to a dysfunctional protein which results in the activation of Wnt signalling (11). This is seen for most sporadic classical CRC cases which suggests a central role for *APC* in disease aetiology. Other cancer-initiating driver alterations leading to MAPK and PI3K pathway disruption led to longer times to tumour formation and usually with less lesions per animal. Interestingly, within the serrated cohort these numbers were very similar between the three genotypes which suggest that oncogenic activation of $Kras^{G12D}$, $Braf^{V637E}$ and $Pik3ca^{H1047R}$ have similar effects in terms of tumour initiation. Even though 2D cell lines have been extensively used for the study of several cancer entities including breast, lung and pancreatic cancer, the use of this model in colon tumours is limited by the low efficiency of the isolation process (204). Furthermore, establishment of culture competent 2D cell lines are only possible from advanced lesions (carcinomas) which would make our study unfeasible since the GEMMs used here develop mostly adenomas and almost never reach the stage of advanced disease. Organoids, instead, provide a much more viable method that reflects the morphology and physiology of the living organ and is not limited to the use of advanced tumours since normal tissue can also be cultured (119). In order to interrogate which genetic mechanisms are involved in the progression of serrated tumours in these mouse models we established a comprehensive living organoids biobank that together with GEMMs provides the advantage of combining *in vivo* and *in vitro* complementary approaches for the study of CRC. Furthermore, the possibility to culture untransformed cells derived from normal tissue from both mouse and human samples further highlights the importance of using organoids for genetic studies. Both human and murine organoid lines were established from normal tissue, hyperplasia and neoplastic lesions at different stages of which captured disease progression in an unprecedented manner (**Figure 8**). Although not mentioned in this thesis, we established several additional organoid cultures from different genetic CRC mouse models used in the laboratory, accounting in total for more than 600 successfully generated lines. Even though these lines were not used in the scope of this thesis, we would like to emphasize the great efforts involved in the creation, to our knowledge, of the largest and most complete murine organoids biobank in scientific research to date.

In-house organoid culture conditions allow the growth of both normal tissue/hyperplastic and tumour-derived murine organoids. Therefore, it is crucial that proper care is taken when the tumour is sampled for cell extraction. The concepts of sampling bias

and ITH are technically related and can, to a great extent, affect the interpretation of results. As stated in the Results section, for every lesion, half of the tissue was used for histological evaluation and half for cell isolation. Macroscopic assessment of lesions upon mouse necropsy does not allow the distinction at the microscopic level between tumour and normal tissue, which might result in a culture containing both normal and tumour organoids (sampling bias). In samples derived from endogenous mice, tumour cells and healthy cells cannot be distinguished morphologically or genetically (e.g. PCR or by means of a reporter-expressing line) in most cases, although in some instances it is possible to distinguish morphologically normal from tumour organoids in a "mixed" culture (which would lead to a dropout culture). Ultimately, when these two options are not possible, another way to prove that the organoids in culture are derived from normal or tumour tissue from endogenous models is far more complex and involves implanting them and monitor both tumour forming capacity and survival of injected animals. In this work however we did not explore this option for the intended purpose of culture purity assessment. Hence, using the aforementioned cultures for downstream analysis might be misleading. We suspect that a few organoids lines in our cohort might have been affected by sampling bias (**Figure 9**). The adenoma-derived lines "TM7513 T1-2", "TM7510 T2-1" (*Kras*^{G12D}), "TM1541 T1" and "TM1179 T1-3" (*Braf*^{V637E}) had features that were not present in all other tumour-derived organoid lines such as very low mutation number, absence of recurrent mutations (with the exception of a *Lama4* mutation in sample "TM1179 T1-3") and almost complete absence of T>C and T>A transversions. These results combined suggest the possibility of a mixed (or low tumour content) organoid culture with an overrepresentation of non-transformed cells. Two opposing studies (129, 144) reported different outcomes when organoids derived from normal and tumour tissue were cultured together. In the former, normal colonic human organoids were seen to outcompete tumour organoids *in vitro* (129), whilst in the second study, wt cells when co-cultured with mouse tumour organoids were selectively eliminated by apoptosis (144). It is important to bear in mind that organoids from different species (human or mouse) differ in many aspects: for instance, human advanced carcinomas show higher levels of aneuploidy as a result of chromosomal rearrangements while mouse cells are genetically more stable (205). Differences in genome stability rates caused by aneuploidy (which leads to apoptosis of tumour cells) between human and mouse cells might explain the reported differences in these studies. Among IDO lines however, since the genotype of the organoid line used for implantation differs from the one from the recipient mouse, we can efficiently assess tumour purity of the culture. Of note, a few IDO lines had to be excluded from this study due to contamination with organoids from the recipient mouse.

Another constraint that requires consideration for this study is the fact that some tumours are intrinsically heterogenous (ITH). Histological assessment of a tumour often supposes the existence of a spectrum of morphological features that can coexist in the same lesion (92). Hence, organoids isolation from such tissues may select for parts of the tumour that are either more proliferative/invasive or for areas that are absent in the half used for histological grading. In our sample cohort, low- and high-grade adenomas were grouped as one (adenomas) and some samples even though graded as adenomas contained foci of more advanced lesions. Particularly, the *Kras*^{iGD} (*Kras ratios*=0.99) adenoma sample in **Figure 9C** presented foci of advanced adenoma in transition to adenocarcinoma at the histological level. Given the fact that iGD was almost exclusively present in the carcinoma group, we cannot exclude that organoids established from this sample were not from the adenoma part but instead from more advanced areas of the lesion.

5.2 Increased gene dosage (iGD) of *Kras*^{G12D} and *Braf*^{V637E}-driven CRC

Genetic characterization of the organoid cohort, helped us to identify the main alterations occurring specifically in tumour cells and pinpoint stage-specific mechanisms of tumour progression (hyperplasia, adenoma and carcinoma). Interestingly, the majority of *Kras* and *Braf*-driven tumour lines showed disruption of Wnt pathway either by *Ctnnb1* or *Apc* mutations. Wnt pathway disruption has been reported for *Braf*^{V637E} (9) however it represented a new finding in the *Kras*^{G12D} setting (8). In CRC, not only the type but also the order of genetic events dictate carcinogenesis. *APC* and *KRAS* mutations are common events in CRC, and in the vast majority of cases *APC* mutation occurs prior to MAPK dysregulation (11). The crosstalk between Wnt signalling and RAS-ERK pathway has been extensively studied, however always assuming the aforementioned order of events (206-209). Jeong et al., (2012) showed that Wnt/β-catenin signalling stabilizes RAS-ERK pathway activation in a process dependent on the phosphorylation of RAS molecules (H, N and KRAS) by GSK3β kinase. Under Wnt stimuli the destruction complex (from which GSK3β is part of) is quenched by Frizzled which leads to accumulation of both β-catenin and RAS molecules (208). Based on this evidence we suggest that following *Kras*^{G12D} oncogenic activation, *Ctnnb1* or *Apc* mutations work as the stimuli for Wnt pathway activation that will further amplify MAPK pathway in these cells and boost tumour progression. Sixty-seven percent of *Kras*^{G12D}-driven carcinomas lost *Cdkn2a* at the CNV level. *CDKN2A* locus encodes two physically linked tumour suppressor genes, p16INK4A and p14ARF which codes for p19Arf in mice (210). It has been shown that in a non-transformed context *Kras*^{G12D} oncogenic activation leads to OIS (211-213). In murine cells

the inactivation of both p53 its direct upstream regulator, p19Arf were shown to be sufficient to bypass *Kras*^{G12D} induced OIS (181, 214). In line with this, all *Kras*^{G12D} and *Braf*^{V637E} carcinoma samples from both genotypes showed either *Trp53* mutations and/or *Cdkn2a* deletions, which suggests that these mechanisms are responsible for overcoming the senescence phenotype induced by the oncogene. Considering all carcinoma-derived lines in the study (endogenous + implanted) *Kras*^{iGD} was reported in 76% of carcinomas compared to 8.6% of adenomas and 0% of hyperplasia. The most common genetic event leading to this was arm gain level of chromosome 6 observed in 79% of *Kras*^{iGD} samples followed by CN neutral LOH in 21% of samples which differs from the percentages reported for mPDAC cell lines (152). The specific gene dosage increase of the mutant *Kras*^{G12D} allele indicates that there is selection for oncogenic signalling intensification during intestinal cancer progression.

BRAF^{V600E}-mutant CRCs occur mostly in the proximal colon, display mucinous histology, a poor differentiation status and show a distinct metastasis pattern (79). *BRAF*^{V600E} oncogenic activation has also been shown to induce OIS through the upregulation of *p16*^{Ink4a} and that *p16*^{Ink4a} down-regulation by promoter methylation is necessary for tumour progression in mice and humans (9, 215, 216). Additionally, Wnt activation in dysplasia progression and *Trp53* inactivation in advanced lesions constitute additional mechanisms by which cells can escape OIS and resume proliferation (9). In the present *Braf*^{V637E}-driven organoid cohort we report similar results where 11/16 tumour samples have a *Ctnnb1* or *Apc* mutation and 3/16 samples (including the carcinoma) are *Trp53* mutant. Strikingly, also in this setting we report allelic imbalances in the progression of *Braf*^{V637E}-driven tumours mostly observed in the carcinoma group. The mechanism leading to iGD was found to be exclusively arm-gain level of chromosome 6 where *Braf* is also located, occurring in 60% of carcinomas. Of note, 1 carcinoma-derived organoid sample of each genotype (*Kras*^{G12D} and *Braf*^{V637E}) showed a decreased oncogenic dosage (*ratios*<0.5) (**Figure 7E**). In these samples, because LOH affects chromosome 6 we hypothesized that the genetic insult selects not the wt but rather the mutant allele, although we don't have an explanation for this finding yet. *BRAF* is mutated in melanoma (~50%), papillary thyroid (~60%), colorectal (~12%) or non-small cell lung cancer (~5%) (217). Recently, copy-number gains harbouring the *BRAF*^{V600E} locus have been reported in therapy-naïve melanoma but not in CRC or lung cancer (218). In *BRAF*^{V600E}-expressing melanoma cases, copy number gains cooperate with the mutation to enhance proliferation by providing a selective advantage to cells with these genetic features. Furthermore, *BRAF* oncogenic amplification has been suggested as a mechanism of targeted therapy resistance in melanoma and lung cancer (219). Cancer therapy regimens have profound effects in tumour architecture causing selective sweeps that eradicate less *fit* clones and

allow resistant clones to take over the tumour (punctuated evolution) (151). Treatment of melanoma and lung cancer PDXs with ERK, MEK and RAF inhibitors selected for *BRAF*-amplified clones that could not be targeted by the drug agents (219). This mechanism has not been reported for CRC though. The different treatment responses shown by distinct cancer entities with the same driver mutation (e.g. $BRAF^{V600E}$) underlies the importance of tissue context and the prospect that different biological mechanisms cooperate with mutations for disease progression and therapy resistance. We therefore suggest that oncogenic amplification of mutant $Kras^{G12D}$ and $Braf^{V637E}$ are common genetic events in the adenoma-carcinoma transition that most likely confer proliferative selective advantage to certain clones within a tumour. Oncogene amplifications are recurrent events in cancer as previously discussed. In the elegant study of Bielski and colleagues (220) the authors reported that around half of all oncogenic driver mutations in human cancers are targeted by additional genetic changes, suggesting that cancer cells gain a selective growth advantage by increasing and tuning the dosage of these mutations. Just like filamentous feathers that were present in non-avian dinosaurs long before modern flying birds were born, the concept of “exaptation” can also be employed here. Within a tumour, allelic imbalances cooperate with a pre-existing oncogenic mutation to produce a mutant allele dosage increase that likely provides a fitness advantage to the evolving malignant clone, which is in line with our observations.

Unlike *APC/KRAS/BRAF*-driven human CRC, no lesion or tumour subtype has been identified with *PIK3CA* mutation as the initiating genetic event. *PIK3CA* mutational status is a prognostic marker for CRC and patients expressing this mutation have decreased overall survival compared to wt tumours (221). In addition, multiple studies report that stage IV CRC patients with *PIK3CA* mutations are resistant to anti-EGFR targeted therapy (222-224). Genetic characterization of $Pik3ca^{H1047R}$-driven tumours revealed absence of Wnt and *Trp53* mutations which reflects findings from previous studies (107). Recently, it was shown that $PIK3CA^{H1047R}$ allelic imbalances confer a stemness phenotype with tumorigenic potential *in vivo* (225). Interestingly, these effects were only seen when both alleles were mutated, and at the transcriptional level there was no difference between WT and heterozygous $PIK3CA^{H1047R}$ cells (225). In our organoids cohort we did not detect gene amplifications at the *Pik3ca/Rosa26* locus (**Figure 3B**) however 43% of carcinomas lost *Cdkn2a* in a similar way to what is reported for the other 2 serrated settings ($Kras^{G12D}$ and $Braf^{V637E}$). Even though it is not reported that *Pik3ca* oncogenic activation leads to senescence, it seems that the same mechanism used to overcome it in *Kras* and *Braf*-driven lesions is partially shared with this genotype. It would be crucial, though, to extend the analysis to more samples in order to draw stronger conclusions.

5.3 Generation of a model for the study of tumour progression *in vivo*

The use of animal models in both basic and translational research changed the way cancer is perceived as a disease. The development of new therapies and treatment strategies was only possible due to a deeper understanding of the relationship between gene expression and phenotype. Only by unveiling basic mechanisms of disease biology, can one know which aspects of tumour behaviour to target. A plethora of cancer mouse models have been developed for the study of cancer and CRC in particular (96). While, these models have provided valuable knowledge about the initiation and progression of CRC, they are not devoid of limitations that need to be considered. Typically, mouse tumour formation occurs 40-89 weeks after birth (**Figure 7C**) which combined with the extensive breeding that is required to reach the desired genotype, amounts for heavy costs related to animal maintenance that may not be supported in some work environments. Most of the aforementioned mouse models develop tumours in the small intestine and thus do not correctly reflect human disease location. Sporadic human CRC occurs typically in the colon and only 5% of all gastrointestinal cancers occur in the small intestine (168). Finally, CRC GEMMs, particularly the ones harbouring single genetic driver mutations develop mostly adenomas and rarely metastasize which constitutes a major caveat for the study of advanced disease (5, 8, 9). In the scope of this thesis, we established an orthotopic (colon) injections-based model that allows us to bypass these constraints (**Figure 10**). By implanting genetically edited organoids or from early-stage tumour lesions (adenoma) into the submucosa of mice we were able to show that progression could be resumed *in vivo* and tumours can develop into advanced localized (carcinoma) and distant disease states (metastases). Several so-called orthotopic CRC models have been established in the past where cells are transplanted in the mouse caecum (111), rectum (113) or through direct damage to the colon mucosa (116). We based our method on the previously published work of Roper J. et al., (2) which offers considerable advantages in comparison to other methods: i) tumours are located in the colon and are formed in the correct tissue layer (lamina propria); ii) tumour formation is considerably fast and recapitulates the histologic features of the implanted line; iii) It is a non-invasive procedure with minimal stress and comorbidities to the animals; iv) tumours are seen in a large percentage of animals; v) organoids require only *Apc* loss for orthotopic engraftment in contrast to what was observed for other studies where engraftment was only possible when additional driver mutations were included (134, 135) and vi) tumour formation can be visually assessed and monitored over time by colonoscopy (non-invasive). Implantation of adenoma-derived organoids into the mouse colon submucosa led to carcinoma formation which was confirmed histologically and genetically. In contrast, no progression was observed when carcinoma-derived organoids

were implanted. Here, it would be expected that more mice developed metastases but we observed similar rates of metastases formation in comparison to mice implanted with adenoma-derived organoids. We hypothesize the lack of metastatic lesions to be related with the survival of mice after implantation. Implanted mice often have symptoms resulting from increasing size of the tumour that eventually leads to the obstruction of the lumen of the colon and prevents proper bowel movements and intestinal clearing in the animal. As a consequence, mice cannot eat and lose weight in a steadily manner until termination criteria are met and the mouse has to be euthanized. Because tumour formation is faster in mice implanted with carcinoma-derived organoids there is also less time for the tumour to invade into distant organs (metastases) which would explain the low number of metastatic lesions in this cohort. We also observed that a few murine and all human adenoma-derived organoid lines did not form tumours after orthotopic transplantation. Although the reasons for these observations were not explored in this work we assume they are linked to the genetic composition of early adenomas in both species. From our experience, successful engraftment and tumour forming capacity supposes the existence of genetic events (such as mutations) that confer increased *fitness* to cells and allow them to survive harsh environments upon transplantation. The aforementioned adenoma-derived organoids might lack these features and therefore engraft without forming tumours in the animal. Concomitantly, we did not see tumours forming after orthotopic transplantation of mouse wt organoids which further supports this hypothesis even though these cells engraft in the mouse colon (data not shown). In fact, the possibility of using organoids derived from healthy tissue has been the focus of great interest in the field of regenerative medicine in recent years (226, 227). It would be definitely of interest to identify the genetic features that enables a cell or cell population to induce tumours *in vivo*, which most likely would give us a deeper understanding of early cancer mechanisms in these genotypes.

Genetic comparison between endogenous and IDOs from different genotypes revealed that clones bearing a specific set of mutations are selected and therefore are most likely necessary for tumour progression while others tend to disappear. Moreover, during progression, tumour clones acquire new mutations that were not initially present. In addition, we report that shared mutations between endogenous and IDO lines have increased frequency which argues in favour of selection of these events. The same adenoma line injected in different animals showed almost the exact same genetic features after transplantation (IDO) including shared mutations that and CNV profile, thus suggesting that tumour progression is a very conserved process (**Figure 12D**). Particularly, several IDOs showed an increased *Trp53* mutation frequency due to LOH occurring on the wt allele of the same locus. Loss of WT *TP53* by LOH is frequently found in human CRCs carrying *TP53* mutations (55). Our results are also in line with a recently published study where it

was shown that LOH occurring on the *Trp53*^{WT} allele was required for metastases formation when *Trp53*-mutant cells were transplanted into mice (228). Furthermore, in *Trp53*-mutant cells, LOH on the wt allele confers protection from anoikis to epithelial cells which suggests that this facilitates tumour cells to invade distant tissues. Concomitantly, in our observations, mutant *Trp53* leads to higher levels of aneuploidy as seen by increased number of alterations at the CNV level before and after transplantations of *Trp53*-mutated organoids (**Figure 12E**). Similar findings have also been reported elsewhere (229).

Interestingly, WES data from both endogenous and IDO matched organoid lines revealed that metastases did not seem to differ significantly at the genetic level from the respective primary tumour lesions from the same mouse. These results are in line to what is observed in human lesions as well, where no genetic differences were observed when comparing primary tumour with the respective liver metastasis tissue (230).

5.4 Transcriptomic signatures of serrated organoids during disease progression

In addition to providing a suitable platform to study CRC progression, the use of an orthotopic injection model increased the number of carcinoma lesions in the cohort. Thus, using our unique set of organoid samples comprising distinct disease stages (hyperplasia → adenoma → carcinoma) driven by 3 distinct genetic alterations (*Kras*^{G12D}, *Braf*^{V637E} and *Pik3ca*^{H1047R}) we inquired which pathways are regulated during tumour progression.

In the *Kras*^{G12D} cohort, adenoma-derived organoids activate mainly EMT programs, together with Wnt pathway upregulation and remodeling of the ECM. EMT is a biologic process that forces an epithelial cell to undergo several biochemical changes from a polarized morphology towards a mesenchymal phenotype, thereby providing them with increased migratory capacity, invasiveness and production of ECM components (231). It has also been shown that CRC lesions (primary tumour vs metastases) show upregulation of distinct ECM-related proteins which underlies different invasion properties of the cells in each lesion (232). In line with this, enriched pathways in the comparison between adenoma and hyperplasia suggest that transformed cells undergo a series of conformational changes in order to become more independent of cell-to-cell contact. Concomitantly, increased proliferation induced by upregulation of Wnt signalling and ECM degradation provides the ideal conditions for an invasion phenotype. Comparison of advanced lesions (carcinoma) with non-transformed tissue (hyperplasia) showed that in addition to EMT, several immune-related pathways are upregulated in the former, including “TNF-alpha Signalling via NF-κB”, “inflammatory response”, “cytokine-cytokine receptor interaction” and “IL6 JAK STAT3 signalling”. Immune-related signatures namely “cytokine-cytokine receptor interaction”

become stronger in the adenoma → carcinoma transition. *KRAS* mutations have been linked to inflammation by inducing the expression of several cytokines, chemokines and tumour-promoting signalling pathways (233). The activation of JAK/STAT3 pathway by IL6 contributes to several tumorigenic processes in *KRAS*-mutant lung (234) and pancreas tissues (235). In the colon, this pathway is also activated in late-stage tumours as shown in this work. Furthermore, distinct immunomodulatory processes involving different molecules take place at different stages of tumour progression as seen by the gene expression dynamics across the cohort. While some genes such as *Csf3* and *Il8r1* were increasingly expressed from adenoma to carcinoma, others like *Ackr3* and *Cxcl2* are specifically upregulated in advanced disease. Particularly, CSF3 and CXCL2 are critical regulators of neutrophil production and activity, found to be highly expressed in CRC. It has been shown that CSF3 increases proliferation and migration of tumour cells *in vitro* (236, 237). IL8R1 also commonly known as CXCR1 is another receptor for IL8 that modulates the tumour microenvironment (TME) towards a pro-inflammatory state by serving as chemotactic factors for neutrophils (238). Recently, *Ackr3* was shown to promote intestinal tumours in a mouse model by direct perturbation of ribosomal RNA biogenesis (239). In addition, several granulocyte/neutrophil chemokines (*Cxcl1*, *Cxcl2*, *Cxcl5*, *Cxcl7*, *Cxcl15* and *Cxcl17*) are upregulated in carcinoma samples in comparison to the remaining samples of the cohort suggesting an important role for neutrophils in the progression of these lesions. In a recent study, oncogenic activation of $Kras^{G12D}$ was shown to induce an immunosuppressive TME *in vivo* by increasing the number of myeloid derived suppressor cells (MDSC) and decreasing T-cell infiltration in CRC (240). Of note, neutrophils have been shown to bear a tumour promoting effect based on the evidence that they frequently infiltrate both mouse and human colon carcinomas. Here they repress adaptive immune anti-tumour responses and hamper the effect of immune checkpoint blockade therapy by suppressing T-cell function through TGFβ activation (241).

We observed EMT signature to be also present in adenoma samples from $Braf^{V637E}$ and $Pik3ca^{H1047R}$ mutant adenomas. Specifically, in $Braf^{V637E}$ adenomas a strong p53 and Wnt signalling activation was reported in our study and by others (9, 216). In line with our results, expression of mutant *BRAF/Braf* has been shown to lead to the loss of stem cells located in the intestinal crypt. Interestingly, Wnt activation and loss of p53 activity can revert this mechanism and boost tumorigenesis (216, 242). These mechanisms seem to be crucial for initial transformation and tumour establishment. However, we showed that tumour progression (adenoma → carcinoma) relies on the upregulation of Myc-regulated genes and cell cycle-related targets of the E2F transcription factor. Myc overexpression is reported in 50-70% of CRC cases and this is thought to be a downstream effect of Wnt/β-catenin/TCF pathway activation (243, 244), however these studies mostly included CIN tumours where

BRAF mutant cases are not included. In *BRAF*^{V600E}- mutant lung tumours, Myc activation collaborates with Wnt signalling and is required for proliferation of early lesions but not necessary for the formation of invasive cancers (245). We therefore hypothesize that the upregulation of Wnt signalling observed in the hyperplasia to adenoma transition might be responsible for the activation of Myc targets in order to sustain tumour proliferation. Moreover, in a recent study, transcriptomic analysis of *BRAF*^{V600E} mutant thyroid cells treated with the BRAF inhibitors Dabrafenib and Vemurafenib showed “Myc targets” and “E2F targets” to be the most downregulated pathways in comparison to non-treated cells (246) which is in line with our observations. Interestingly, even though a different cancer model with the same driver genetic mutation was used, the same pathways seem to be regulated suggesting that the mechanisms of progression in different cancer entities with the same driver mutation might at least partially overlap.

Given the high heterogeneity and great overlap of carcinoma samples with hyperplasia and adenomas after PCA, in the *Pik3ca*^{H1047R} cohort, we focused our analysis on the processes involved in the transformation of hyperproliferative tissue to dysplasia (hyperplasia → adenoma). Most regulated pathways were found to be related to inflammatory- and immune-related biological processes. Strikingly, the top upregulated pathways in adenoma organoid samples were “IFNα response” and “IFNγ response” followed by “EMT”. Although IFN responses are usually regarded as enhancers of immune responses by either killing tumour cells or improving ICB therapy, they can show opposing effects on anti-tumour responses (191). In fact, both IFN-I (mediated mostly by IFNα) and IFN-II (mediated by IFNγ) responses can contribute to tumour immune evasion by inducing the expression of PD-L1, indoleamine 2,3-dioxygenase (IDO), and arginase in the tumour microenvironment that have suppressive effects on $CD4^+$ and $CD8^+$ T cells. Consequently, the tumour is shielded from immune cells. In fact, this is considered to be an adaptation mechanism of malignant cells upon “sensing” an inflammatory immune microenvironment that “threatens” tumour cells. Moreover, the elevated expression of these molecules can be found on both cancer cells and immune cells which suggests a synergistic effect between the tumour and the TME for tumour maintenance (192, 247).

Even though Wnt signalling is not often associated with *PIK3CA* activity in CRC, Wnt5a was the most upregulated gene in adenomas in comparison to hyperplasia. Wnt5a is a highly evolutionary conserved non-canonical Wnt ligand that plays a critical role in regulating several biological processes during embryogenesis. Nonetheless, its role in cancer remains ill-defined and was shown to function as either an oncogene or TSG in a context dependent manner (248). In other cancer entities and diseases, interaction between *WNT5a* and *PIK3CA* has been reported (249-251). Of note, Zhao and colleagues (249) showed that upregulation of *WNT5a* increased inflammation and oxidative stress via the

PI3K/AKT/ NF-κB axis in polycystic ovary syndrome. Furthermore, Wnt5a induces tumour cell migration in osteosarcoma (250) and in gastric cancer cells by direct phosphorylation of PI3K/AKT (251). Our data suggest that *Wnt5a* upregulation might cooperate with oncogenic *Pik3ca*^{H1047R} to induce cell transformation and boost cell migration (EMT signature). Across all genotypes, heterogeneity within histologic groups is expected and was observed. Strikingly, carcinoma samples from this cohort showed a markedly distinct transcriptional pattern that overlapped partially with non-transformed tissue samples (hyperplasia and normal tissue). Reasons for this are elusive though. Sampling bias can explain extreme behaviours of specific samples (lack of SNVs, decreased oncogenic ratios, similarity to other tissues on a transcriptional level) however this is very unlikely to happen in a larger number of samples such as the carcinomas included in this cohort (n=14). Another explanation, assuming no sampling bias took place, could be that progression from early to advanced *Pik3ca*^{H1047R}-driven tumours activate pathways and mechanisms that are also characteristic of non-transformed cells. Although this does not seem likely, also considering the observations for the other 2 cohorts (*Kras*^{G12D} and *Braf*^{V637E}) which show differentially regulated pathways in the different stages of progression, we cannot exclude this possibility.

In sum, we postulate that initial cellular transformation (hyperplasia → adenoma) in *Kras*^{G12D} and *Braf*^{V637E} organoids depends mostly on the acquisition of conformational changes leading to loss of cell-cell contact, suppression of epithelial features and acquisition of more mesenchymal ones (EMT) and increased motility provided by enhanced proliferation (Wnt signalling). At the same time communication between tumour cells and the TME is established which leads to the activation of pro-inflammatory and pro-tumorigenic immune-related pathways that are progressively activated in order to shield the tumour from the immune system. Particularly in *Kras*^{G12D}-driven cancers this communication becomes stronger and mediates the transition from adenoma → carcinoma suggesting that immune cells (e.g. neutrophils) are “called” to the tumour and suppress the activity of cytotoxic immune cells (e.g. T cells).

Progression of *Braf*^{V637E} tumours depends mostly on the upregulation of cell cycle *Myc*-associated genes. Conversely, in the *Pik3ca* setting, IFN-driven inflammation seems to be at the basis of malignant transformation since it is the most upregulated pathway in adenomas in comparison to hyperplasia. At the same time, conformational changes also occur in these cells (EMT) however seem to be secondary to immune evasion and inflammation. Recent studies revealed that both tumour and immune cell compartment can influence each other, suggesting a potential role of immune microenvironment in EMT and tumour metastasis (252, 253). In line with our results, Liu and colleagues (230) compared the transcriptomes of primary tumours from patients who had also metastases with tumours

from patients without metastases and showed that EMT signalling was upregulated in the former.

Transcriptomic analysis of organoids allows the identification of cell-intrinsic signatures and pathways that are regulated during disease progression. It becomes apparent that this process depends not only on the genetic composition of transformed cells but to a great extent also on the communication with the TME. Therefore, the experiments performed in the scope of this thesis comprising mostly organoids offer perhaps a limited view of one of the most complex biological processes known to date: tumorigenesis. In order to obtain a more comprehensive view of the interaction between tumour cells and the TME we additionally sequenced the transcriptomes of a cohort of more than 350 murine tissue samples comprising normal tissue, hyperplasia, adenomas and carcinomas from all 4 genotypes used in this thesis (*Apc*$^{fle1-15}$, *Kras*^{G12D}, *Braf*^{V637E} and *Pik3ca*^{H1047R}). The application of bioinformatic deconvolution tools prompts the attribution of gene signatures to specific cell types thereby unveiling the contribution of each cell compartment (e.g. epithelial or immune origin) in cancer progression (254, 255). The GEMMs included in this study all rely on the germline loss of *Apc* or activation of mutant *Kras*^{G12D}, *Braf*^{V637E} and *Pik3ca*^{H1047R} which does not reflect the somatic acquisition of mutations observed in sporadic human CRC. Therefore we also performed transcriptomic, proteomic and secretome characterization of 20 Cre-inducible organoid lines (108) (7 *R26-CreER*;*Kras*^{G12D}, 6 *R26-CreER;Braf*^{V637E}, 5 *R26-CreER;Pik3ca*^{H1047R} and 2 *R26-CreER* used as control) before and after oncogenic activation mediated by Cre activation (Tamoxifen treatment). This elegant experiment allows the interrogation of the pathways and biological mechanisms that are activated, precisely when the mutation occurs. This experimental setting constitutes a “one-step-back” approach in understanding tumorigenesis since it is done in a purely normal cellular context, only perturbed by the “oncogenic switch”. Both the aforementioned experiments offer a great complement to the present study however it was not possible to include this data at the time the thesis was written.

5.5 Human CRC modulation *in vivo*

The study of human cancers using human-derived samples often rely on the use of archived tissue (FFPE) or fresh tissue for downstream analysis, generation of cell lines which usually require additional genetic modifications, or PDXs using patient-derived tissue or cell lines. There are however, several limitations to the use of these approaches mostly if a setting is needed where tumours have to be maintained longitudinally for therapeutic purposes for instance (256). The possibility to directly transfer human tumours into mice and propagate them *in vivo* offers unique opportunities, making PDXs one of the most

widely used and valuable models for cancer research and drug discovery, albeit with caveats that need consideration. One main limitation seems to be the loss of genomic stability in tumours that some studies report when using PDXs (257, 258). *In vivo* passaging of PDXs leads to the acquisition of additional CNVs not present in the initial tumour which can alter disease behaviour (258). An ideal model to study human tumours *ex vivo* would need to obey certain conditions such as i) showing longitudinal genetic stability over time, ii) offer the possibility to treat and monitor tumour progression and preferentially iii) easy to establish.

In this work, we successfully showed the feasibility of using mice as an avatar model for the study of human CRC, as the genomic, transcriptomic and histological profiles of patient tumours are recapitulated in mice after PDOX. Thus, this model may be suitable for drug screening approaches in order to test different therapy regimens which patients could benefit more from. The use of the organoid system offers several advantages in comparison to other methods. Firstly, organoids comprise the epithelial fraction of the tumour where no other sources of cellular contamination are present, allowing the study of a population of pure cancer cells. This can also be a limitation if research is focused on the interaction of cancer cells with the TME, however this was not within the scope of our work. Secondly, organoids remain stable in culture over several rounds of passaging. It was suggested however that due to the high levels of aneuploidy exhibited by some CRCs, normal tissue cells in the culture can overcome the culture (129). In this work, we did not report this issue as the tumour piece used for organoids isolation was always carefully selected by a pathologist and consisted exclusively of tumour cells. Furthermore, we showed that tumour organoids from different histologic subtypes reflect the main genetic features of the tumour in the patient which indicates no major clonal selection events took place upon *in vitro* establishment of the culture in contrast to what is reported for 2D cell lines (259).

We additionally showed that by implanting carcinoma-derived organoids orthotopically into mice, local and distant disease could be recapitulated. Routes of human CRC metastization include primarily the liver and then lung, which were the sites where metastases were formed in mice as well. Strikingly, the histology of the patient primary tumour was reflected in the tumours resulting from organoids implantation in the mouse. Subsequent genetic analysis of PDOX-derived organoids revealed that mutations found in patient tissue (FFPE) and respective organoid culture were also found in the PDOX lines. NGS showed that, with very few exceptions, both the presence and VAF of driver mutations remain unaltered between primary tumour and PDOX-derived organoids. An expected increase in mutational VAF was observed between organoids from primary tumour and FFPE tissue, since organoids comprise exclusively the epithelial tumour fraction, in contrast to tissue section where other stromal and immune cells are present. Of note, the NGS

strategy employed in this study does not allow the identification of CNVs in the samples, therefore we cannot exclude that there are no changes at this level as reported previously (258). One obvious advantage of using organoids instead of tissue for a PDX approach relies on the fact that organoids can be cultured *in vitro*, cryopreserved and thawed at different timepoints without affecting the stability and overall composition of the culture. In contrast, tissue PDXs can be only passaged *in vivo* which requires several animals in a continuous experiment with no possibility to stop and resume it at a later timepoint. In this work, organoids were injected into several mice only once and we did not perform additional *in vivo* passages of the same organoid line. Transcriptomic analysis of primary tumour and respective PDOX-derived organoids lines revealed no differences between the two states. Furthermore, RNA profiling of organoids faithfully reflected the histological subtypes included in this study. For 2 histological subgroups (NOS and Micropapillary) we saw that often they cluster together when PCA was performed. Classification of NOS tumours is a general term for the lack of specific histological findings (155). Our data shows they are heterogeneous at the transcriptional level which might underlie the existence of other subtypes or microscopic features not identified to date. Micropapillary CRC is graded as such if at least 5% of cells with micropapillary features are found in the tissue slide, and usually coexist with other histologic types. The presence of micropapillary features and conventional adenocarcinoma (NOS) in a lesion is a common finding (199). Our data suggests that these features can be detected at the transcriptional level and allow discrimination of samples within the same histological group. In our cohort, adenomas were very similar at the transcriptional level which is line with the fact they were derived from lesions with the same histological grading. In contrast, non-implanted carcinomas were starkly more different between them as it would be expected.

The aim of this part of the work was prompted by the question regarding the genetic stability of organoids after mouse transplantations. Thus, we, did not focus on the analysis of genes and pathways that are regulated between the CRCs subtypes and adenomas in the cohort. This analysis is currently undergoing and might shed light on some ongoing questions such as why don't adenoma-derived organoids engraft in mice in contrast to carcinomas. Our hypothesis points to the fact that adenoma samples lack the necessary genetic features to engraft in a more adverse environment (mouse colon). Thus, through genetic manipulation by CRISPR/Cas9, for instance, the effect of specific genes could be addressed and possibly bypass this limitation to allow organoids engraftment. In this setting we foresee that the orthotopic injection mouse model could be used for the study of human CRC progression in a similar way as for the mouse counterpart.

Collectively, these results show that our established model based on orthotopic injections of carcinoma-derived organoids in immunodeficient mice is a suitable avatar

model for the study of *ex vivo* human CRC. We foresee the use of this model to expand beyond genetic and transcriptomic studies, and suggest that therapeutic approaches can also be exploited. The feasibility of drug testing in tissue PDXs has been extensively explored in recent years using conventional chemotherapy and targeted immunotherapy by boosting innate immune cells (260-262). Moreover, the use of patient-derived rectal cancer organoids has been proven to be a suitable model for predicting individual patient responses to chemoradiation (263, 264). With the resources created in this work, both *in vitro* and *in vivo* approaches can be complemented for the testing of new treatment options or to predict drug responses in CRC patients.

6 Summary and outlook

A major challenge of therapeutically targeting serrated CRC comes from the limited view of the underlying molecular processes leading to tissue transformation and tumour progression. It becomes apparent that the mechanisms that distinct tumour entities with the same oncogenic mutation use to evolve partially overlap. The loss of *Cdkn2a* and upregulation of p53 pathway are both seen in the progression of $Kras^{G12D}$-mutant PDAC (152) and in CRC (8) as a way to overcome senescence. Notwithstanding, we are still lacking an explanation of why do tumours with the same driver mutation (e.g. $BRAF^{V600E}$ in melanoma and CRC) achieve significant responses in one setting (melanoma) but not in the other (CRC) when the mutation is therapeutically targeted. Clearly, not only the genetic and epigenetic events define a lesion but also the tissue context most likely influenced by the TME, seems to affect tumour behaviour. Therefore, simply by studying human tumour lesions, the reasons underlying these differences are not likely to be unveiled. To understand disease aetiology, characterization is only the first step and must be followed by the use of suitable models that reflect human disease where also functional studies can be employed.

The major findings reported in this thesis are summarized in **Figure 22**. The establishment of an *in vitro* epithelial cellular system (organoids) to study tumour progression elucidated the genotype-specific molecular events that are activated in different stages of disease progression. Genetically, early $Kras^{G12D}$ and $Braf^{V637E}$ driven serrated lesions activate Wnt signalling by the acquisition of mutations in either *Ctnnb1* or *Apc* which cooperates with EMT for neoplastic transformation of tumour lesions. Oncogenic amplification of the mutant allele together with loss of a functional p53 protein likely drives the progression of these tumours by modulating the tumour-immune cell interaction in the TME thereby shielding the tumour from its action. Although an EMT signature is also present in *Pik3ca*-driven tumorigenesis our results put inflammation and immune regulation as the main biological events responsible for the hyperplasia-adenoma transition. Furthermore, we developed a system that faithfully recapitulates key genetic and histological features of murine and human tumours and can be used for disease modulation *ex vivo*. Also, we showed that different human subtypes of CRC have distinct transcriptomic profiles that highlight discrepancies within each histological subgroup. The work developed in the scope of this thesis covers basic and functional aspects of serrated CRC by incorporating *in vitro* cellular with *in vivo* functional models to study tumour progression. In the next paragraphs we refer to aspects and experiments, most of them ongoing, that in our view would complement this project.

We are currently expanding our genetic analysis (WES and lcWGS) by including a larger cohort of both endogenous and IDO serrated samples. We wish to confirm the findings from this work and hopefully identify additional recurrent genetic mechanisms that define each stage of tumour progression particularly in the *Pik3ca* setting. Our observations suggest that progression of *Kras* and *Braf* mutant tumours is driven by oncogenic amplification that mediates the cross-talk of cancer cells with immune cells thereby "hiding" the tumour from the cytotoxic activity of the adaptive immune system. This is another elegant example of the elaborate complexity of the tumorigenic process, in which malignant cells make use of other components of the organism (self) in order to sustain itself much like a parasite infecting a host. Although this has been identified as a hallmark of cancer (265) the identification of the processes and cellular intervenients remain, to a certain extent, elusive. We foresee that the transcriptomic data from the mouse tumour tissue mentioned in the **Discussion** chapter will give us further insights of how the interaction of cancer cells with the tumour microenvironment takes place, potentially helping us to define which immune cells are more frequent in a given disease stage and what part could they play in tumour progression.

Our findings show that many cytokines and cytokine receptors (for instance *Ackr3*, *Cxcl2*, *Il8r*, *Cxcl15*, *Csf2*, *Csf3*) are differentially regulated in distinct stages of tumour progression. Given the known function of these molecules in the recruitment of myeloid cells (e.g. neutrophils and macrophages) we plan to functional dissect their role in the transition from early to advanced disease. Hence, we will create individual *knockouts* for each cytokine in organoids using CRISPR/Cas9 in *Kras* and *Braf*-mutant tumour organoids with different oncogene doses (Het and iGD) and subsequently address the *in vivo* tumorigenic potential by orthotopically injecting them into immunocompetent mice. Downstream analysis by 3´pA RNA-seq and FACS of the immune cell populations in the tumour tissue of mice injected with *knockout* organoids compared with wt organoids will help us understand the contribution of each cytokine in disease progression. As a complementary approach, in order to interrogate the biological relevance of cytokine receptor signaling on intestinal tumour cell fitness, we plan to perform a pooled *in vitro* Cas12 CRISPR screen targeting all CXC chemokine/Csf receptors *(Csf1r*, *Csf2ra*, *Csf2rb*, *Csfr3*, *Cxcr1*, *Cxcr2*, *Cxcr3*, *Cxcr4*, *Cxcr5*, *Cxcr6*, *Ackr3)* alone or as double- and triple-*knockout* combinations. The use of the Cas12 system allows the combinatorial genetic targeting through processing of multiple sgRNAs from a single transcript (266).

In addition, we are currently addressing the biological consequences of iGD on a cellular level. Therefore, we plan to establish an *in vitro* organoids-based assay in which the oncogenic levels of $Kras^{G12D}$ and $Braf^{V637E}$ can be modulated and controlled to further understand the impact of oncogene amplification on a transcriptional level. Together with

the experiments performed using the inducible CreER^{T2} (tamoxifen-inducible) organoid lines (mentioned in the Discussion chapter) where profiling of transcriptomes, global proteomes, phospho-proteomes and secretomes was done precisely when the oncogenic mutation is “switched on” in an otherwise wt environment, we believe we will get a significantly more comprehensive view of the programs and events that precede malignant transformation in the serrated setting.

Carcinogenesis is a far more complex program that includes many more aspects than the ones explored in this thesis. Epigenetics in tumorigenesis constitute an additional set of mechanisms that together with DNA alterations are crucial for tumour formation and maintenance. We are currently investigating the epigenetic landscape of tumour organoid lines from different disease stages by ChIP-seq, ATAC-seq and HI-C as complementary approaches to identify transcription factors and epigenetic patterns that are associated with iGD, cancer cell metastasis or organ tropism of metastasis. We will focus our downstream analyses on the identification of super-enhancers (SEs) because of their prominent role in the regulation of genes controlling developmental programs, cell identity and cell-type specific functions. Also here, the *in vitro knockout* of relevant transcription factors (TFs) in tumour and hyperplasia-derived organoid lines could potentially affect tumour progression. This hypothesis would be also complemented with *in vivo* experiments in a similar fashion as aforementioned.

Finally, by using our established orthotopic injections mouse model, one could even think of testing different therapy approaches (e.g. ICB, immunotherapy, chemotherapy) to prevent adenoma-carcinoma transition. The feasibility of this method and the resources generated in the scope of this work extend beyond the mouse context. A similar approach can be employed for the human counterpart where also advanced aspects of tumorigenesis were shown to be faithfully recapitulated by the model. *Ex vivo* tumour treatment can also be employed using the mouse as an avatar and address aspects of tumour biology that would otherwise not be possible by using solely *in vitro* approaches. In addition, we plan to extend our human organoid cohort and properly define the transcriptomic features that truly identify a given histologic subtype.

In sum, this work complements existing knowledge of the biology of serrated CRC by combining several technologies and experimental approaches where different aspects of tumour biology could be connected and therefore better understood. Initial efforts were put in overcoming the caveats of previous models in the lab (GEMMs) that were not suitable at that time to address our research questions.

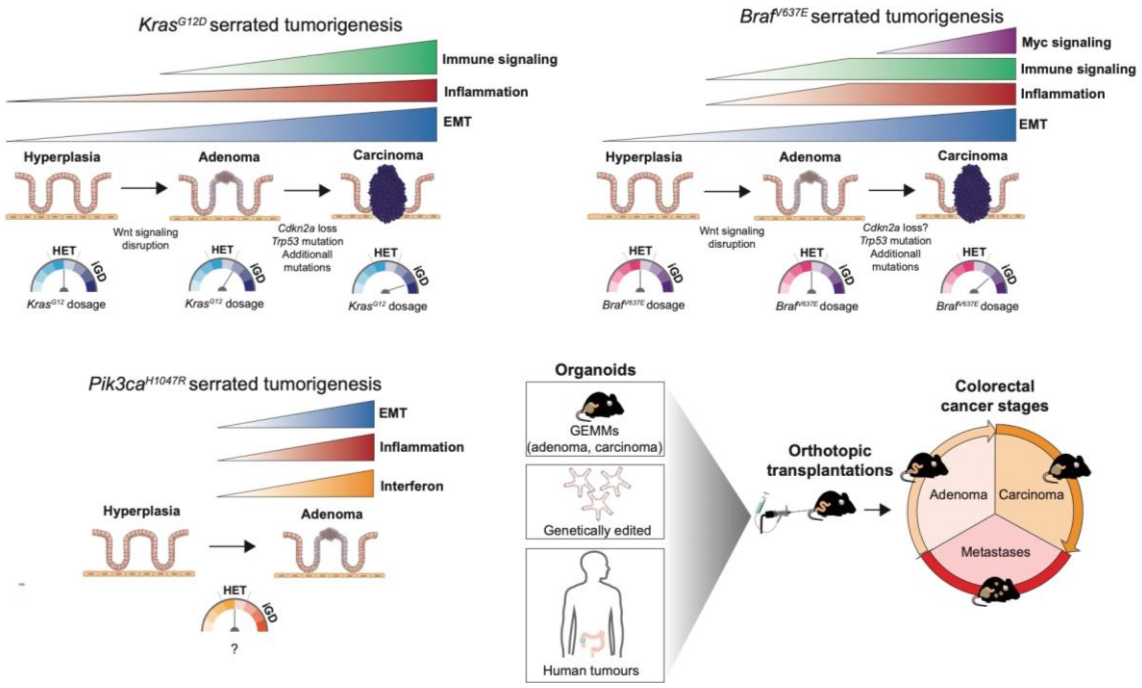


Figure 22 | Graphical summary of the main findings in the scope of the work of this thesis. From early on, *Kras*^{G12D} and *Braf*^{V637E}-driven tumours acquire mutations in Wnt pathway effectors that lead to its overactivation. These mutations cooperate with EMT and inflammatory signals in initial stages of transformation which remain active thorugh progression. The transition from early to advanced disease is mediated by the acquisition of additional mutational events such as loss of *Cdkn2a* and *Trp53* mutations and the interaction with the immune system orchestrated by oncogenic amplifications (*Kras*^{G12D}) or upregulation of Myc target genes and cell cycle effectors (*Braf*^{V637E}). Oncogenic amplifications of the *Braf*^{V637E} allele are seen exclusively in carcinomas albeit to a lesser extent when compared to the *Kras*^{G12D} cohort. Neoplasic transformation in the *Pik3ca*^{H1047R} setting is mostly dependant on interferon signaling in addition to EMT and inflammation. Finally the cellular resources and tools generated in this thesis allows the *ex vivo* and *in vivo* modeling of distinct stages of both murine and human CRC.

By bypassing the constraints of long mouse survival times, lack of advanced disease and the “non-orthotopic” nature of CRC models in the lab, the tools generated here offered us the right experimental setting to develop the project. At the same time the study of advanced disease was made possible by injecting organoids derived from early-stage tumours into recipient mice so that tumorigenesis is resumed. Additionally, the use of organoids allows the study of non-transformed cells that carry an oncogenic mutation (*Kras*^{G12D}, *Braf*^{V637E}, *Pik3ca*^{H1047R}) or loss of TSG (*Apc*) which is a clear advantage in comparison with 2D cell lines where only cells from advanced tumours can be cultured.

We hope that the resources generated here, can be used to spark additional questions and projects so that one day cancer patients could benefit from these “small” steps, which ultimately should be the main drive of cancer research.

7 References

Uncategorized References

1. Ferlay J EM, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer 2020 [Available from: https://gco.iarc.fr/today, .
2. Roper J, Tammela T, Cetinbas NM, Akkad A, Roghanian A, Rickelt S, et al., In vivo genome editing and organoid transplantation models of colorectal cancer and metastasis. Nat Biotechnol. 2017;35(6):569-76.
3. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.
4. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al., The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin. 2017;67(2):93-9.
5. Cheung AF, Carter AM, Kostova KK, Woodruff JF, Crowley D, Bronson RT, et al., Complete deletion of Apc results in severe polyposis in mice. Oncogene. 2010;29(12):1857-64.
6. Vendramin R, Litchfield K, Swanton C. Cancer evolution: Darwin and beyond. EMBO J. 2021;40(18):e108389.
7. Gehart H, Clevers H. Tales from the crypt: new insights into intestinal stem cells. Nat Rev Gastroenterol Hepatol. 2019;16(1):19-34.
8. Bennecke M, Kriegl L, Bajbouj M, Retzlaff K, Robine S, Jung A, et al., Ink4a/Arf and oncogene-induced senescence prevent tumor progression during alternative colorectal tumorigenesis. Cancer Cell. 2010;18(2):135-46.
9. Rad R, Cadinanos J, Rad L, Varela I, Strong A, Kriegl L, et al., A genetic progression model of Braf(V600E)-induced intestinal tumorigenesis reveals targets for therapeutic intervention. Cancer Cell. 2013;24(1):15-29.
10. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al., The repertoire of mutational signatures in human cancer. Nature. 2020;578(7793):94-101.
11. Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. N Engl J Med. 2009;361(25):2449-60.
12. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
13. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017;66(4):683-91.
14. Kasi PM, Shahjehan F, Cochuyt JJ, Li Z, Colibaseanu DT, Merchea A. Rising Proportion of Young Individuals With Rectal and Colon Cancer. Clin Colorectal Cancer. 2019;18(1):e87-e95.
15. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. Lancet. 2019;394(10207):1467-80.
16. Nassar D, Blanpain C. Cancer Stem Cells: Basic Concepts and Therapeutic Implications. Annu Rev Pathol. 2016;11:47-76.
17. He J, Efron JE. Screening for colorectal cancer. Adv Surg. 2011;45:31-44.
18. Boland CR, Sinicrope FA, Brenner DE, Carethers JM. Colorectal cancer prevention and treatment. Gastroenterology. 2000;118(2 Suppl 1):S115-28.
19. Van Cutsem E, Nordlinger B, Cervantes A. Advanced colorectal cancer: ESMO Clinical Practice Guidelines for treatment. Ann Oncol. 2010;21 Suppl 5:v93-7.
20. Edwards MS, Chadda SD, Zhao Z, Barber BL, Sykes DP. A systematic review of treatment guidelines for metastatic colorectal cancer. Colorectal Dis. 2012;14(2):e31-47.
21. Gryfe R. Inherited colorectal cancer syndromes. Clin Colon Rectal Surg. 2009;22(4):198-208.
22. Bogaert J, Prenen H. Molecular genetics of colorectal cancer. Ann Gastroenterol. 2014;27(1):9-14.
23. Darwich AS, Aslam U, Ashcroft DM, Rostami-Hodjegan A. Meta-analysis of the turnover of intestinal epithelia in preclinical animal species and humans. Drug Metab Dispos. 2014;42(12):2016-22.

24. Miron N, Cristea V. Enterocytes: active cells in tolerance to food and microbial antigens in the gut. Clin Exp Immunol. 2012;167(3):405-12.
25. Miller H, Zhang J, Kuolee R, Patel GB, Chen W. Intestinal M cells: the fallible sentinels? World J Gastroenterol. 2007;13(10):1477-86.
26. Birchenough GM, Johansson ME, Gustafsson JK, Bergstrom JH, Hansson GC. New developments in goblet cell mucus secretion and function. Mucosal Immunol. 2015;8(4):712-9.
27. Gribble FM, Reimann F. Function and mechanisms of enteroendocrine cells and gut hormones in metabolism. Nat Rev Endocrinol. 2019;15(4):226-37.
28. Gerbe F, Sidot E, Smyth DJ, Ohmoto M, Matsumoto I, Dardalhon V, et al., Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites. Nature. 2016;529(7585):226-30.
29. Sato T, van Es JH, Snippert HJ, Stange DE, Vries RG, van den Born M, et al., Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. Nature. 2011;469(7330):415-8.
30. Ritsma L, Ellenbroek SIJ, Zomer A, Snippert HJ, de Sauvage FJ, Simons BD, et al., Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging. Nature. 2014;507(7492):362-5.
31. Clevers HC, Bevins CL. Paneth cells: maestros of the small intestinal crypts. Annu Rev Physiol. 2013;75:289-311.
32. Nusse R, Clevers H. Wnt/beta-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. Cell. 2017;169(6):985-99.
33. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, Cozijnsen M, et al., Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature. 2007;449(7165):1003-7.
34. de Lau W, Barker N, Low TY, Koo BK, Li VS, Teunissen H, et al., Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. Nature. 2011;476(7360):293-7.
35. Storm EE, Durinck S, de Sousa e Melo F, Tremayne J, Kljavin N, Tan C, et al., Targeting PTPRK-RSPO3 colon tumours promotes differentiation and loss of stem-cell function. Nature. 2016;529(7584):97-100.
36. Alexandre C, Baena-Lopez A, Vincent JP. Patterning and growth control by membrane-tethered Wingless. Nature. 2014;505(7482):180-5.
37. Farin HF, Jordens I, Mosa MH, Basak O, Korving J, Tauriello DV, et al., Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. Nature. 2016;530(7590):340-3.
38. Massague J. TGFbeta signalling in context. Nat Rev Mol Cell Biol. 2012;13(10):616-30.
39. He XC, Zhang J, Tong WG, Tawfik O, Ross J, Scoville DH, et al., BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt-beta-catenin signaling. Nat Genet. 2004;36(10):1117-21.
40. Sheng H, Shao J, Townsend CM, Jr., Evers BM. Phosphatidylinositol 3-kinase mediates proliferative signals in intestinal epithelial cells. Gut. 2003;52(10):1472-8.
41. Bi L, Okabe I, Bernard DJ, Wynshaw-Boris A, Nussbaum RL. Proliferative defect and embryonic lethality in mice homozygous for a deletion in the p110alpha subunit of phosphoinositide 3-kinase. J Biol Chem. 1999;274(16):10963-8.
42. Morrison DK. MAP kinase pathways. Cold Spring Harb Perspect Biol. 2012;4(11).
43. Guo YJ, Pan WW, Liu SB, Shen ZF, Xu Y, Hu LL. ERK/MAPK signalling pathway and tumorigenesis. Exp Ther Med. 2020;19(3):1997-2007.
44. Ryan MB, Corcoran RB. Therapeutic strategies to target RAS-mutant cancers. Nat Rev Clin Oncol. 2018;15(11):709-20.
45. Yao Z, Yaeger R, Rodrik-Outmezguine VS, Tao A, Torres NM, Chang MT, et al., Tumours with class 3 BRAF mutants are sensitive to the inhibition of activated RAS. Nature. 2017;548(7666):234-8.
46. Koch U, Lehal R, Radtke F. Stem cells living with a Notch. Development. 2013;140(4):689-704.
47. Demitrack ES, Samuelson LC. Notch regulation of gastrointestinal stem cells. J Physiol. 2016;594(17):4791-803.
48. Sancho R, Cremona CA, Behrens A. Stem cell and progenitor fate in the mammalian intestine: Notch and lateral inhibition in homeostasis and disease. EMBO Rep. 2015;16(5):571-81.
49. Kim TH, Escudero S, Shivdasani RA. Intact function of Lgr5 receptor-expressing intestinal stem cells in the absence of Paneth cells. Proc Natl Acad Sci U S A. 2012;109(10):3932-7.
50. Durand A, Donahue B, Peignon G, Letourneur F, Cagnard N, Slomianny C, et al., Functional intestinal stem cells after Paneth cell ablation induced by the loss of transcription factor Math1 (Atoh1). Proc Natl Acad Sci U S A. 2012;109(23):8965-70.
51. Medema JP. Cancer stem cells: the challenges ahead. Nat Cell Biol. 2013;15(4):338-44.
52. Vogelstein ERFaB. A Genetic Model for Colorectal Tumorigenesis. Cell. 1990;61(759-67).

53. Malki A, ElRuz RA, Gupta I, Allouch A, Vranic S, Al Moustafa AE. Molecular Mechanisms of Colon Cancer Progression and Metastasis: Recent Insights and Advancements. Int J Mol Sci. 2020;22(1).
54. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. Gastroenterology. 2010;138(6):2059-72.
55. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(7407):330-7.
56. Worthley DL, Whitehall VL, Spring KJ, Leggett BA. Colorectal carcinogenesis: road maps to cancer. World J Gastroenterol. 2007;13(28):3784-91.
57. Garrido-Ramos MA. Satellite DNA: An Evolving Topic. Genes (Basel). 2017;8(9).
58. Li GM. Mechanisms and functions of DNA mismatch repair. Cell Res. 2008;18(1):85-98.
59. Markowitz S, Wang J, Myeroff L, Parsons R, Sun L, Lutterbaugh J, et al., Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science. 1995;268(5215):1336-8.
60. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, et al., A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res. 1998;58(22):5248-57.
61. Hegde M, Ferber M, Mao R, Samowitz W, Ganguly A, Working Group of the American College of Medical G, et al., ACMG technical standards and guidelines for genetic testing for inherited colorectal cancer (Lynch syndrome, familial adenomatous polyposis, and MYH-associated polyposis). Genet Med. 2014;16(1):101-16.
62. Wong JJ, Hawkins NJ, Ward RL. Colorectal cancer: a model for epigenetic tumorigenesis. Gut. 2007;56(1):140-8.
63. Magzoub MM, Prunello M, Brennan K, Gevaert O. The impact of DNA methylation on the cancer proteome. PLoS Comput Biol. 2019;15(7):e1007245.
64. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, et al., CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet. 2006;38(7):787-93.
65. Ogino S, Kawasaki T, Kirkner GJ, Suemoto Y, Meyerhardt JA, Fuchs CS. Molecular correlates with MGMT promoter methylation and silencing support CpG island methylator phenotype-low (CIMP-low) in colorectal cancer. Gut. 2007;56(11):1564-71.
66. Nosho K, Shima K, Irahara N, Kure S, Baba Y, Kirkner GJ, et al., DNMT3B expression might contribute to CpG island methylator phenotype in colorectal cancer. Clin Cancer Res. 2009;15(11):3663-71.
67. Ogino S, Brahmandam M, Kawasaki T, Kirkner GJ, Loda M, Fuchs CS. Combined analysis of COX-2 and p53 expressions reveals synergistic inverse correlations with microsatellite instability and CpG island methylator phenotype in colorectal cancer. Neoplasia. 2006;8(6):458-64.
68. Rowan AJ, Lamlum H, Ilyas M, Wheeler J, Straub J, Papadopoulou A, et al., APC mutations in sporadic colorectal tumors: A mutational "hotspot" and interdependence of the "two hits". Proc Natl Acad Sci U S A. 2000;97(7):3352-7.
69. Powell SM, Petersen GM, Krush AJ, Booker S, Jen J, Giardiello FM, et al., Molecular diagnosis of familial adenomatous polyposis. N Engl J Med. 1993;329(27):1982-7.
70. Carothers AM, Melstrom KA, Jr., Mueller JD, Weyant MJ, Bertagnolli MM. Progressive changes in adherens junction structure during intestinal adenoma formation in Apc mutant mice. J Biol Chem. 2001;276(42):39094-102.
71. Heinen CD, Goss KH, Cornelius JR, Babcock GF, Knudsen ES, Kowalik T, et al., The APC tumor suppressor controls entry into S-phase through its ability to regulate the cyclin D/RB pathway. Gastroenterology. 2002;123(3):751-63.
72. Jaiswal AS, Narayan S. Assembly of the base excision repair complex on abasic DNA and role of adenomatous polyposis coli on its functional activity. Biochemistry. 2011;50(11):1901-9.
73. Vigil D, Cherfils J, Rossman KL, Der CJ. Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? Nat Rev Cancer. 2010;10(12):842-57.
74. Van Cutsem E, Cervantes A, Nordlinger B, Arnold D, Group EGW. Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2014;25 Suppl 3:iii1-9.
75. Molina-Cerrillo J, San Roman M, Pozas J, Alonso-Gordoa T, Pozas M, Conde E, et al., BRAF Mutated Colorectal Cancer: New Treatment Approaches. Cancers (Basel). 2020;12(6).

76. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, et al., Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N Engl J Med. 2011;364(26):2507-16.
77. Kopetz S, Desai J, Chan E, Hecht JR, O'Dwyer PJ, Maru D, et al., Phase II Pilot Study of Vemurafenib in Patients With Metastatic BRAF-Mutated Colorectal Cancer. J Clin Oncol. 2015;33(34):4032-8.
78. Rowland A, Dias MM, Wiese MD, Kichenadasse G, McKinnon RA, Karapetis CS, et al., Meta-analysis of BRAF mutation as a predictive biomarker of benefit from anti-EGFR monoclonal antibody therapy for RAS wild-type metastatic colorectal cancer. Br J Cancer. 2015;112(12):1888-94.
79. Taieb J, Lapeyre-Prost A, Laurent Puig P, Zaanan A. Exploring the best treatment options for BRAF-mutant metastatic colon cancer. Br J Cancer. 2019;121(6):434-42.
80. Longacre TA, Fenoglio-Preiser CM. Mixed hyperplastic adenomatous polyps/serrated adenomas. A distinct form of colorectal neoplasia. Am J Surg Pathol. 1990;14(6):524-37.
81. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. Gastroenterology. 2010;138(6):2088-100.
82. Higuchi T, Sugihara K, Jass JR. Demographic and pathological characteristics of serrated polyps of colorectum. Histopathology. 2005;47(1):32-40.
83. Spring KJ, Zhao ZZ, Karamatic R, Walsh MD, Whitehall VL, Pike T, et al., High prevalence of sessile serrated adenomas with BRAF mutations: a prospective study of patients undergoing colonoscopy. Gastroenterology. 2006;131(5):1400-7.
84. Bettington M, Walker N, Clouston A, Brown I, Leggett B, Whitehall V. The serrated pathway to colorectal carcinoma: current concepts and challenges. Histopathology. 2013;62(3):367-86.
85. Campisi J. Suppressing cancer: the importance of being senescent. Science. 2005;309(5736):886-7.
86. Fukuyama R, Niculaita R, Ng KP, Obusez E, Sanchez J, Kalady M, et al., Mutated in colorectal cancer, a putative tumor suppressor for serrated colorectal cancer, selectively represses beta-catenin-dependent transcription. Oncogene. 2008;27(46):6044-55.
87. Kambara T, Simms LA, Whitehall VL, Spring KJ, Wynter CV, Walsh MD, et al., BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. Gut. 2004;53(8):1137-44.
88. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology. 2007;50(1):113-30.
89. Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, et al., The consensus molecular subtypes of colorectal cancer. Nat Med. 2015;21(11):1350-6.
90. Muller MF, Ibrahim AE, Arends MJ. Molecular pathological classification of colorectal cancer. Virchows Arch. 2016;469(2):125-34.
91. Fessler E, Medema JP. Colorectal Cancer Subtypes: Developmental Origin and Microenvironmental Regulation. Trends Cancer. 2016;2(9):505-18.
92. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al., Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012;366(10):883-92.
93. Litchfield K, Stanislaw S, Spain L, Gallegos LL, Rowan A, Schnidrig D, et al., Representative Sequencing: Unbiased Sampling of Solid Tumor Tissue. Cell Rep. 2020;31(5):107550.
94. Isella C, Brundu F, Bellomo SE, Galimi F, Zanella E, Porporato R, et al., Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. Nat Commun. 2017;8:15107.
95. Rosenberg DW, Giardina C, Tanaka T. Mouse models for the study of colon carcinogenesis. Carcinogenesis. 2009;30(2):183-96.
96. Burtin F, Mullins CS, Linnebacher M. Mouse models of colorectal cancer: Past, present and future perspectives. World J Gastroenterol. 2020;26(13):1394-426.
97. Moser AR, Pitot HC, Dove WF. A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. Science. 1990;247(4940):322-4.
98. Lakso M, Sauer B, Mosinger B, Jr., Lee EJ, Manning RW, Yu SH, et al., Targeted oncogene activation by site-specific recombination in transgenic mice. Proc Natl Acad Sci U S A. 1992;89(14):6232-6.
99. Madison BB, Dunbar L, Qiao XT, Braunstein K, Braunstein E, Gumucio DL. Cis elements of the villin gene control expression in restricted domains of the vertical (crypt) and horizontal (duodenum, cecum) axes of the intestine. J Biol Chem. 2002;277(36):33275-83.

100. el Marjou F, Janssen KP, Chang BH, Li M, Hindie V, Chan L, et al., Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. Genesis. 2004;39(3):186-93.
101. Collado M, Gil J, Efeyan A, Guerra C, Schuhmacher AJ, Barradas M, et al., Tumour biology: senescence in premalignant tumours. Nature. 2005;436(7051):642.
102. Pakneshan S, Salajegheh A, Smith RA, Lam AK. Clinicopathological relevance of BRAF mutations in human cancer. Pathology. 2013;45(4):346-56.
103. Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, et al., High frequency of mutations of the PIK3CA gene in human cancers. Science. 2004;304(5670):554.
104. Liu P, Cheng H, Santiago S, Raeder M, Zhang F, Isabella A, et al., Oncogenic PIK3CA-driven mammary tumors frequently recur via PI3K pathway-dependent and PI3K pathway-independent mechanisms. Nat Med. 2011;17(9):1116-20.
105. Engelman JA, Chen L, Tan X, Crosby K, Guimaraes AR, Upadhyay R, et al., Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. Nat Med. 2008;14(12):1351-6.
106. Eser S, Reiff N, Messer M, Seidler B, Gottschalk K, Dobler M, et al., Selective requirement of PI3K/PDK1 signaling for Kras oncogene-driven pancreatic cell plasticity and cancer. Cancer Cell. 2013;23(3):406-20.
107. Yueh AE, Payne SN, Leystra AA, Van De Hey DR, Foley TM, Pasch CA, et al., Colon Cancer Tumorigenesis Initiated by the H1047R Mutant PI3K. PLoS One. 2016;11(2):e0148730.
108. Feil S, Valtcheva N, Feil R. Inducible Cre mice. Methods Mol Biol. 2009;530:343-63.
109. Bresalier RS, Raper SE, Hujanen ES, Kim YS. A new animal model for human colon cancer metastasis. Int J Cancer. 1987;39(5):625-30.
110. Cespedes MV, Espina C, Garcia-Cabezas MA, Trias M, Boluda A, Gomez del Pulgar MT, et al., Orthotopic microinjection of human colon cancer cells in nude mice induces tumor foci in all clinically relevant metastatic sites. Am J Pathol. 2007;170(3):1077-85.
111. Fumagalli A, Drost J, Suijkerbuijk SJ, van Boxtel R, de Ligt J, Offerhaus GJ, et al., Genetic dissection of colorectal cancer progression by orthotopic transplantation of engineered cancer organoids. Proc Natl Acad Sci U S A. 2017;114(12):E2357-E64.
112. Fu XY, Besterman JM, Monosov A, Hoffman RM. Models of human metastatic colon cancer in nude mice orthotopically constructed by using histologically intact patient specimens. Proc Natl Acad Sci U S A. 1991;88(20):9345-9.
113. Donigan M, Norcross LS, Aversa J, Colon J, Smith J, Madero-Visbal R, et al., Novel murine model for colon cancer: non-operative trans-anal rectal injection. J Surg Res. 2009;154(2):299-303.
114. Zigmond E, Halpern Z, Elinav E, Brazowski E, Jung S, Varol C. Utilization of murine colonoscopy for orthotopic implantation of colorectal cancer. PLoS One. 2011;6(12):e28858.
115. Bettenworth D, Mucke MM, Schwegmann K, Faust A, Poremba C, Schafers M, et al., Endoscopy-guided orthotopic implantation of colorectal cancer cells results in metastatic colorectal cancer in mice. Clin Exp Metastasis. 2016;33(6):551-62.
116. O'Rourke KP, Loizou E, Livshits G, Schatoff EM, Baslan T, Manchado E, et al., Transplantation of engineered organoids enables rapid generation of metastatic mouse models of colorectal cancer. Nat Biotechnol. 2017;35(6):577-82.
117. Kishimoto H, Momiyama M, Aki R, Kimura H, Suetsugu A, Bouvet M, et al., Development of a clinically-precise mouse model of rectal cancer. PLoS One. 2013;8(11):e79453.
118. Hite N, Klinger A, Hellmers L, Maresh GA, Miller PE, Zhang X, et al., An Optimal Orthotopic Mouse Model for Human Colorectal Cancer Primary Tumor Growth and Spontaneous Metastasis. Dis Colon Rectum. 2018;61(6):698-705.
119. Drost J, Clevers H. Organoids in cancer research. Nat Rev Cancer. 2018;18(7):407-18.
120. Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, Stange DE, et al., Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature. 2009;459(7244):262-5.
121. Jung P, Sato T, Merlos-Suarez A, Barriga FM, Iglesias M, Rossell D, et al., Isolation and in vitro expansion of human colonic stem cells. Nat Med. 2011;17(10):1225-7.
122. Sato T, Stange DE, Ferrante M, Vries RG, Van Es JH, Van den Brink S, et al., Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology. 2011;141(5):1762-72.
123. Yin X, Farin HF, van Es JH, Clevers H, Langer R, Karp JM. Niche-independent high-purity cultures of Lgr5+ intestinal stem cells and their progeny. Nat Methods. 2014;11(1):106-12.
124. Huch M, Gehart H, van Boxtel R, Hamer K, Blokzijl F, Verstegen MM, et al., Long-term culture of genome-stable bipotent stem cells from adult human liver. Cell. 2015;160(1-2):299-312.

125. Boj SF, Hwang CI, Baker LA, Chio, II, Engle DD, Corbo V, et al., Organoid models of human and mouse ductal pancreatic cancer. Cell. 2015;160(1-2):324-38.
126. Bartfeld S, Bayram T, van de Wetering M, Huch M, Begthel H, Kujala P, et al., In vitro expansion of human gastric epithelial stem cells and their responses to bacterial infection. Gastroenterology. 2015;148(1):126-36 e6.
127. Rock JR, Onaitis MW, Rawlins EL, Lu Y, Clark CP, Xue Y, et al., Basal cells as stem cells of the mouse trachea and human airway epithelium. Proc Natl Acad Sci U S A. 2009;106(31):12771-5.
128. Sachs N, de Ligt J, Kopper O, Gogola E, Bounova G, Weeber F, et al., A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. Cell. 2018;172(1-2):373-86 e10.
129. van de Wetering M, Francies HE, Francis JM, Bounova G, Iorio F, Pronk A, et al., Prospective derivation of a living organoid biobank of colorectal cancer patients. Cell. 2015;161(4):933-45.
130. Vlachogiannis G, Hedayat S, Vatsiou A, Jamin Y, Fernandez-Mateos J, Khan K, et al., Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. Science. 2018;359(6378):920-6.
131. Fujii M, Shimokawa M, Date S, Takano A, Matano M, Nanki K, et al., A Colorectal Tumor Organoid Library Demonstrates Progressive Loss of Niche Factor Requirements during Tumorigenesis. Cell Stem Cell. 2016;18(6):827-38.
132. Roerink SF, Sasaki N, Lee-Six H, Young MD, Alexandrov LB, Behjati S, et al., Intra-tumour diversification in colorectal cancer at the single-cell level. Nature. 2018;556(7702):457-62.
133. Blokzijl F, de Ligt J, Jager M, Sasselli V, Roerink S, Sasaki N, et al., Tissue-specific mutation accumulation in human adult stem cells during life. Nature. 2016;538(7624):260-4.
134. Drost J, van Jaarsveld RH, Ponsioen B, Zimberlin C, van Boxtel R, Buijs A, et al., Sequential cancer mutations in cultured human intestinal stem cells. Nature. 2015;521(7550):43-7.
135. Matano M, Date S, Shimokawa M, Takano A, Fujii M, Ohta Y, et al., Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. Nat Med. 2015;21(3):256-62.
136. de Sousa e Melo F, Kurtova AV, Harnoss JM, Kljavin N, Hoeck JD, Hung J, et al., A distinct role for Lgr5(+) stem cells in primary and metastatic colon cancer. Nature. 2017;543(7647):676-80.
137. Sakamoto N, Feng Y, Stolfi C, Kurosu Y, Green M, Lin J, et al., BRAF(V600E) cooperates with CDX2 inactivation to promote serrated colorectal tumorigenesis. Elife. 2017;6.
138. Fessler E, Drost J, van Hooff SR, Linnekamp JF, Wang X, Jansen M, et al., TGFbeta signaling directs serrated adenomas to the mesenchymal colorectal cancer subtype. EMBO Mol Med. 2016;8(7):745-60.
139. Yan HHN, Lai JCW, Ho SL, Leung WK, Law WL, Lee JFY, et al., RNF43 germline and somatic mutation in serrated neoplasia pathway and its association with BRAF mutation. Gut. 2017;66(9):1645-56.
140. Seino T, Kawasaki S, Shimokawa M, Tamagawa H, Toshimitsu K, Fujii M, et al., Human Pancreatic Tumor Organoids Reveal Loss of Stem Cell Niche Factor Dependence during Disease Progression. Cell Stem Cell. 2018;22(3):454-67 e6.
141. Cattaneo CM, Dijkstra KK, Fanchi LF, Kelderman S, Kaing S, van Rooij N, et al., Tumor organoid-T-cell coculture systems. Nat Protoc. 2020;15(1):15-39.
142. Pleguezuelos-Manzano C, Puschhof J, Rosendahl Huber A, van Hoeck A, Wood HM, Nomburg J, et al., Mutational signature in colorectal cancer caused by genotoxic pks(+) E. coli. Nature. 2020;580(7802):269-73.
143. Hubert CG, Rivera M, Spangler LC, Wu Q, Mack SC, Prager BC, et al., A Three-Dimensional Organoid Culture System Derived from Human Glioblastomas Recapitulates the Hypoxic Gradients and Cancer Stem Cell Heterogeneity of Tumors Found In Vivo. Cancer Res. 2016;76(8):2465-77.
144. Krotenberg Garcia A, Fumagalli A, Le HQ, Jackstadt R, Lannagan TRM, Sansom OJ, et al., Active elimination of intestinal cells drives oncogenic growth in organoids. Cell Rep. 2021;36(1):109307.
145. Nowell PC. The clonal evolution of tumor cell populations. Science. 1976;194(4260):23-8.
146. Saito T, Niida A, Uchi R, Hirata H, Komatsu H, Sakimura S, et al., A temporal shift of the evolutionary principle shaping intratumor heterogeneity in colorectal cancer. Nat Commun. 2018;9(1):2884.
147. Uchi R, Takahashi Y, Niida A, Shimamura T, Hirata H, Sugimachi K, et al., Integrated Multiregional Analysis Proposing a New Model of Colorectal Cancer Evolution. PLoS Genet. 2016;12(2):e1005778.

148. Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, et al., A Big Bang model of human colorectal tumor growth. Nat Genet. 2015;47(3):209-16.
149. Cross W, Kovac M, Mustonen V, Temko D, Davis H, Baker AM, et al., The evolutionary landscape of colorectal tumorigenesis. Nat Ecol Evol. 2018;2(10):1661-72.
150. Niida A, Hasegawa T, Innan H, Shibata T, Mimori K, Miyano S. A unified simulation model for understanding the diversity of cancer evolution. PeerJ. 2020;8:e8842.
151. Niida A, Mimori K, Shibata T, Miyano S. Modeling colorectal cancer evolution. J Hum Genet. 2021;66(9):869-78.
152. Mueller S, Engleitner T, Maresch R, Zukowska M, Lange S, Kaltenbacher T, et al., Evolutionary routes and KRAS dosage define pancreatic cancer phenotypes. Nature. 2018;554(7690):62-8.
153. Sarkisian CJ, Keister BA, Stairs DB, Boxer RB, Moody SE, Chodosh LA. Dose-dependent oncogene-induced senescence in vivo and its evasion during mammary tumorigenesis. Nat Cell Biol. 2007;9(5):493-505.
154. Jackson EL, Willis N, Mercer K, Bronson RT, Crowley D, Montoya R, et al., Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. Genes Dev. 2001;15(24):3243-8.
155. Nagtegaal ID, Odze RD, Klimstra D, Paradis V, Rugge M, Schirmacher P, et al., The 2019 WHO classification of tumours of the digestive system. Histopathology. 2020;76(2):182-8.
156. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018;28(11):1747-56.
157. Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, et al., CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol. 2019;37(3):224-6.
158. Lange S, Engleitner T, Mueller S, Maresch R, Zwiebel M, Gonzalez-Silva L, et al., Analysis pipelines for cancer genome sequencing in mice. Nat Protoc. 2020;15(2):266-315.
159. Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I. The impact of amplification on differential expression analyses by RNA-seq. Sci Rep. 2016;6:25533.
160. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al., Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015;161(5):1202-14.
161. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
162. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417-25.
163. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014;11(8):783-4.
164. Roper J, Tammela T, Akkad A, Almeqdadi M, Santos SB, Jacks T, et al., Colonoscopy-based colorectal cancer modeling in mice with CRISPR-Cas9 genome editing and organoid transplantation. Nat Protoc. 2018;13(2):217-34.
165. Pfarr N, Penzel R, Endris V, Lier C, Flechtenmacher C, Volckmar AL, et al., Targeted next-generation sequencing enables reliable detection of HER2 (ERBB2) status in breast cancer and provides ancillary information of clinical relevance. Genes Chromosomes Cancer. 2017;56(4):255-65.
166. Endris V, Penzel R, Warth A, Muckenhuber A, Schirmacher P, Stenzinger A, et al., Molecular diagnostic profiling of lung cancer specimens with a semiconductor-based massive parallel sequencing approach: feasibility, costs, and performance compared with conventional sequencing. J Mol Diagn. 2013;15(6):765-75.
167. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-67.
168. Barsouk A, Rawla P, Barsouk A, Thandra KC. Epidemiology of Cancers of the Small Intestine: Trends, Risk Factors, and Prevention. Med Sci (Basel). 2019;7(3).
169. Arnold A, Tronser M, Sers C, Ahadova A, Endris V, Mamlouk S, et al., The majority of beta-catenin mutations in colorectal cancer is homozygous. BMC Cancer. 2020;20(1):1038.
170. Lemieux E, Cagnol S, Beaudry K, Carrier J, Rivard N. Oncogenic KRAS signalling promotes the Wnt/beta-catenin pathway through LRP6 in colorectal cancer. Oncogene. 2015;34(38):4914-27.
171. Cai J, Feng D, Hu L, Chen H, Yang G, Cai Q, et al., FAT4 functions as a tumour suppressor in gastric cancer by modulating Wnt/beta-catenin signalling. Br J Cancer. 2015;113(12):1720-9.
172. Nikolaev SI, Sotiriou SK, Pateras IS, Santoni F, Sougioultzis S, Edgren H, et al., A single-nucleotide substitution mutator phenotype revealed by exome sequencing of human colon adenomas. Cancer Res. 2012;72(23):6279-89.

173. Oh JH, Jang SJ, Kim J, Sohn I, Lee JY, Cho EJ, et al., Spontaneous mutations in the single TTN gene represent high tumor mutation burden. NPJ Genom Med. 2020;5:33.
174. Alcantara Torres M, Rodriguez Merlo R, Repiso Ortega A, de Lucas Veguillas A, Valle Munoz J, Sanchez Simon R, et al., DNA aneuploidy in colorectal adenomas. Role in the adenoma-carcinoma sequence. Rev Esp Enferm Dig. 2005;97(1):7-15.
175. Pillaire MJ, Selves J, Gordien K, Gourraud PA, Gentil C, Danjoux M, et al., A 'DNA replication' signature of progression and negative outcome in colorectal cancer. Oncogene. 2010;29(6):876-87.
176. Diep CB, Kleivi K, Ribeiro FR, Teixeira MR, Lindgjaerde OC, Lothe RA. The order of genetic events associated with colorectal cancer progression inferred from meta-analysis of copy number changes. Genes Chromosomes Cancer. 2006;45(1):31-41.
177. Favazza LA, Parseghian CM, Kaya C, Nikiforova MN, Roy S, Wald AI, et al., KRAS amplification in metastatic colon cancer is associated with a history of inflammatory bowel disease and may confer resistance to anti-EGFR therapy. Mod Pathol. 2020;33(9):1832-43.
178. Valtorta E, Misale S, Sartore-Bianchi A, Nagtegaal ID, Paraf F, Lauricella C, et al., KRAS gene amplification in colorectal cancer and impact on response to EGFR-targeted therapy. Int J Cancer. 2013;133(5):1259-65.
179. Boretto M, Cox B, Noben M, Hendriks N, Fassbender A, Roose H, et al., Development of organoids from mouse and human endometrium showing endometrial epithelium physiology and long-term expandability. Development. 2017;144(10):1775-86.
180. Weiss L, Grundmann E, Torhorst J, Hartveit F, Moberg I, Eder M, et al., Haematogenous metastatic patterns in colonic carcinoma: an analysis of 1541 necropsies. J Pathol. 1986;150(3):195-203.
181. Serrano M, Lin AW, McCurrach ME, Beach D, Lowe SW. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell. 1997;88(5):593-602.
182. Ferbeyre G, de Stanchina E, Lin AW, Querido E, McCurrach ME, Hannon GJ, et al., Oncogenic ras and p53 cooperate to induce cellular senescence. Mol Cell Biol. 2002;22(10):3497-508.
183. Aubrey BJ, Kelly GL, Janic A, Herold MJ, Strasser A. How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? Cell Death Differ. 2018;25(1):104-13.
184. Wajant H, Pfizenmaier K, Scheurich P. Tumor necrosis factor signaling. Cell Death Differ. 2003;10(1):45-65.
185. Sparmann A, Bar-Sagi D. Ras-induced interleukin-8 expression plays a critical role in tumor growth and angiogenesis. Cancer Cell. 2004;6(5):447-58.
186. Zhu Z, Aref AR, Cohoon TJ, Barbie TU, Imamura Y, Yang S, et al., Inhibition of KRAS-driven tumorigenicity by interruption of an autocrine cytokine circuit. Cancer Discov. 2014;4(4):452-65.
187. Ling J, Kang Y, Zhao R, Xia Q, Lee DF, Chang Z, et al., KrasG12D-induced IKK2/beta/NF-kappaB activation by IL-1alpha and p62 feedforward loops is required for development of pancreatic ductal adenocarcinoma. Cancer Cell. 2012;21(1):105-20.
188. Griffith JW, Sokol CL, Luster AD. Chemokines and chemokine receptors: positioning cells for host defense and immunity. Annu Rev Immunol. 2014;32:659-702.
189. Perillo B, Di Donato M, Pezone A, Di Zazzo E, Giovannelli P, Galasso G, et al., ROS in cancer therapy: the bright side of the moon. Exp Mol Med. 2020;52(2):192-203.
190. Ashton TM, McKenna WG, Kunz-Schughart LA, Higgins GS. Oxidative Phosphorylation as an Emerging Target in Cancer Therapy. Clin Cancer Res. 2018;24(11):2482-90.
191. Minn AJ, Wherry EJ. Combination Cancer Therapies with Immune Checkpoint Blockade: Convergence on Interferon Signaling. Cell. 2016;165(2):272-5.
192. Du W, Frankel TL, Green M, Zou W. IFNgamma signaling integrity in colorectal cancer immunity and immunotherapy. Cell Mol Immunol. 2021.
193. Andreu P, Colnot S, Godard C, Laurent-Puig P, Lamarque D, Kahn A, et al., Identification of the IFITM family as a new molecular marker in human colorectal tumors. Cancer Res. 2006;66(4):1949-55.
194. Nan J, Wang Y, Yang J, Stark GR. IRF9 and unphosphorylated STAT2 cooperate with NF-kappaB to drive IL6 expression. Proc Natl Acad Sci U S A. 2018;115(15):3906-11.
195. McDonald SL, Silver A. The opposing roles of Wnt-5a in cancer. Br J Cancer. 2009;101(2):209-14.
196. Kikuchi A. Tumor formation by genetic mutations in the components of the Wnt signaling pathway. Cancer Sci. 2003;94(3):225-9.

197. Morin PJ, Sparks AB, Korinek V, Barker N, Clevers H, Vogelstein B, et al., Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. Science. 1997;275(5307):1787-90.
198. Gonzalez RS, Huh WJ, Cates JM, Washington K, Beauchamp RD, Coffey RJ, et al., Micropapillary colorectal carcinoma: clinical, pathological and molecular properties, including evidence of epithelial-mesenchymal transition. Histopathology. 2017;70(2):223-31.
199. Jakubowska K, Guzinska-Ustymowicz K, Pryczynicz A. Invasive micropapillary component and its clinico-histopathological significance in patients with colorectal cancer. Oncol Lett. 2016;12(2):1154-8.
200. Friedman K, Brodsky AS, Lu S, Wood S, Gill AJ, Lombardo K, et al., Medullary carcinoma of the colon: a distinct morphology reveals a distinctive immunoregulatory microenvironment. Mod Pathol. 2016;29(5):528-41.
201. Henderson RH, French D, Maughan T, Adams R, Allemani C, Minicozzi P, et al., The economic burden of colorectal cancer across Europe: a population-based cost-of-illness study. Lancet Gastroenterol Hepatol. 2021;6(9):709-22.
202. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, et al., Colorectal cancer statistics, 2017. CA Cancer J Clin. 2017;67(3):177-93.
203. De Palma FDE, D'Argenio V, Pol J, Kroemer G, Maiuri MC, Salvatore F. The Molecular Hallmarks of the Serrated Pathway in Colorectal Cancer. Cancers (Basel). 2019;11(7).
204. Weeber F, Ooft SN, Dijkstra KK, Voest EE. Tumor Organoids as a Pre-clinical Cancer Model for Drug Discovery. Cell Chem Biol. 2017;24(9):1092-100.
205. Hoevenaar WHM, Janssen A, Quirindongo AI, Ma H, Klaasen SJ, Teixeira A, et al., Degree and site of chromosomal instability define its oncogenic potential. Nat Commun. 2020;11(1):1501.
206. Kim SE, Choi KY. EGF receptor is involved in WNT3a-mediated proliferation and motility of NIH3T3 cells via ERK pathway activation. Cell Signal. 2007;19(7):1554-64.
207. Park KS, Jeon SH, Kim SE, Bahk YY, Holmen SL, Williams BO, et al., APC inhibits ERK pathway activation and cellular proliferation induced by RAS. J Cell Sci. 2006;119(Pt 5):819-27.
208. Jeong WJ, Yoon J, Park JC, Lee SH, Lee SH, Kaduwal S, et al., Ras stabilization through aberrant activation of Wnt/beta-catenin signaling promotes intestinal tumorigenesis. Sci Signal. 2012;5(219):ra30.
209. Hwang JH, Yoon J, Cho YH, Cha PH, Park JC, Choi KY. A mutant KRAS-induced factor REG4 promotes cancer stem cell properties via Wnt/beta-catenin signaling. Int J Cancer. 2020;146(10):2877-90.
210. Russo AA, Tong L, Lee JO, Jeffrey PD, Pavletich NP. Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. Nature. 1998;395(6699):237-43.
211. Bihani T, Mason DX, Jackson TJ, Chen SC, Boettner B, Lin AW. Differential oncogenic Ras signaling and senescence in tumor cells. Cell Cycle. 2004;3(9):1201-7.
212. Lundberg AS, Hahn WC, Gupta P, Weinberg RA. Genes involved in senescence and immortalization. Curr Opin Cell Biol. 2000;12(6):705-9.
213. Volonte D, Sedorovitz M, Cespedes VE, Beecher ML, Galbiati F. Cell autonomous angiotensin II signaling controls the pleiotropic functions of oncogenic K-Ras. J Biol Chem. 2021;296:100242.
214. Kamijo T, Zindy F, Roussel MF, Quelle DE, Downing JR, Ashmun RA, et al., Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. Cell. 1997;91(5):649-59.
215. Carragher LA, Snell KR, Giblett SM, Aldridge VS, Patel B, Cook SJ, et al., V600EBraf induces gastrointestinal crypt senescence and promotes tumour progression through enhanced CpG methylation of p16INK4a. EMBO Mol Med. 2010;2(11):458-71.
216. Reischmann N, Andrieux G, Griffin R, Reinheckel T, Boerries M, Brummer T. BRAF(V600E) drives dedifferentiation in small intestinal and colonic organoids and cooperates with mutant p53 and Apc loss in transformation. Oncogene. 2020;39(38):6053-70.
217. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, et al., Mutations of the BRAF gene in human cancer. Nature. 2002;417(6892):949-54.
218. Gopal P, Sarihan EI, Chie EK, Kuzmishin G, Doken S, Pennell NA, et al., Clonal selection confers distinct evolutionary trajectories in BRAF-driven cancers. Nat Commun. 2019;10(1):5143.
219. Xue Y, Martelotto L, Baslan T, Vides A, Solomon M, Mai TT, et al., An approach to suppress the evolution of resistance in BRAF(V600E)-mutant cancer. Nat Med. 2017;23(8):929-37.

220. Bielski CM, Donoghue MTA, Gadiya M, Hanrahan AJ, Won HH, Chang MT, et al., Widespread Selection for Oncogenic Mutant Allele Imbalance in Cancer. Cancer Cell. 2018;34(5):852-62 e4.
221. Kato S, Iida S, Higuchi T, Ishikawa T, Takagi Y, Yasuno M, et al., PIK3CA mutation is predictive of poor survival in patients with colorectal cancer. Int J Cancer. 2007;121(8):1771-8.
222. Sartore-Bianchi A, Martini M, Molinari F, Veronese S, Nichelatti M, Artale S, et al., PIK3CA mutations in colorectal cancer are associated with clinical resistance to EGFR-targeted monoclonal antibodies. Cancer Res. 2009;69(5):1851-7.
223. Therkildsen C, Bergmann TK, Henrichsen-Schnack T, Ladelund S, Nilbert M. The predictive value of KRAS, NRAS, BRAF, PIK3CA and PTEN for anti-EGFR treatment in metastatic colorectal cancer: A systematic review and meta-analysis. Acta Oncol. 2014;53(7):852-64.
224. Wu S, Gan Y, Wang X, Liu J, Li M, Tang Y. PIK3CA mutation is associated with poor survival among patients with metastatic colorectal cancer following anti-EGFR monoclonal antibody therapy: a meta-analysis. J Cancer Res Clin Oncol. 2013;139(5):891-900.
225. Madsen RR, Knox RG, Pearce W, Lopez S, Mahler-Araujo B, McGranahan N, et al., Oncogenic PIK3CA promotes cellular stemness in an allele dose-dependent manner. Proc Natl Acad Sci U S A. 2019;116(17):8380-9.
226. Sampaziotis F, Muraro D, Tysoe OC, Sawiak S, Beach TE, Godfrey EM, et al., Cholangiocyte organoids can repair bile ducts after transplantation in the human liver. Science. 2021;371(6531):839-46.
227. Okamoto R, Shimizu H, Suzuki K, Kawamoto A, Takahashi J, Kawai M, et al., Organoid-based regenerative medicine for inflammatory bowel disease. Regen Ther. 2020;13:1-6.
228. Nakayama M, Hong CP, Oshima H, Sakai E, Kim SJ, Oshima M. Loss of wild-type p53 promotes mutant p53-driven metastasis through acquisition of survival and tumor-initiating properties. Nat Commun. 2020;11(1):2333.
229. Redman-Rivera LN, Shaver TM, Jin H, Marshall CB, Schafer JM, Sheng Q, et al., Acquisition of aneuploidy drives mutant p53-associated gain-of-function phenotypes. Nat Commun. 2021;12(1):5184.
230. Liu J, Cho YB, Hong HK, Wu S, Ebert PJ, Bray SM, et al., Molecular dissection of CRC primary tumors and their matched liver metastases reveals critical role of immune microenvironment, EMT and angiogenesis in cancer metastasis. Sci Rep. 2020;10(1):10725.
231. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009;119(6):1420-8.
232. Voss H, Wurlitzer M, Smit DJ, Ewald F, Alawi M, Spohn M, et al., Differential regulation of extracellular matrix proteins in three recurrent liver metastases of a single patient with colorectal cancer. Clin Exp Metastasis. 2020;37(6):649-56.
233. Hamarsheh S, Gross O, Brummer T, Zeiser R. Immune modulatory effects of oncogenic KRAS in cancer. Nat Commun. 2020;11(1):5439.
234. Brooks GD, McLeod L, Alhayyani S, Miller A, Russell PA, Ferlin W, et al., IL6 Trans-signaling Promotes KRAS-Driven Lung Carcinogenesis. Cancer Res. 2016;76(4):866-76.
235. Lesina M, Kurkowski MU, Ludes K, Rose-John S, Treiber M, Kloppel G, et al., Stat3/Socs3 activation by IL-6 transsignaling promotes progression of pancreatic intraepithelial neoplasia and development of pancreatic cancer. Cancer Cell. 2011;19(4):456-69.
236. Morris KT, Khan H, Ahmad A, Weston LL, Nofchissey RA, Pinchuk IV, et al., G-CSF and G-CSFR are highly expressed in human gastric and colon cancers and promote carcinoma cell proliferation and migration. Br J Cancer. 2014;110(5):1211-20.
237. Park J, Wysocki RW, Amoozgar Z, Maiorino L, Fein MR, Jorns J, et al., Cancer cells induce metastasis-supporting neutrophil extracellular DNA traps. Sci Transl Med. 2016;8(361):361ra138.
238. Oladipo O, Conlon S, O'Grady A, Purcell C, Wilson C, Maxwell PJ, et al., The expression and prognostic impact of CXC-chemokines in stage II and III colorectal cancer epithelial and stromal tissue. Br J Cancer. 2011;104(3):480-7.
239. Yang J, Li Y-N, Pan T, Miao R-R, Zhang Y-Y, Wu S-H, et al., Atypical chemokine receptor 3 (ACKR3) induces the perturbation of rRNA biogenesis: a novel mechanism of colorectal tumorigenesis. 2021.
240. Liao W, Overman MJ, Boutin AT, Shang X, Zhao D, Dey P, et al., KRAS-IRF2 Axis Drives Immune Suppression and Immune Therapy Resistance in Colorectal Cancer. Cancer Cell. 2019;35(4):559-72 e7.
241. Germann M, Zangger N, Sauvain MO, Sempoux C, Bowler AD, Wirapati P, et al., Neutrophils suppress tumor-infiltrating T cells in colon cancer via matrix metalloproteinase-mediated activation of TGFbeta. EMBO Mol Med. 2020;12(1):e10681.

242. Riemer P, Sreekumar A, Reinke S, Rad R, Schafer R, Sers C, et al., Transgenic expression of oncogenic BRAF induces loss of stem cells in the mouse intestine, which is antagonized by beta-catenin activity. Oncogene. 2015;34(24):3164-75.
243. Finch AJ, Soucek L, Junttila MR, Swigart LB, Evan GI. Acute overexpression of Myc in intestinal epithelium recapitulates some but not all the changes elicited by Wnt/beta-catenin pathway activation. Mol Cell Biol. 2009;29(19):5306-15.
244. Rochlitz CF, Herrmann R, de Kant E. Overexpression and amplification of c-myc during progression of human colorectal cancer. Oncology. 1996;53(6):448-54.
245. Juan J, Muraguchi T, Iezza G, Sears RC, McMahon M. Diminished WNT -> beta-catenin -> c-MYC signaling is a barrier for malignant progression of BRAFV600E-induced lung tumors. Genes Dev. 2014;28(6):561-75.
246. Bonaldi E, Gargiuli C, De Cecco L, Micali A, Rizzetti MG, Greco A, et al., BRAF Inhibitors Induce Feedback Activation of RAS Pathway in Thyroid Cancer Cells. Int J Mol Sci. 2021;22(11).
247. Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. Cancer Cell. 2015;27(4):450-61.
248. Asem MS, Buechler S, Wates RB, Miller DL, Stack MS. Wnt5a Signaling in Cancer. Cancers (Basel). 2016;8(9).
249. Zhao Y, Zhang C, Huang Y, Yu Y, Li R, Li M, et al., Up-regulated expression of WNT5a increases inflammation and oxidative stress via PI3K/AKT/NF-kappaB signaling in the granulosa cells of PCOS patients. J Clin Endocrinol Metab. 2015;100(1):201-11.
250. Zhang A, He S, Sun X, Ding L, Bao X, Wang N. Wnt5a promotes migration of human osteosarcoma cells by triggering a phosphatidylinositol-3 kinase/Akt signals. Cancer Cell Int. 2014;14(1):15.
251. Liu J, Zhang Y, Xu R, Du J, Hu Z, Yang L, et al., PI3K/Akt-dependent phosphorylation of GSK3beta and activation of RhoA regulate Wnt5a-induced gastric cancer cell migration. Cell Signal. 2013;25(2):447-56.
252. Chockley PJ, Keshamouni VG. Immunological Consequences of Epithelial-Mesenchymal Transition in Tumor Progression. J Immunol. 2016;197(3):691-8.
253. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, et al., Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science. 2006;313(5795):1960-4.
254. Vallania F, Tam A, Lofgren S, Schaffert S, Azad TD, Bongen E, et al., Leveraging heterogeneity across multiple datasets increases cell-mixture deconvolution accuracy and reduces biological and technical biases. Nat Commun. 2018;9(1):4735.
255. Sturm G, Finotello F, Petitprez F, Zhang JD, Baumbach J, Fridman WH, et al., Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. Bioinformatics. 2019;35(14):i436-i45.
256. Gargiulo G. Next-Generation in vivo Modeling of Human Cancers. Front Oncol. 2018;8:429.
257. Squatrito M, Holland EC. DNA damage response and growth factor signaling pathways in gliomagenesis and therapeutic resistance. Cancer Res. 2011;71(18):5945-9.
258. Ben-David U, Ha G, Tseng YY, Greenwald NF, Oh C, Shih J, et al., Patient-derived xenografts undergo mouse-specific tumor evolution. Nat Genet. 2017;49(11):1567-75.
259. Borrell B. How accurate are cancer cell lines? Nature. 2010;463(7283):858.
260. Hidalgo M, Bruckheimer E, Rajeshkumar NV, Garrido-Laguna I, De Oliveira E, Rubio-Viqueira B, et al., A pilot clinical study of treatment guided by personalized tumorgrafts in patients with advanced cancer. Mol Cancer Ther. 2011;10(8):1311-6.
261. Zhang M, Hutter G, Kahn SA, Azad TD, Gholamin S, Xu CY, et al., Anti-CD47 Treatment Stimulates Phagocytosis of Glioblastoma by M1 and M2 Polarized Macrophages and Promotes M1 Polarized Macrophages In Vivo. PLoS One. 2016;11(4):e0153550.
262. Theocharides AP, Jin L, Cheng PY, Prasolava TK, Malko AV, Ho JM, et al., Disruption of SIRPalpha signaling in macrophages eliminates human acute myeloid leukemia stem cells in xenografts. J Exp Med. 2012;209(10):1883-99.
263. Yao Y, Xu X, Yang L, Zhu J, Wan J, Shen L, et al., Patient-Derived Organoids Predict Chemoradiation Responses of Locally Advanced Rectal Cancer. Cell Stem Cell. 2020;26(1):17-26 e6.
264. Ganesh K, Wu C, O'Rourke KP, Szeglin BC, Zheng Y, Sauve CG, et al., A rectal cancer organoid platform to study individual responses to chemoradiation. Nat Med. 2019;25(10):1607-14.
265. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.

266. Gier RA, Budinich KA, Evitt NH, Cao Z, Freilich ES, Chen Q, et al., High-performance CRISPR-Cas12a genome editing for combinatorial genetic screening. Nat Commun. 2020;11(1):3455.