

# Intestinal stem cells differentiation and its application in cancer therapy

Présentée le 10 décembre 2021

Faculté des sciences de la vie  
Unité du Prof. Huelsken  
Programme doctoral en approches moléculaires du vivant

pour l'obtention du grade de Docteur ès Sciences

par

**Maxim NORKIN**

Acceptée sur proposition du jury

Prof. D. Constam, président du jury  
Prof. J. Hülsken, directeur de thèse  
Prof. T. Petrova, rapporteuse  
Prof. P. Nowak-Sliwinska, rapporteuse  
Prof. F. Radtke, rapporteur  
Prof. J. P. Medema, rapporteur

# Acknowledgements

First of all, I would like to thank my thesis supervisor, Joerg for his great support and advice during the project. Joerg always had great ideas on how to overcome problems in difficult times. He was always giving me experimental freedom but was there when I needed some advice. He would stay with me until the latest moments of our deadlines for paper submissions, grant applications, and administrative paperwork when needed. Also, thanks for inviting all of PhD students and co-workers to his place where he would arrange a warm welcome and tasty barbecues.

Secondly, I would like to thank my friends and colleagues – Angela, Elena, Fanny, Laurent, Luisa, Mathieu, Megan and Pierre for the very supportive and fun atmosphere in the lab, for the great lunches and coffees we had together, for the nice picnics, some hiking and skiing together. I just wish there would have been more of that. Also, I would like to thank all of you for the fruitful discussions and constructive comments during the lab meetings and personal meetings.

Further, I would like to thank a former postdoc Paloma for her support during the project, advices she gave to me on how to manipulate organoids, thoughtful discussions about cancer stem cells. I would also like to thank her for the opportunities she gave me to be a part of her work that resulted in several joint publications. I really enjoyed working with you and looking forward to further cooperation and joint projects.

I would like to thank my family, my mother Ludmila, and my sister Marina who have enormous belief in me and always supported me in difficult times, trying to cheer me up when I had some problems and giving me important life advice. It's always a pleasure to return home when you know you are always welcome, and people are waiting for you. I wish my father would also be here to see this great moment of my life and see me defending my PhD thesis.

Further, I would like to thank my big Russian community friends. Special thanks to Emil and Irina and Timur for taking me to lots of nice activities we did together: adventurous walks in the forests on sunny or rainy weather, difficult and intermediate hikes with an obligatory stop for a meal with tea; many skiing days and weekends we spend together; playing tennis and even going to different countries together. Thanks for inviting me over for dinners, it's always a pleasure to have long discussions about the current standing of the things in our world. Further, I would like to thank my other big group of Russian-speaking friends Maiia, Natalia, Anna R., Vitalijs, Marina, Adele for having wonderful lunches at EPFL at the start of my PhD and before pandemic times



which hit us hard. Further I would extend my thanks to Nastya (Lipa), Anna V., Denis, Lisa and Viktor for being such a nice atmospheric company to spend time together lately.

I would like to thank my PhD friends and colleagues in the life science department: Amber, Mateusz, Patrik, Aspasia, Anita, Katarina, Fabio, Nadia, Jelena, Daniel, Gerrard for good talks, coffee breaks, scientific discussions. I would also like to thank Julius, Ophélie and Jenny, my bouldering partners at EPFL for having nice times together at TOTEM.

I also want to thank my previous colleagues and friends from Institute of Molecular Biology in Mainz (Germany) - Jury, Lira, Margarita, Ekaterina, Arthur, Fernando, Mariangela - for their continuous support and our friendship through the time of my PhD in Lausanne, whether we met in Lausanne, Mainz, Groningen, Moscow or via zoom.

I would also like to thank the members of my candidacy exam and my PhD committee: Profs. D. Constam, F. Radtke, T. Halazonetis, T. Petrova, P. Nowak-Sliwinska, J.P. Medema. It's a great honor to have you on my committee. Thanks for advice and comments on my project. I would also like to thank Prof. P. Goenczy for being my mentor, for his great support and advice during my PhD lifetime.

Finally, I would like to thank the core facilities at EPFL: GECF and Bastien for their help in my project; FCCF, HCF, BIOP and BCF for their help during various parts of the project.

# Abstract

Colorectal cancer (CRC) is one of the deadliest cancers worldwide. Despite extensive research in the field, new therapies targeting CRC are in urgent need. CRC is a highly heterogeneous disease with high levels of intra- and inter-patient heterogeneity. Cancer stem cells are the driving force of tumor development and metastasis. One of the approaches to treat CRC is differentiation therapy whereby induction of cancer stem cell differentiation may lead to tumor regression. Intestinal normal and cancer organoids represent an ideal model for studying the differentiation process *ex vivo*. Gene expression profiling of drug-treated organoids can be used for quantification of the differentiation process with high precision. RNA-seq based drug-screening platforms in organoids are not yet established. Here we developed a high-throughput high-content targeted RNA-seq-based platform (TORNADO-seq) for monitoring gene expression of large gene signatures in intestinal organoids which allowed us to quantify and evaluate complex cellular phenotypes and cellular alterations upon drug treatment. We applied TORNADO-seq in drug screening in wt and cancer intestinal organoids. We found many differentiation-inducing drugs in wt organoids, including some drugs specifically enriching for certain cell phenotypes, e.g. enteroendocrine cells. Further, we applied identified differentiation-inducing drugs to cancer organoids ( $APC^{\text{lof}}$ ;  $KRAS^{\text{G12D}}$ ;  $TP53^{\text{lof}}$ ) and discovered that the majority of those differentiation-inducing drugs also targeted cancer organoids. Differentiation induction seems to be a common mode of action of cancer-targeting drugs. Based on gene expression profiles of treated cancer organoids, we were able to propose and confirm the mechanism of action of several drugs indicating that cancer cells may be sensitive to alterations in the cholesterol biosynthesis pathway. In general, utilization of TORNADO-seq allows predicting connections between cellular phenotypes and signaling pathways in organoids which can be used for hypothesis generation in a variety of biological systems.

The second part of the thesis is devoted to the investigation of mechanisms underlying the accumulation of genetic alterations in CRC tumors. It is not known whether cancer driver genes affect the rate of mutation acquisition. Here we performed exome sequencing of single cell-derived organoids obtained from tumors with either  $APC^{\text{lof}}$ ,  $APC^{\text{lof}}$ ;  $KRAS^{\text{G12D}}$  or  $APC^{\text{lof}}$ ;  $KRAS^{\text{G12D}}$ ;  $TP53^{\text{lof}}$  mutation phenotypes that reflect early stages of CRC development. We discovered that presence of KRAS and P53 mutations didn't affect the number of single nucleotide substitutions (SNSs) in tumors, while the

latter was increased in APC mutated tumors compared to wt cells and was highly correlated with the tumor age. P53 inactivation was highly correlated with the presence of large-scale (>10 Mb) copy number alterations (CNAs).

Keywords: colorectal cancer; differentiation therapy; high-throughput screening; high-content screening; organoids; drug screening; RNA-seq.

# Résumé

Le cancer colorectal (CCR) est l'un des cancers les plus mortels dans le monde. Le CCR est une maladie hautement hétérogène avec des niveaux élevés d'hétérogénéité intra- et inter-patients. Les cellules souches cancéreuses sont importantes au développement tumoral et des métastases. L'une des approches pour traiter le CCR est la thérapie de différenciation, l'induction de la différenciation des cellules souches cancéreuses pouvant ainsi conduire à une régression tumorale. Les organoïdes issues d'intestins sains et cancéreux représentent un modèle idéal pour étudier le processus de différenciation *ex vivo*. Le profilage de l'expression génique des organoïdes traités par le médicament peut être utilisé pour quantifier avec une grande précision le processus de différenciation. Les plates-formes de criblage de médicaments utilisant le séquençage d'ARN (RNAseq) dans les organoïdes ne sont pas encore établies. Nous avons ainsi développé une plate-forme basée sur l'ARN-seq ciblé à haut débit (TORNADO-seq) pour suivre l'expression génique de larges signatures géniques dans les organoïdes intestinaux, ce qui nous a permis de quantifier et d'évaluer de complexes phénotypes cellulaires et altérations cellulaires lors de traitements médicamenteux. Nous avons appliqué TORNADO-seq au criblage de médicaments sur des organoïdes issues d'intestins sains ou cancéreux. Nous avons trouvé de nombreux médicaments induisant la différenciation dans les organoïdes non cancéreux, y compris certains médicaments enrichissant spécifiquement certains phénotypes cellulaires, par exemple les cellules entéroendocrines. De plus, nous avons appliqué des médicaments induisant la différenciation identifiés sur des organoïdes cancéreux ( $APC^{lof}$ ;  $KRAS^{G12D}$ ;  $TP53^{lof}$ ) et découvert que la majorité de ces médicaments induisant la différenciation ciblaient également les organoïdes cancéreux. Sur la base des profils d'expression génique des organoïdes cancéreux traités, nous avons pu proposer et confirmer le mécanisme d'action de plusieurs médicaments indiquant que les cellules cancéreuses peuvent être sensibles aux altérations de la voie de biosynthèse du cholestérol. En général, l'utilisation de TORNADO-seq permet de prédire les connexions entre les phénotypes cellulaires et les voies de signalisation dans les organoïdes, ce qui permettrait la génération d'hypothèses dans une variété de systèmes biologiques.

La deuxième partie de la thèse est consacrée à l'étude des mécanismes sous-jacents à l'accumulation d'altérations génétiques dans les tumeurs du CCR. Il n'est pas connu si les gènes moteurs du cancer affectent le taux d'acquisition de mutations. Nous

avons ainsi effectué le séquençage de l'exome d'organoïdes dérivés de cellules uniques obtenus à partir de tumeurs avec des phénotypes de mutation  $APC^{lof}$ ,  $APC^{lof};KRAS^{G12D}$  ou  $APC^{lof};KRAS^{G12D};TP53^{lof}$  qui reflètent les premiers stades du développement du CCR. Nous avons découvert que la présence des mutations KRAS et P53 n'affectait pas le nombre de substitutions de nucléotides simples (SNS) dans les tumeurs, alors que cette dernière était augmentée par la mutation APC en comparaison aux cellules non mutées et était fortement corrélée avec l'âge de la tumeur. L'inactivation de P53 était fortement corrélée à la présence d'altérations du nombre de copies (CNA) à grande échelle (> 10 Mb).

Mots-clés: colorectal cancer; differentiation therapy; high-throughput screening; high-content screening; organoids; drug screening; RNA-seq.

# Abbreviations

2D - Two-Dimensional

3D - Three-Dimensional

5-FU – Fluorouracil

AKP – Mouse genotype including APC<sup>lof</sup>:KRAS<sup>G12</sup>:TP53<sup>lof</sup>

ALP - Alkaline phosphatase

APC – Adenomatous polyposis coli

APL - Acute promyelocytic leukemia

Ato - Atorvastatin

ATRA - All-trans retinoic acid

BCF – Biomolecular core facility

BIOP – Bio Imaging and optics core facility

CIMP - CpG island methylator phenotype

CIN – Chromosomal instability

CMS -Consensus molecular subtype

CRC – Colorectal cancer

CRE-ERT2 – Tamoxifen inducible fusion protein of Cre enzyme and estrogen receptor

CRIS - Colorectal cancer intrinsic signature

CRISPR - Clustered regularly interspaced short palindromic repeats

CSCs – Cancer stem cells

Dauno - Daunorubicin

DSS - Dextran sulfate sodium

EECs – Enteroendocrine cells

eGFP - Enhanced green fluorescent protein

EMT - Epithelial-mesenchymal transition

ENU - N-Ethyl-N-nitrosourea

FAP - Familial adenomatous polyposis

FCS - Fetal Calf Serum

FCCF – Flow cytometry core facility

FDA – Federal Drug Agency

FOLFIRI – Chemotherapy regimen including folinic acid, 5-FU and irinotecan

FOLFOX - Chemotherapy regimen including folinic acid, 5-FU and oxaliplatin  
GEMM – Genetically engineered mouse model  
GECF – Gene expression core facility  
HCF – Histology core facility  
HTS - High-throughput screening  
IBD – Inflammatory bowel disease  
IDG – Drug cocktail including IWP-2, DAPT and Gefetinib  
IV - Drug cocktail including IWP-2, VPA  
IF Immunofluorescence  
IHC - Immunohistochemistry  
IRES - Internal Ribosome Entry Site  
ISCs – Intestinal stem cells  
Lgr5 - Leucine rich repeat containing G protein-coupled receptor  
LICs - Leukemia-initiating cells  
Lova - Lovastatin  
Min – Multiple intestinal neoplasia  
MMR-D - Mismatch repair- deficiency  
MOA – Mechanism of action  
MSI – Microsatellite instability  
MSS – Microsatellite stable  
NGS – Next generation sequencing  
NMF - Non-negative matrix factorization  
OS – Overall survival  
PAS - Periodic acid-Schiff  
PDOs – Patient-derived organoids  
PDX - Patient-derived Xenograft  
RLU - Relative luminescence units  
SCNA - Somatic copy number alterations  
SERM - Selective estrogen-receptor modulator  
scRNA-seq – Single-cell RNA-seq  
SSP – Sessile serrated polyp  
TAs – Transit-amplifying cells  
TGF- $\beta$  - Transforming Growth Factor- $\beta$

TORNADO-seq - Targeted **ORgANOiD** sequencing

UMAP - Uniform Manifold Approximation and Projection for dimension reduction

UMI – Unique molecular identifier

VPA – Valproic acid



# Table of contents

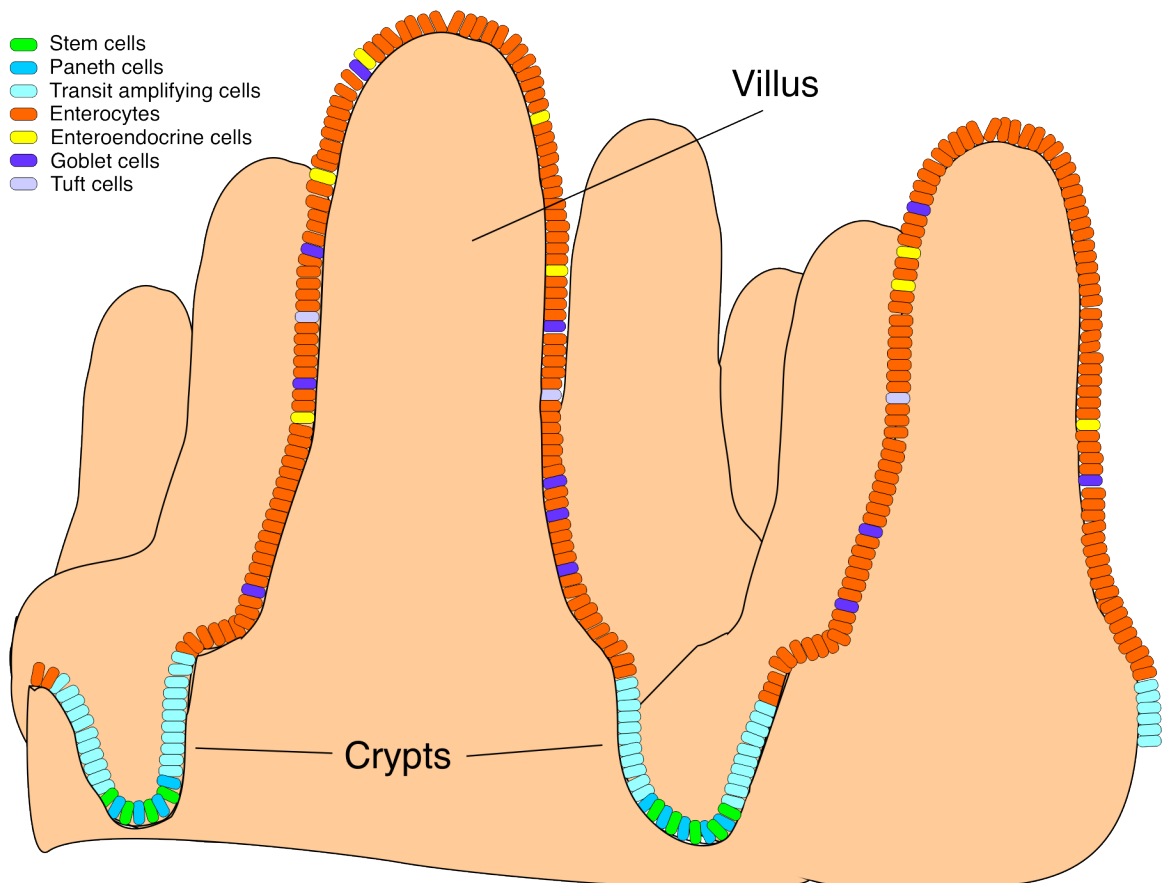
Acknowledgments .....	2
Abstract.....	4
Resume.....	6
Abbreviations .....	8
Table of contents .....	11
Chapter 1 Introduction .....	13
1.1 Overview of intestinal epithelium organization .....	13
1.2 Intestinal wt organoids .....	15
1.3 Intestinal epithelial cell types.....	16
1.4 Colorectal cancer .....	21
1.4.1 General information .....	21
1.4.2 Risks and nature of colorectal cancer.....	21
1.4.3 Characterization of colorectal cancer development.....	22
1.4.4 Subtypes and classification of colorectal cancer patients.....	24
1.4.5 Most common mutations in colorectal cancer.....	26
1.4.6 Colorectal cancer treatment options .....	26
1.5 Differentiation therapy .....	28
1.5.1 Differentiation therapy in acute promyelocytic leukemia.....	28
1.5.2 Basics for differentiation therapy in colorectal cancer .....	30
1.5.3 Examples of differentiation therapy in colorectal cancer .....	32
1.6 Colorectal cancer models.....	33
1.7 Colorectal cancer patient-derived organoids.....	36
1.8 Screenings in colorectal cancer .....	37
1.8.1 Reporter-based assays in cell lines .....	38
1.8.2 High-throughput screenings in 3D cultures.....	39
1.8.3 High-content imaging-based screenings.....	40
1.8.4 CRISPR screenings in organoids .....	40
1.8.5 Screenings in patient-derived organoids .....	41
1.8.6 RNA-seq based screening platforms .....	43
1.9 Prospects of drug screening methods in organoids .....	43
1.10 Objectives of the project.....	44
1.10.1 The first part of the thesis project.....	44
1.10.2 The second part of the thesis project.....	44
Chapter 2 Results .....	46
2.1 Overview .....	46
2.2 High-content, targeted RNA-seq screening in organoids for drug discovery in colorectal cancer ...	47
2.2.1 Graphical abstract.....	48
2.2.2 Summary.....	49

2.2.3	Introduction .....	50
2.2.4	Results .....	52
2.2.4.1	Screen design and validation .....	52
2.2.4.2	Identification of differentiation-inducing drugs in wt intestinal organoids .....	55
2.2.4.3	Comperhensive analysis of drug-induced differentiation patterns .....	58
2.2.4.4	Evaluating differentiation-inducing drugs in cancer organoids .....	60
2.2.4.5	Treatment responses of cancer organoids exhibit distinct signatures .....	63
2.2.4.6	Drug classification helps uncover mechanism of action against cancer organoids ..	64
2.2.5	Discussion .....	66
2.2.6	Acknoledgments and author contribution .....	68
2.2.7	STAR methods .....	69
2.2.8	Quantification and statistical analysis .....	76
2.2.9	Key resource table .....	76
2.2.10	References .....	80
2.2.11	Supplementary figures and tables .....	84
2.2.11.1	Supplementary figures .....	84
2.2.11.2	Supplementary tables .....	95
2.3	Genomic Instability Profiles at the Single Cell Level in Mouse CRCs of Defined Genotypes .....	104
2.3.1	Graphical abstract .....	105
2.3.2	Abstract .....	105
2.3.3	Introduction .....	106
2.3.4	Results .....	108
2.3.4.1	Clonal organoid cultures derived from single tumoral cells .....	108
2.3.4.2	Single nucleotide substitutions – Prevalence .....	110
2.3.4.3	Single nucleotide substitutions – Link to genotype and Distribution .....	113
2.3.4.4	Single nucleotide substitutions – Mutational signature .....	114
2.3.4.5	Copy number alterations .....	116
2.3.4.6	Evolution of tumour clones in mouse IV .....	118
2.3.5	Discussion .....	120
2.3.6	Conclusions .....	121
2.3.7	Materials and methods .....	123
2.3.8	References .....	127
2.3.9	Supplementary figures and tables .....	133
Chapter 3	Conclusions and outlook .....	148
3.1	The first part of the thesis .....	148
3.2	The second part of the thesis .....	150
Chapter 4	Bibliography .....	152
Chapter 5	Annexes .....	165
5.1	Single-Cell Studies of Intestinal Stem Cell Heterogeneity During Homeostasis and Regeneration	166
Curriculum Vitae	.....	179

# 1 Introduction

## 1.1 Overview of Intestinal epithelium organization

The small intestine is one of the most rapidly renewing organs in a mammalian organism [Gehart H., Clevers H., 2019]. The small intestine of a mouse represents a highly organized and repetitive structure (Fig. 1). The surface epithelium can be divided into two main parts: villi (big finger-like protrusions) and crypts (smaller in size invaginations surrounded by villi). There are 4 main cell types present in the villus: enterocytes, enteroendocrine cells, goblet cells, and tuft cells. Enterocytes are absorptive cells and are the most frequent cell type present in the intestinal villus region. There are three types of secretory cells - enteroendocrine cells (which secrete hormones), goblet cells (which secrete mucus), and tuft cells (which secrete interleukins and opioids). The crypt compartment contains the stem cell niche, where stem cells reside which can give rise to all types of differentiated cells.

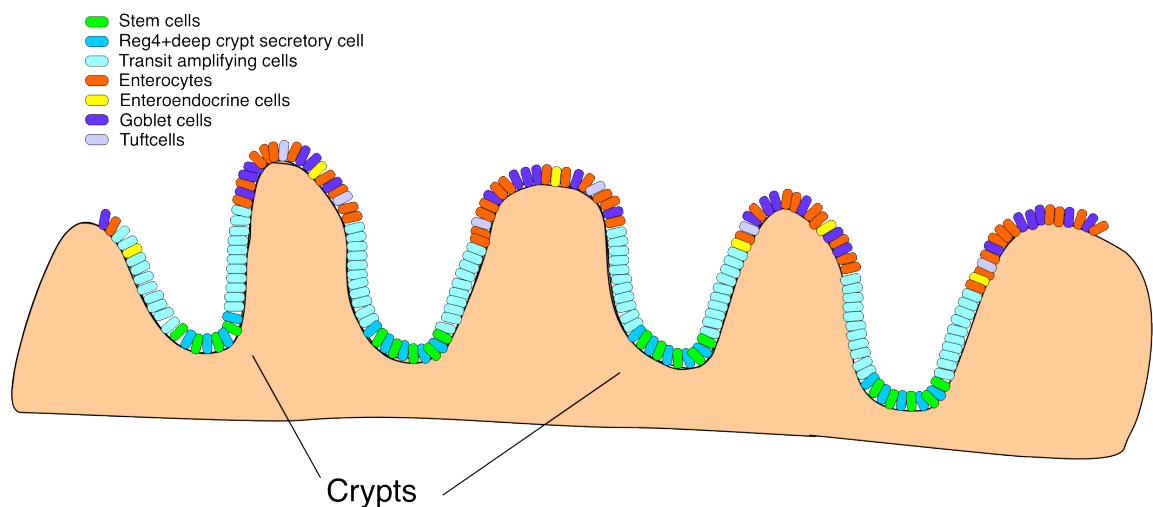


**Figure 1.** Schematic representation of intestinal epithelium organisation. Villus and crypt regions are highlighted.

Stem cells are marked by high expression of leucine-rich repeat-containing G-protein-coupled receptor 5 (Lgr5) and located at the bottom of the crypts where they are

surrounded by Paneth cells. Paneth cells are another type of differentiated cell that provides important growth factors such as Wnt and EGF important for stem cell niche maintenance. Paneth cells are also a part of the epithelial defence system against bacteria: they release various defensin proteins – small peptides with antimicrobial and immune signaling activities. Various pathways including Notch, Wnt, BMP, and EGF signaling pathways are essential for adult intestinal stem cell homeostasis and maintenance. Dividing stem cells start to migrate from the bottom of the crypt up to the top of the villi and while leaving this niche they start to differentiate. Differentiated cells move up along the villus axis to the top of the villus where they undergo apoptosis. This renewing cycle in the gut takes 4-5 days for enterocytes in the mouse and leads to a continuous replacement of the epithelium [Crosnier C et al., 2006].

The large intestine or colon is a posterior part of the gastrointestinal tract. It absorbs water and remaining nutrients, as well as compacts feces for excretion [Nick Barker, 2013]. The structure of the colon is remarkably different from the small intestine: the colon epithelium consists of many crypts and lacks a villus compartment. Moreover, there are major differences in cell type abundances: the colon is highly enriched by goblet cells and lacks Paneth cells. Alternative cell niche population, Reg4<sup>+</sup> deep crypt secretory cells was discovered in the colon which performs similar functions as Paneth cells in the small intestine: they are located at the bottom of the crypt and express Notch ligands Dll1/Dll4 as well as EGF [Sasaki N. et al., 2016]. Interestingly, unlike Paneth cells, Reg4<sup>+</sup> deep crypt secretory cells don't produce Wnt3, therefore *ex vivo* colonic organoid cultures require additional supplementation with Wnt3 ligand for successful growth. Similarly, to the small intestine, Lgr5 stem cells are also present in the colonic crypts and give rise to all differentiated cells of the colonic epithelium (Fig. 2).

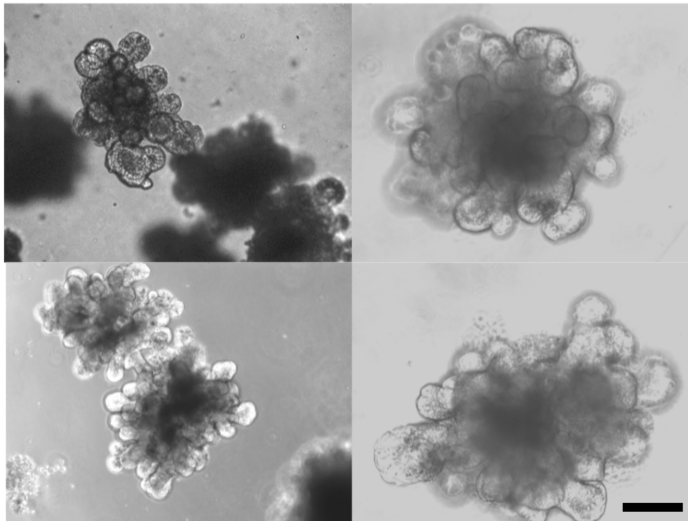


**Figure 2.** Schematic representation of colonic epithelium organisation.

## 1.2 Intestinal wt organoids

The research on intestinal development and homeostasis is important for our understanding of how the processes are orchestrated in the gut in health and how they can be distorted in disease. During the last decades, a growing number of experimental data brought us to a better understanding of the networks and pathways implicated in stem cell maintenance, differentiation, and cell turnover. A lot of this information was obtained from *in vivo* mouse models (e.g. knock-out and lineage tracing experiments). While providing reliable data, *in vivo* experiments are time-consuming and labor-intensive. At the same time, the utilization of cell lines for *in vitro* experiments provides only limited insight of the actual system. Thus, it became a necessity to use more reliable and more representative models of an organ than simple 2D cell lines. Development of extracellular matrix-based gels in the 1990s allowed to perform experiments with cells in 3D and boosted research in the organoid field [Barcellos-Hoff M.H. et al., 1989; Petersen O.W. et al., 1992; Bissell M.J. et al., 2005; Simian M. et al., 2017]. Nowadays, organoids are widely accepted as a useful research tool. A variety of screening experimental setups and platforms involving organoid cultures were established in the recent decade.

A key achievement in generating a 3D *ex vivo* model of intestinal development was obtained by Toshiro Sato from the Clevers group [Sato et al., 2009], who created a long-term culturing method of Lgr5<sup>+</sup> stem cells. The latter could form 3D structures (organoids) in the presence of essential growth factors - Rspo, EGF, and Noggin (the factors which imitate the stem cell niche of the intestine where they are released by supporting Paneth cells and the surrounding stroma). Organoids preserve crypt-villus morphology and consist of all cell types found in the intestinal epithelium (Fig. 3), therefore being a highly representative model of the intestinal epithelium. Organoids could be grown in the culture for months, therefore, being a useful research tool and allowing to perform long-term experiments. Application of organoids in the research boosted studies of intestinal stem cell differentiation. Several studies described the involvement of new pathways and drug combinations allowing to differentiate stem cell towards one particular cell type [X. Yin et al., 2013; O. Basak et al., 2017]. For example, the importance of EGF signaling for EEC differentiation was unravelled by the Clevers group: the EEC phenotype can be obtained by simultaneous blocking of Wnt, Notch and EGFR pathways [O. Basak et al., 2017].

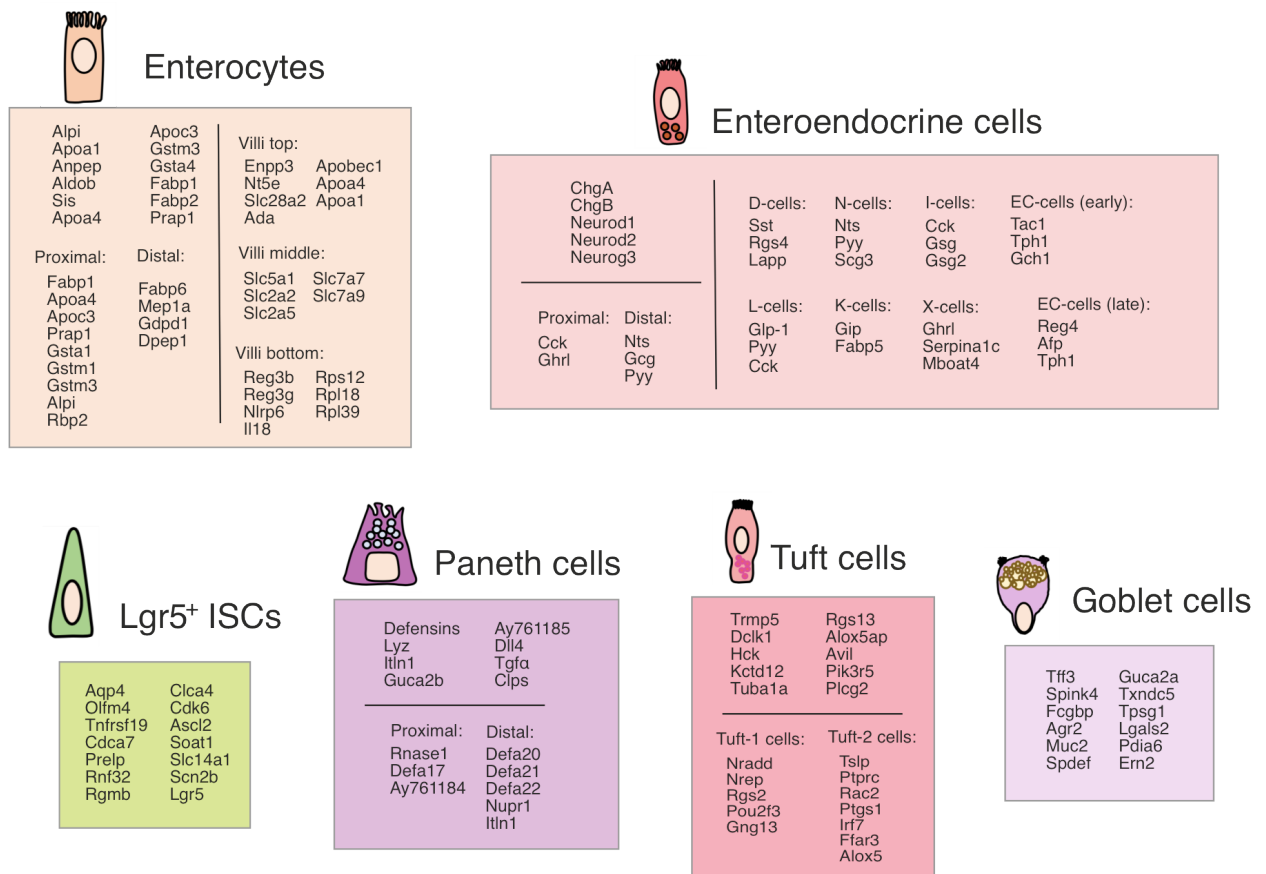


**Figure 3.** Intestinal organoids. The black bar corresponds to 100  $\mu$ m.

### 1.3 Intestinal epithelial cell types

[This chapter (1.3.) is cited here from my publication: “Single-Cell Studies of Intestinal Stem Cell Heterogeneity During Homeostasis and Regeneration. M. Norkin\*, C. Capdevila\*, R. I Calderon, T. Su, M. Trifas, P. Ordóñez-Morán, K. S Yan (2020). *Methods Mol Biol* 2171:155-167”. The full publication can be found in the Annex]

Constant tissue renewal of the intestinal epithelium during homeostasis is fueled by continuously dividing intestinal stem cells (ISCs) that reside at the base of the crypts [Barker N. et al., 2007; Gehart H., Clevers H., 2019; Clevers H., 2013]. These crypt-base columnar ISCs have been molecularly defined by the expression of the R-spondin receptor *Lgr5*, a leucine rich repeat containing G protein-coupled receptor [Barker N. et al., 2007]. Single-cell transcriptomic analysis reveals that *Lgr5*-eGFP<sup>+</sup> ISCs isolated from *Lgr5*-eGFP-IRES-CreERT2 mice are comprised of distinct clusters of both cycling and non-cycling Wnt- and R-spondin-dependent ISC as well as nascent transit-amplifying cell populations, consistent with multiple reports suggesting heterogeneity within the ISC pool [Buczacki S.J. Et al., 2013; Barriga F.M. Et al., 2017; von Furstenberg R.J. et al., 2014]. Many initial markers of ISCs were identified based on microarray and bulk RNA-seq profiling of cells isolated using *Lgr5* reporter expression. Subsequently, transcriptional profiling using scRNA-seq technology helped to confirm these markers and also allowed the identification of new ones, with varying degrees of characterization and validation (Fig. 4): *Aqp4*, *Olfm4*, *Tnfrsf19*, *Cdca7*, *Prelp*, *Rnf32*, *Cdk6*, *Rgmb*, *Clca4*, *Scn2b*, *Slc14a1*, *Ascl2* and *Soat1* [Haber A.L. et al., 2017; Grun D. et al., 2018; Grun. D. et al., 2016; van der Flier L.G. et al., 2009; van der Flier L.G. et al., 2009; Munoz J. et al., 2012].



**Figure 4.** Markers of small intestinal stem cells and their progeny. Summary of putative marker genes of different intestinal cell types based on transcriptional profiling.

Under homeostatic conditions, Lgr5<sup>+</sup> ISCs regenerate the intestinal epithelium [Barker N. et al., 2007]. In contrast, a distinct, label-retaining cell at the +4 cell position in intestinal crypts has been reported and proposed to serve as a quiescent/slowly cycling, reserve stem cell population that comes into play upon loss of Lgr5<sup>+</sup> ISCs [Sangiorgi E. et al., 2008; Tian H. et al., 2011; Montgomery R.K. et al., 2011; Yan K.S. et al., 2012; Potten C.S. et al., 1974; Potten C.S. et al., 1975; Potten C.S. et al., 1977]. Even though multiple markers have been described for the +4 cell [Sangiorgi E. et al., 2008; Tian H. et al., 2011; Montgomery R.K. et al., 2011; Powell A.E. et al., 2012; Yan K.S. et al., 2012], more recent studies show that many of these overlap with Lgr5<sup>+</sup> ISCs, or are expressed very broadly throughout the crypt [Munoz J. et al., 2012; Li N. et al., 2014; Itzkovitz S. et al., 2011]. While it remains unclear if there are other professional reserve stem cell populations that co-exist within the crypt alongside the Lgr5<sup>+</sup> ISCs, there is increasing evidence of diverse cell types that are capable of exhibiting cellular plasticity to become activated upon injury conditions to rapidly regenerate the epithelium following loss of Lgr5<sup>+</sup> ISCs [Santos A.J.M. et al., 2018]. Indeed, some recent observations suggest that it is not a single type of quiescent stem cells but rather a number of functionally distinct,

non-*Lgr5*-expressing cell type that reverts to a stem cell state in an injury-inducible fashion [Yan K.S. et al., 2017; Sangiorgi E. et al., 2008; Yan K.S. et al., 2012].

Globally, along the crypt-villus axis, the progenitor cells derived from *Lgr5*<sup>+</sup> ISCs and their immediate transit amplifying (TA) progeny reside in the crypts close to the stem cells and move into the villi upon differentiation. Absorptive enterocytes are the most abundant cell lineage in the intestinal epithelium. Additionally, there are multiple secretory cell lineages including mucus-producing goblet cells, Paneth cells, enteroendocrine cells, and rare tuft cells [Clevers H., 2013]. All these differentiated cell lineages contribute to specific functions of the intestine. The advantage of scRNA-seq methodology is that it can be applied to distinguish specific gene-expression in individual sub-populations and it can help to infer the hierarchical relationships of individual lineages. (Fig. 4).

Tuft cells are a rare cell type involved in chemosensory function. Tuft cells express proteins known to be involved in taste signal transduction and also in an immune response against parasite infection [Gerbe F. et al., 2016; Howitt M.R. et al., 2016]. Tuft cells highly express such markers as *Trpm5*, *Dclk1*, *Rgs13*, *Alox5ap*, *Avil*, *Hck*, *Kctd12*, *Tuba1a*, *Pik3r* and *Plcg2* [Haber A.L. et al., 2017; Grun D. et al., 2018; Grun. D. et al., 2016; Ting H-A. et al., 2019; Gerbe F. et al., 2009]. Recent findings using scRNA-seq demonstrated the existence of two distinct subtypes of tuft cells: tuft-1 cells are highly enriched in genes related to neuronal development, while tuft-2 cells show upregulation of genes related to immunity. Specifically, tuft-1 cells are enriched for *Nradd*, *Gng13*, *Nrep*, *Rgs2*, and *Pou2f3* [Haber A.L. et al., 2017] (Fig. 4). Conversely, tuft-2 cells showed higher expression levels of the TH2-promoting cytokine thymic stromal lymphopoietin (*Tslp*) and *Ptprc* (pan-immune cell marker CD45), as well as enrichment in transcripts for *Rac2*, *Ptgs1*, *Irf7*, *Ffar3* and *Alox5*.

Enterocytes are the most abundant cell type in the intestinal epithelium, making up more than 80% of intestinal epithelial cells [de Santa B.P. et al., 2003]. Enterocytes are predominantly located in the villus region and function in the hydrolysis, absorption, and transport of nutrients [van der Flier L.G. et al., 2009]. Many enterocyte markers were identified using bulk RNA-seq [Kazakevych J. et al., 2017] and scRNA-seq [Haber A.L. et al., 2017; Grun D. et al., 2018; Grun. D. et al., 2016; Ting H-A. et al., 2019; Moor AE. et al., 2018], among them *Alpi*, *Apoa1*, *Anpep*, *Aldob*, *Sis*, *Apoa4*, *Prap1*, *Apoc3*, *Gstm3*, *Gsta4*, *Fabp* and *Fabp2*. Recently, scRNA-seq was used to identify enterocyte progenitor populations [Grun D. et al., 2018; Grun. D. et al., 2016], to examine enterocyte regional diversity through the gut [Haber A.L. et al., 2017], and to investigate the spatial allocation of distinct functional classes of enterocytes along the intestinal crypt-villus axis [Gassler



N. et al., 2006; Moor AE. et al., 2018]. It has been shown that earlier enterocyte progenitors express high transcript levels of ribosomal proteins (*Rn45s*, *Rps19* and *Rps12*), *Dmbt1*, *Reg3g*, *Ube2c* and low levels of those for enterocyte-specific genes (*Prap1*, *Apoa1*, *Apoa4*, *Apoc3*, etc.) [Haber A.L. et al., 2017]. Late progenitor cells lost the expression of ribosomal proteins with concurrent elevation of enterocyte-specific transcripts. Finally, mature enterocytes were characterized by further upregulation of enterocyte-specific mRNA transcripts.

A study of regional enterocyte markers reveals that *Fabp1*, *Apoa4*, *Apoc3*, *Gsta1*, *Gstm3*, *Gstm1*, *Alpi*, *Prap1* and *Rbp2* are highly expressed in the proximal part of the intestine, while *Fabp6*, *Mep1a*, *Dpep1* and *Gdpd1* are predominantly expressed in the distal small intestine [Haber A.L. et al., 2017]. This is consistent with differential absorptive functions along the longitudinal proximal-to-distal gut axis. Moreover, a combination of high-throughput scRNA-seq and bulk RNA-seq of laser-microdissected crypt-villus sections revealed that enterocyte transcriptional signatures differ along the crypt-villus axis consistent with their zonation [Moor AE. et al., 2018]. Enterocytes at the bottom of the crypt showed enrichment in ribosomal/proliferation signatures (*Rps12*, *Rpl18* and *Rpl39*) and antimicrobial program peptides (*Reg3b*, *Reg3g*, *Nlrp6* and *Il18*) [Moor AE. et al., 2018]. Enterocytes in the middle of the crypt-villus axis are enriched in the genes responsible for the processing and absorption of various nutrients, especially amino acids (*Slc7a7* and *Slc7a9*) and carbohydrates (*Slc5a1*, *Slc2a2* and *Slc2a5*) [Moor AE. et al., 2018]. Enterocytes at the top of the villus exhibited a distinct expression program: enrichment in cell adhesion signature (*Egfr*, *Klf4*, *Fos* and *Junb*), purine catabolism genes (*Enpp3*, *Nt5e*, *Slc28a2*, and *Ada*), and apolipoproteins/cholesterol processing genes (*Apobec1*, *Apoa4* and *Apoa1*) (Fig. 4) [Moor AE. et al., 2018].

Goblet cells are secretory cells present in both the small intestine and colon. They produce and secrete mucus into the intestinal lumen, which facilitates the migration of chyme through the gut and contributes to a physical barrier that prevents microorganism and toxins from direct contact with the mucosa. In addition, goblet cells are also shown to be expanded in response to parasite infection [Haber A.L. et al., 2017; Artis D. et al., 2004; Biton M. et al., 2011]. Several scRNA-seq experiments revealed many markers of goblet cells [Haber A.L. et al., 2017; Grun D. et al., 2018; Grun. D. et al., 2016]: *Tff3*, *Spink4*, *Fcgbp*, *Agr2*, *Muc2*, *Txndc5*, *Tpsg1*, *Spdef*, *Guca2a*, *Lgals2*, *Pdia6* and *Ern2*. *Spdef* has been shown to have an essential role in goblet cell differentiation [Noah TK. et al., 2010; Gregorieff A. et al., 2009] (Fig. 4).

Paneth cells are intermingled with Lgr5<sup>+</sup> ISCs at the crypt base and function in antimicrobial defense and metabolic regulation of ISCs. They interact with ISCs via signaling pathways such as Notch, Wnt3, and EGF [Sato T. et al., 2011; Pellegrinet L. et al., 2011; Yilmaz O.H. et al., 2012]. While the first Paneth cell markers (lysozyme and defensins) were reported decades ago [Ghoos Y. et al., 1971; Porter E.M. et al., 1997], recent RNA-seq studies on sorted CD24<sup>+</sup> Paneth cells reveal additional transcripts differentially enriched in Paneth cells relative to ISCs: *Dll4*, *Tgfa*, *Wnt11*, *Clps*, *AY761185*, *Itln1* and *Guca2b* [Sato T. et al., 2011]. Recent scRNA-seq studies additionally showed that Paneth cells also exhibit regional diversity, and the expression profiles of these cells are different in distal and proximal part of the small intestine: for example, *Rnase1*, *Defa17* and *AY761184* are highly expressed in the duodenum, while *Defa20*, *Defa21*, *Defa22*, *Nupr1* and *Itln1* show the highest expression in ileum [Haber A.L. et al., 2017]. Other Paneth-specific markers that show uniform expression through the gut include *Lyz1*, *Ang4*, *Clps*, and *Habp2* (Fig. 4) [Haber A.L. et al., 2017]. It still remains unclear how and why Paneth cells are heterogeneous, and the functional implications on their interactions with ISCs and on the other Paneth cell functions.

Enteroendocrine cells (EECs) produce hormones that regulate digestion, metabolism, and participate in chemosensation, including nutrient detection. EECs represent a very small fraction of epithelial cells that are widely dispersed throughout the intestinal epithelium and have been elusive to molecular characterization due to their marked cellularity heterogeneity. While EECs are located throughout the entire intestinal epithelium, they also exhibit regionalized patterns of gene expression. Recently, scRNA-seq studies have identified new putative markers for the EECs subpopulations and further underscored the heterogeneity of this lineage. At least 10 different EEC subtypes have been identified either *in vivo* or *in vitro* intestinal organoid culture [Haber A.L. et al., 2017; Gehart H. et al., 2019; Grun D. et al., 2018; Grun. D. et al., 2016; Gehart H., Clevers H., 2019; Basak O. et al., 2017]. *Neurog3*, *Neurod1* and *Neurod2* are known EEC markers specific to EEC progenitor populations [Haber A.L. et al., 2017]. In addition, the EECs markers *ChgA* and *ChgB* were shown to be predominantly expressed in the most abundant EEC type – enterochromaffin (EC) cells [Haber A.L. et al., 2017; Gehart H. et al., 2019]. Many more putative novel markers were identified for each of the EEC subpopulations, using sequencing technology, including scRNA-seq: *Gip*, *Fabp5* (K-cells); *Reg4*, *Afp*, *Tph1* (EC-cells (late)); *Sst*, *Lapp*, *Rgs4* (D-cells); *Ghrl*, *Serpina1c*, *Mboat4* (X-cells); *Cck*, *Gsg*, *Gsg-2* (I-cells); *Glp-1*, *Pyy*, *Cck* (L-cells); *Nts*, *Pyy*, *Scg3* (N-cells); *Tac1*, *Tph1*, *Gch1* (EC cells (early)) [Haber A.L. et al., 2017; Gehart H. et al., 2019; Grun D. et

al., 2018; Grun. D. et al., 2016; Gehart H., Clevers H., 2019; Basak O. et al., 2017; Habib A.M. et al., 2012]. Some of the markers (e.g. *Cck*, *Gcg*) are shown to be expressed in several EEC subtypes simultaneously. Furthermore, while some markers are predominantly expressed in the proximal small intestine (*Cck*, *Ghrl*), others are more localized in the distal region (jejunum, ileum: *Nts*, *Gcg*, *Pyy*). *Sct* is expressed in all subtypes of EECs and detected in both proximal and distal part of the gastrointestinal tract [Gehart H. et al., 2019; Egerod K.L. et al., 2012] (Fig. 4). These EEC lineage markers and their distribution need further confirmatory experimental validation and biological interpretation. The diverse repertoire of EEC subsets found by scRNA-seq likely reflects the multi-faceted functions of these cells that orchestrate chemosensation, digestion, metabolism, and communication with other organs via their hormone products.

## 1.4 Colorectal cancer

### 1.4.1 General information

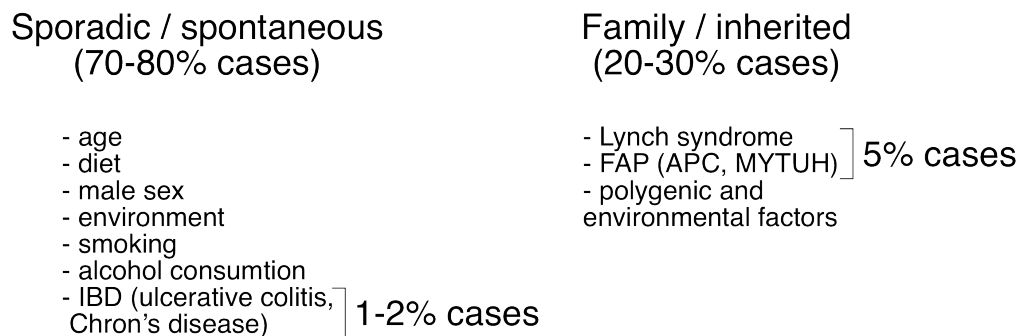
Colorectal cancer is one of the most commonly diagnosed cancers worldwide. Almost 2 million people around the globe were diagnosed with colorectal cancer (CRC) in 2020, resulting in nearly 1 million deaths as of 2020; up from 490,000 in 1990 [Yue Xi et al., 2021]. Due to the implementation of earlier testing, the numbers of diagnosed cases rise from year to year. The most common way to treat colon cancer is surgical removal of the tumor. But often patients are diagnosed with cancer in later stages (III-IV) of its development when the disease has already spread to distant organs. 5-year OS rate for the patients diagnosed with metastasis is only 14%. Cancer cells predominantly metastasize to the liver, lungs, peritoneum, or ovaries [Leah H. et al., 2021; Xie Y.H. et al., 2020].

### 1.4.2 Risks and nature of colorectal cancer

Colorectal cancer cases show both - sporadic (70% of all cases) and hereditary nature (20-30% of all cases) (Fig. 5). The most common inherited risk conditions (about 5% of all CRC cases) are Lynch syndrome and familial adenomatous polyposis (FAP). The remaining family cases have presumably polygenic or shared environmental risks [Nora B. Henrikson, 2015] but one can not exclude the possibility of insufficient testing of family members in the past. FAP is characterized by the development of many colorectal polyps throughout life. This is typically caused by the presence of mutations in a tumor

suppressor gene - adenomatous polyposis coli (APC) - a gene that controls cell proliferation (but APC is also most commonly mutated in sporadic cases). FAP can be further distinguished by the presence of autosomal recessive mutations (alterations in both copies of the gene) in the DNA repair gene MUTYH which causes increased mutational burden in those patients. The lifetime risk for developing CRC in FAP condition is 100% (reached by age 40-50 years). Similarly, DNA repair gene mutations occur in Lynch syndrome. Mutations in MLH1, MSH2, and MSH6 genes account for more than 95% of all Lynch syndrome cases. Patients carrying this disease (MLH1 and MSH2) have up to 80% lifetime risk of developing colorectal cancer [Jasperson K.W. et al., 2010].

## Colorectal cancer



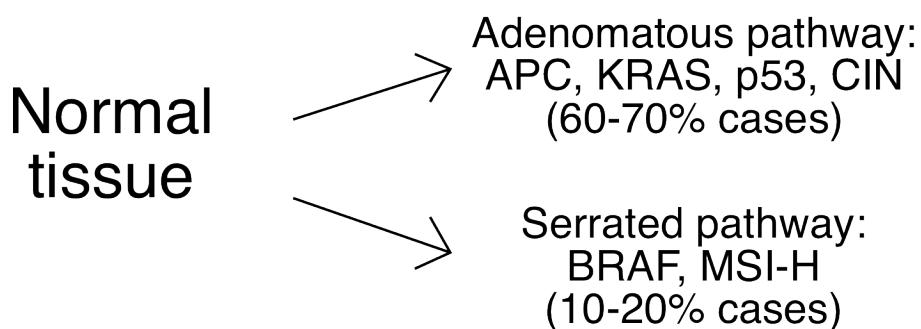
**Figure 5.** Colorectal cancer risks.

Most CRC cases arise spontaneously (70% of all cases). Main non-hereditary risk factors include age, male sex, diet, environmental factors, smoking, excess alcohol consumption. Inflammatory bowel disease (ulcerative colitis and Crohn's disease), which is caused both by environmental and genetic factors and characterised by chronic inflammation in the colon and small intestine, accounts for 1-2% of all CRC cases. The risk of IBD patients for developing CRC increases with the duration of the disease and reaches 18% after 30 years of disease [Kim E.R. et al., 2014].

### 1.4.3 Characterisation of colorectal cancer development

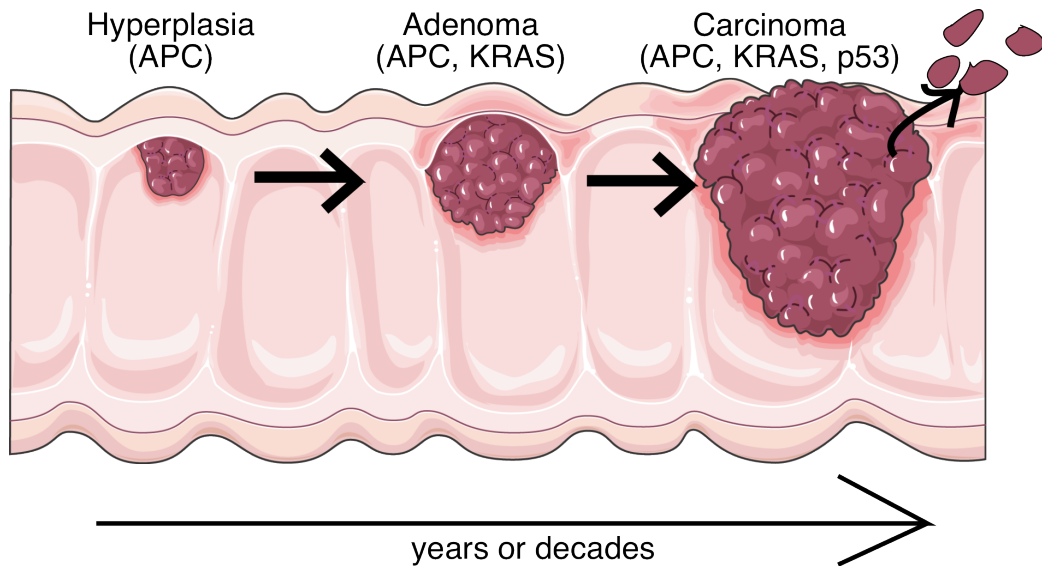
Colorectal cancer is characterized by a multistep carcinogenesis process which includes an adenoma-carcinoma progression sequence. The model was proposed by Fearon and Vogelstein in 1990 and implied a successive acquirement of mutations with the tumor development [Fearon, E. R. et al., 1990]. There are two main pathways of acquiring genetic changes which lead to CRC: the adenomatous pathway and the

serrated pathway (Fig. 6). Both pathways have their typical histological features (adenomas or sessile serrated polyps (SSPs)) and distinguished molecular origins. The adenomatous pathway is the most common and accounts for 60-70% of cases. Its tumors originate from adenomas with tubular histology and have alterations in the APC gene as a first driving mutation. With time, these tumors acquire chromosomal instability (CIN) – the accelerated rate of gains or losses of the chromosomal parts. In contrast, SSPs have very different histological features: these tumors are flat and have serrated glands. Serrated pathway cancer development begins with BRAF mutation. Serrated tumors are characterized with features such as microsatellite instability (MSI), and CpG island methylator phenotype (CIMP) pathway (Fig. 6).



**Figure 6.** Two genetic pathways of CRC development.

Adenomatous pathway: the initial event in tumor formation is believed to be the inactivation of the APC gene (Fig .7), which is also the most common mutation in colon cancer [Kuipers E.J. et al., 2015; cBioportal]. APC is a member of the Wnt signaling pathway and encodes a protein that negatively regulates the level of beta-catenin protein. Nonsense mutations in the APC gene lead to constitutive activation of Wnt signaling (via beta-catenin-mediated transcriptional activation) which leads to a stem cell-like/proliferative cell behavior and the formation of adenomas. Inactivation of the APC gene is often followed by activating mutations in KRAS, an upstream effector protein in the Raf-MEK-ERK pathway. This leads to acceleration in cell cycle progression and further tumor growth. At later time points, inactivation of the tumor suppressor gene p53, the cell cycle gatekeeper, allows tumor cells to escape cell cycle arrest and apoptosis. These events lead to further accumulation of mutations, increased CIN, and progression to the carcinoma stage which is characterized by metastatic dissemination (Fig. 7). The whole process of tumor development might take years or even decades.



**Figure 7.** CRC development stages.

#### 1.4.4 Subtypes and classification of colorectal cancer patients

Colorectal cancer is a heterogeneous disease with many possible mutation combinations which can drive tumor development. Tumors can be distinguished by the genomic instability status - CIN (SCNA), MSI/MSS or CIMP high or low tumors; they can be distinguished by the presence of BRAF and KRAS WT or mutant variants; they can be low or rich in immune infiltration; they can have different MMR-D (mismatch repair-deficiency) status. All these factors influence therapy efficiency and the survival of the patients. Unfortunately, despite progress in molecular biology approaches in the last 10 years and a lot of information gathered on tumor development, histopathological analysis is still the primary tool for therapy decision-making. Nevertheless, several attempts were made to summarize the knowledge about CRC and to classify colorectal cancer patients into particular groups to improve therapy outcome, i.e. in order to decide what group of patients would benefit from receiving a certain drug or drug combination. One of the earliest findings came shortly after the FDA approval of the first targeted therapy – cetuximab – a chimeric IgG1 monoclonal antibody targeting EGFR [Cunningham D. et al., 2004]. Clinical trials showed that only a small proportion of CRC patients benefit from the treatment. It turned out that Cetuximab is primarily efficient for KRAS WT tumors, while KRAS mutant tumors do not respond to the treatment because of continuous activation of downstream EGFR signaling in those lesions [Lièvre A., et al, 2006, Fiore Di F. et al., 2007, De Roock W. et al., 2008]. This discovery led to the widespread application of molecular profiling for the KRAS status in patients prior to therapy regime selection alongside the histopathological evaluation of the tumor samples.

Further, the first molecular CRC classification was performed based on tumor gene expression profiles [Guinney J. et al. 2015]. An international consortium of 6 laboratories combined their knowledge and skills to perform the clustering of more than 4000 patient samples into 4 consensus molecular subtypes of colorectal cancer: CMS1, CMS2, CMS3, and CMS4 (Fig. 8 [Guinney J. et al. 2015]). The work was based and driven by previous research attempts to classify CRC tumors based on gene expression [Sadanandam A. et al., 2013; Melo F. De S E et al, 2013; Marisa L. et al., 2013; Roepman P. et al, 2014,].

CMS1 MSI immune	CMS2 Canonical	CMS3 Metabolic	CMS4 Mesenchymal
14%	37%	13%	23%
MSI, CIMP high, hypermethylation	SCNA high	Mixed MSI status, SCNA low, CIMP low	SCNA high
<i>BRAF</i> mutations		<i>KRAS</i> mutations	
Immune infiltration and activation	WNT and MYC activation	Metabolic deregulation	Stromal infiltration, TGF- $\beta$ activation, angiogenesis
Worse survival after relapse			Worse relapse-free and overall survival

**Figure 8.** Consensus molecular subtypes of colorectal cancer. Proposed taxonomy of colorectal cancer, reflecting significant biological differences in the gene expression-based molecular subtypes [Guinney J. et al. 2015].

CMS1 subtype is characterized predominantly by hypermutated BRAF tumors, with high immune infiltration, high MSI status, high CIMP, and low SCNA status. The patients from this cohort have the worst survival rate after relapse. CMS2 subtype or canonical subgroup is the major group with a high percentage of APC mutations, activated Wnt and Myc pathways, SCNA high, MSI and CIMP low. CMS3 or the metabolic subgroup has the highest incidents of KRAS mutations and the tumors are enriched by multiple signatures of metabolic dysregulation. Mesenchymal CMS4 subtype has the highest rate of stromal infiltration, high SCNA levels, and TGF- $\beta$  activation levels, with the worst overall survival rate.

While providing important information on markers for CRC subtypes via gene expression profiles, the classification has yet to confirm its prognostic relevance for therapy decision-making. While some progress was achieved in predicting patient survival and appropriate therapy according to CMS classification [Del Rio, M. et al., 2017],

many clinical trials reported conflicting results regarding the efficacy of targeted agents [Stintzing S. et al., 2019; Okita A. et al., 2018; Y. Buikhuisen J.Y. et al, 2020]. Moreover, CMS relevance for metastatic lesions was questioned and an alternative classification was proposed for those tumors [Mendelaar P.A. et al., 2021]. Some discrepancies in CMS classification might be attributed to the fact that gene expression profiles were obtained not only from epithelial tumor cells but the whole tumor sample including surrounding immune and stromal cells [Y. Buikhuisen J.Y. et al, 2020]. Thus, an additional classification system for the primary tumors was introduced later on describing 5 CRC intrinsic signature (CRIS) clusters and was based solely on epithelial cell transcripts. Not surprisingly, all these classification methods have limited overlap as they used cells of different origin in their analysis. While these classification efforts are a huge step forward in implementing a new strategy for therapy decision making, it might well be that the level of heterogeneity in CRC patient tumors is much higher than anticipated and cannot be described using only cohorts of 4-5 clusters. In the nearest future, a more personalized approach using different types of PDO platforms will be the next important step in CRC therapy decision making that might bring new insights to the field.

#### 1.4.5 Most common mutations in CRC

According to the cBioportal website, which accumulates thousands of human colorectal cancer metadata, the most common mutations in CRC are APC, TP53, KRAS, SMAD4, and PIK3CA. Other common mutations in CRC tumors and the corresponding pathways are MAPK signaling: KRAS, NRAS, BRAF, EGFR, ERBB2, ERBB3, FLT3; Wnt signaling: APC, RSPO3, FBXW7, RNF43, ZNRF3, SOX9, TCF7L2, AXIN2, CTNNB1, TP53, and ATM are the apoptosis-related genes, PI3K/AKT signaling – PIK3CA, PTEN, AKT1, AKT2, PIK3CD; TGF- $\beta$  signaling – SMAD4, SMAD2, TGFBR1, TGFBR2; Chromatin modifiers: ARID1a, ARID2, ARID1B, and ARID4B.

Several studies establishing patient-derived organoids revealed a similar pattern of most common mutations occurring in patient samples: KRAS, APC, TP53, BRAF, PIK3CA, ATM, SMAD4, SOX9, FBXW7 [van de Wetering M. et al., 2015; Schuette M. et al., 2017; Ooft N.S. et al., 2019].

Analysis of metastatic samples reveals that the most common mutations driving highly aggressive tumors include APC, KRAS, SMAD4, TP53, SRC, PIK3CA, SMARCA4 [cBioportal; Vlachogiannis G. et al., 2018].

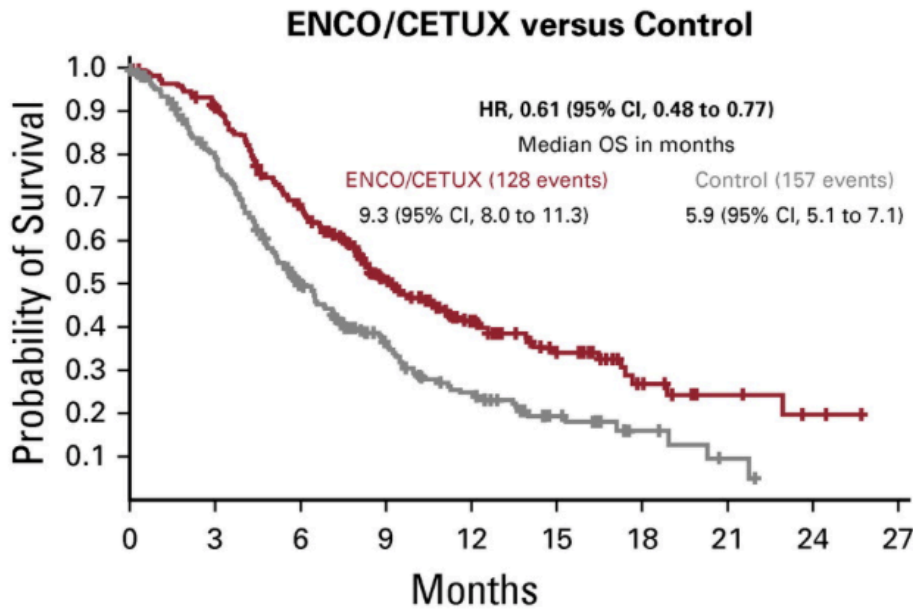
#### 1.4.6 Colorectal cancer treatment options



Early-stage (I-II) CRC patients undergo surgery as the main therapy option. More advanced stages (II-III) are complementarily combined with chemotherapy. The majority of stage IV patients undergo chemotherapy treatment without surgical intervention. Until the early 2000s, 5-FU and Leucovorin were the main first-line conventional therapies for advanced CRC stages when oxaliplatin and Irinotecan-based regimens such as FOLFOX and FOLFIRI were introduced for patient treatments [Gustavsson B. et al., 2015]. The implementation of Irinotecan and Oxaliplatin lead to increased survival of metastatic patients from 12 months to 18-21 months [Cunningham D. et al., 2004]. In 2004 the first targeted therapy was introduced – cetuximab - a chimeric IgG1 monoclonal antibody targeting EGFR receptor [Cunningham D. et al., 2004]. It turned out that it's effective only for a cohort of KRAS and BRAF wt patients [Lièvre A., et al, 2006; Fiore Di F. et al., 2007; De Roock W. et al., 2008]. Moreover, the effect on survival of those patients is moderate with only 4 extra months of median overall survival (FOLFIRI vs FOLFIRI + Cetuximab: 21 months vs 24.9 months) [E. V. Cutsem et al., 2009]. Another EGFR targeting agent Panitumumab was approved by FDA in 2006 and is interchangeable with Cetuximab (median overall survival, 23.8 months with FOLFOX plus panitumumab vs 19.4 months with FOLFOX alone [Douillard JY. Et al., 2014]).

BRAF V600E patients have worse survival outcome compared to BRAF WT patients with 15-20% shorter median OS [Taieb J. et al., 2019]. BRAF mutations are found in 5-10% of CRC tumors and BRAF V600E accounts for a vast majority of metastatic CRC cases (98%) [Taieb J. et al., 2019]. Current therapy treatment is not any different from other CRC cases and includes fluoropyrimidine and either irinotecan or oxaliplatin. Only in 2020, FDA approved a combination of encorafenib (BRAF inhibitor) and cetuximab for BRAF V600E metastatic patients based on phase 3 BEACON CRC trial results (Fig. 9) [Kopetz S. et al., 2019].

After successful implementation of the first targeted therapy against EGFR in 2004, several drugs targeting the VEGF pathway were also approved by the FDA for the treatment of metastatic disease in 2004-2015 - Bevacizumab, Ziv-aflibercept, Regorafenib, Ramucirumab [Hurwitz, H. et al., 2004; Van Cutsem E et al., 2012; Grothey A. et al., 2012; Tabernero J., et al., 2015]. Bevacizumab is a humanized IgG monoclonal antibody targeting VEGF-A. Bevacizumab increased overall survival for metastatic patients from 15.6 months to 20.3 months in phase III AVF2107 trials [Hurwitz, H. et al., 2004]. Recently one more drug targeting VEGF pathway has been approved by the FDA – Ramucirumab - a humanized monoclonal IgG antibody targeting VEGFR-2 [Tabernero, J. et al., 2015].



**Figure 9.** Overall survival results in phase 3 BEACON CRC clinical trial. ENCO/CETUX, encorafenib plus cetuximab; HR, hazard ratio; OS, overall survival. [Tabernero J. et al., 2021].

Another milestone in treating CRC patients was achieved with the FDA approval of checkpoint inhibitors in 2017-2018. Pembrolizumab (PD-1 like antibody), nivolumab (PD-1 antibody), and ipilimumab (CTLA-4 inhibitor) were approved for treating metastatic CRCs with MMR-deficient and MSI-H tumors where they showed successful prolongation of overall survival. The effect though extends to tumors with only a high mutational burden which is probably explained by the expression of abnormal antigens which can be further recognized by immune cells.

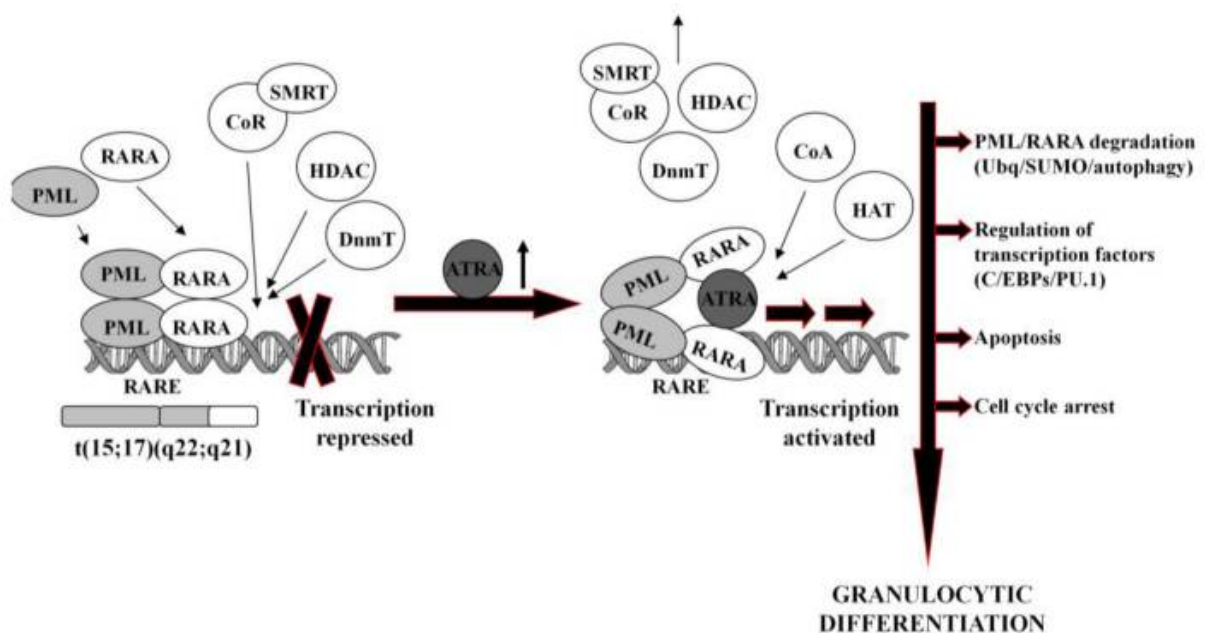
Recent advances in screening procedures and novel drug treatments undoubtedly improved patient lives and their survival rate. At the same time, if we compare the median survival rates from the late 1990s to nowadays the achieved effect is obviously very small. This is especially relevant for metastatic patients where median overall survival increased from 12 months to only 25 months. Despite many attempts in the development of targeted drugs over the past years, more specific targeted therapy is still missing and urgently needed for CRC treatments.

## 1.5 Differentiation therapy

### 1.5.1 Differentiation therapy in acute promyelocytic leukemia

One of the most successful examples of differentiation therapy in cancer is retinoic acid derivatives in the treatment of acute promyelocytic leukemia (APL). APL is a type of acute myeloid leukemia which is characterized by the expanded presence of

promyelocytes (immature blood-forming cells) in the blood and bone marrow. The alteration is caused by the impaired maturation process of those blood cells and the majority of lethal cases are driven by blood clots. Mechanistically, almost all cases involve chromosomal translocation which results in the production of PML-RARa fusion proteins [Hansen LA. et al., 2000]. In homeostatic conditions, RARa is a nuclear transcription factor that is getting activated upon the binding of steroid hormones driving consequential heterodimerization with RXR transcriptional factors. RARa-RXR transcriptional complexes regulate differentiation and growth of different cell lineages including promyelocytes. In the disease, the differentiation function of the RARa-RXR complexes is impaired by the dominant-negative function of PML-RARa fusion protein and the expression of corresponding differentiation genes is repressed (Fig. 10). All-trans retinoic acid (ATRA) drugs bind to PML-RARa fusion protein causing conformational changes which lead to transcriptional activation of the differentiation program and maturation of promyelocytes into granulocytes [Gocek E. et al., 2011].



**Figure 10.** Molecular mechanism of ATRA-induced differentiation of APL cells [Gocek E. et al., 2011].

Another paper highlights the dual role of retinoids in APL treatment – differentiation of promyelocytes and the elimination of leukemia-initiating cells (LICs), a stem cell subset of promyelocytes [Nasr R. et al., 2008]. The authors elucidated the important role of cyclic AMP signaling in those processes which are important for the destruction of the PML-RARa complex and eradication of LICs. They claim that LICs elimination is more relevant for the APL treatment than differentiation of promyelocytes as only the latter itself doesn't cure patients. Even though it seems that the effect on LICs is important, the dual action

of retinoic acid is not separable into differentiation vs. LIC depletion and the contribution of promyelocytes differentiation is still substantial for disease control. Moreover, the LIC elimination process requires higher doses of retinoic acid, and the mechanism of clearance is yet not understood, which might well include the execution of some differentiation programs. Thus, the importance of retinoic acid as a differentiation agent is invaluable.

The discovery of retinoids for the treatment of APL completely changed the landscape of disease prognosis. Before the implementation of differentiation therapy, the disease was highly fatal and only about 40% of patients could be cured. ATRA treatments lead to a drastic increase in patient survival with more than 95% of complete remissions which highlights the importance and the actual potential of differentiation therapy.

### 1.5.2 Basics for differentiation therapy in colorectal cancer

Tumors are composed of different types of cancer cells that contribute to tumor heterogeneity [Li H. et al., 2017]. Among those populations, there are cancer stem cells (CSCs) that drive tumor growth and metastatic progression. In the primary tumor, they are marked by *Lgr5* gene expression, which is one of the markers for CSCs in CRC. Differentiated cells are another prominent cancer cell population that can be distinguished, for example, by KRT20 expression. Different intermediate progenitor cell populations are also detected in CRC tumors [Baker A.M. et al., 2015]. Cancer cells differ both by the level of proliferation and level of differentiation and usually, these two features are inversely correlated in tumors. Thus, using CRC patient tumor samples and corresponding human organoid xenografts in mouse models it has been shown that *Lgr5* stem cells and KRT20 differentiated cells are localized in different parts of the tumor – stem cells occupying the tumor border while differentiated cells reside in the inner tumor regions [Shimokawa M. et al., 2017]. Moreover, while differentiated cells are mainly Ki67 negative, stem cells are reported to be Ki67 positive [de Sousa e Melo et al., 2017; Shimokawa et al., 2017]. At the same time, it was shown that there are rare KRT20<sup>+</sup> Ki67<sup>+</sup> and *Lgr5*<sup>+</sup> Ki67<sup>-</sup> populations and more prevalent *Lgr5*<sup>-</sup> KRT20<sup>+</sup> Ki67<sup>+</sup> cells which might comprise intermediate progenitor cells residing in the tumor [Shimokawa M. et al., 2017; Baker A.M. et al., 2015]. The fact that both differentiated cells and stem cells are endogenously present in tumors already demonstrates that tumor cells are plastic. Generally, the vast majority of differentiated cells in tumors either disappear or stay as single-cell colonies over a long period of time, while stem cells and their progenitors proliferate giving rise to colonies that ultimately contain differentiated cells. At the same

time, as it was shown by the Sato group in lineage tracing experiments, a small proportion of KRT20<sup>+</sup> differentiated cells (0.7%) is also able to give rise to colonies of tumor cells that are positive for Ki67 staining and most likely express Lgr5. Moreover, various stress conditions expand the capacity of differentiated cells to give rise to Lgr5<sup>+</sup> cells which were demonstrated by specific ablation of Lgr5<sup>+</sup> cells in tumors [Shimokawa et al., 2017]. It was suggested that transformation might be triggered by the freed stem cell niche after Lgr5<sup>+</sup> cell elimination. Interestingly, in *in vitro* conditions, KRT20<sup>+</sup> and Lgr5<sup>+</sup> cells had a similar clonogenic capacity which highlights the intrinsic stemness potential of both cell populations. These findings highlight the similarity between the tumor architecture and the crypt-villus hierarchy in normal intestinal epithelium where it was also shown that under stress-induced conditions (targeted ablation, DSS treatment) merely all differentiated cell populations are able to give rise to Lgr5<sup>+</sup> stem cells [van Es JH. Et al., 2012; Buczacki SJ. et al., 2013; Tetteh PW. Et al., 2016; Jadhav U. et al, 2017; Yu S. et al., 2018].

A deeper understanding of cancer cell plasticity and the connection between Lgr5<sup>+</sup> stem cells and Lgr5<sup>-</sup> differentiated cells (which are highly enriched for differentiation markers including KRT20 [Fumagalli A. et al., 2020]) was achieved with the works from Frederic J. de Sauvage and Jacco van Rheenen groups. The Sauvage group demonstrated that Lgr5 cells are indispensable for the growth of metastasis but not for the primary tumor. Ablation of Lgr5<sup>+</sup> cells lead to a drastic decrease in metastasis formation but didn't cause tumor regression due to a constant replenishment of the Lgr5<sup>+</sup> stem cell pool by Lgr5<sup>-</sup> cells [de Sousa e Melo et al., 2017]. The van Rheenen group expanded this knowledge by establishing that Lgr5<sup>-</sup> cells are predominantly responsible for metastatic dissemination from primary tumor but compelled for transformation back to Lgr5<sup>+</sup> cells upon arrival to the metastatic niche for fueling metastatic tumor outgrowth. Using growth factor-free *in vitro* experiments, the authors also showed that the ability of differentiated cells to give rise to stem cells is a cell-intrinsic feature rather than dictated by stem cell niche factors [Fumagalli A. et al., 2020].

Altogether, colorectal tumors are heterogeneous and composed of cells with different proliferation and differentiation statuses. Differentiated cells are endogenously present in colorectal tumors. Cancer stem cells are the driving force of primary and metastatic tumor development. Differentiated cells preferentially vanish from the tumor site during its development. Moreover, it's well documented that patients with well-differentiated tumors have better overall survival rates compared to poorly differentiated colorectal cancers [Xiao H. et al., 2013]. Stem cells, their progenitor cells, and

differentiated cells are plastic and can transform one into another with a certain frequency even in homeostatic conditions. This means that stem cells are able to respond to differentiation-inducing signals and that the ratio of stem cell to differentiated phenotypes in CRC could be altered exogenously (e.g. by using drugs).

These remarkable discoveries highlighted several important aspects for introducing differentiation therapy. Differentiation therapy is one of the approaches to treat CRC tumors which involves induction of cancer stem cell differentiation which blocks tumor development and metastasis. This therapy enriches tumors with differentiated cells which would be eliminated spontaneously. One predicament in this therapy is cell plasticity: it's known that differentiated cells are able to transform back to stem cells, especially in those cases when stem cells are actively eliminated. Thus, successful differentiation therapy should simultaneously induce cancer stem cell differentiation and block signals which allow differentiated cells to revert back to the stem cell state.

### 1.5.3 Examples of differentiation therapy in colorectal cancer

The first step in developing a differentiation therapy for CRC was previously done in our lab [Ordóñez-Morán P. et al., 2015]. This research uncovered the mechanism of Hoxa5 driven differentiation in colorectal cancer using mouse and human experimental models. The Hoxa5 gene drives terminal differentiation of the mouse small intestinal epithelial cells. It negatively regulates Wnt signaling in differentiated cells and is missing in stem cells. Hoxa5 gene expression is downregulated in CRC tumors, but its induction leads to tumor differentiation, and inhibits progression and metastasis. Hoxa5 gene expression can be activated by retinoic acid, a well-known inducer of differentiation. Thus, this research clearly demonstrated a mechanistic insight into the tumor differentiation process in CRC and potential drugs which could target it.

Hoxa5 is not the only example of tumor suppressor genes involved in differentiation. Id2 is another protein with similar activity: depending on the environmental context it can act as a repressor or an activator of tumor growth. For a few decades, it was believed that Id2 is an inhibitor of differentiation. Indeed, in the WT small intestine it is highly expressed at embryonic stages of development however reduced in the adult [Lasorella A. et al., 2014]. Later investigations showed that Id2 negatively regulates Wnt signaling and its knock-out caused increased numbers of Lgr5<sup>+</sup> cells in the small intestine at embryonic stages and spontaneous neoplasia in the adult [Nigmatullina L. et al., 2017]. Thus, Id2 is showing similar effects as Hoxa5 in affecting intestinal differentiation. Further steps for making Id2 druggable are still missing.

NDRG2 gene is another example of a tumor suppressor gene that facilitates the differentiation of CRC [Shen L., et al., 2018]. Researchers showed that NDRG2 promotes differentiation through silencing of Skp2 E3 ligase, which leads to the stabilization of p21 and p27. CDK inhibitors p21 and p27 regulate cell cycle progression which might underlie NDRG2-driven differentiation. Mechanistically, NDRG2 abrogates  $\beta$ -catenin nuclear translocation which reduces the activity of TCF/ $\beta$ -catenin driven Skp2 transcription. Clues for making NDRG2 expression druggable are still missing and remain to be discovered.

Overall, few examples provided insights into the differentiation process in CRC, but only the Hoxa5-induced differentiation study [Ordóñez-Morán P. et al., 2015] provided possible drug candidates which could be used for CRC therapy. Thus, the discovery of differentiation-inducing drugs for potential use in the clinic is still challenging but of high interest.

## 1.6 Colorectal cancer models

Progress in cancer research could not be achieved without proper models to study it. Thus, experimental need boosted the development of corresponding colon cancer models. Cell lines were one of the first experimental models of colon cancer and some of them were established already in the 1970-1980s (HT29, HCT116, Caco-2, SW480, LoVo) according to a PubMed search and available publications [e.g. von Kleist S. et al., 1975; Brattain M.G. et al., 1984]. Many of those are immortalized cell lines derived from human adenomas, carcinomas, or adenocarcinomas. Nowadays there are nearly 90 different cell line models of human colon cancer available at the American Type Culture Collection (ATCC) harboring different spectra of genetic mutations representing different subtypes of CRC. Those very early but also limited cancer models nevertheless provided a huge amount of information for the development of the field and are still widely used nowadays.

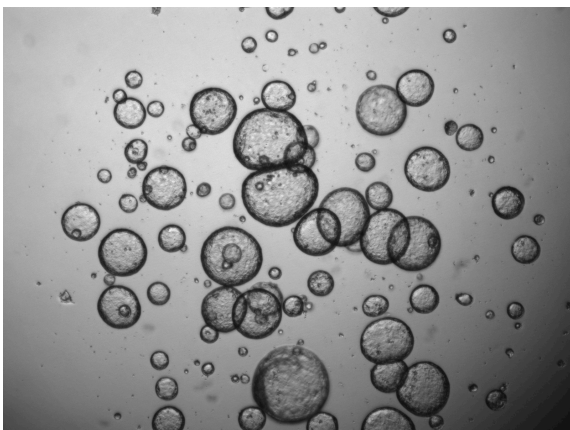
Further research development demanded more representative models of CRC than simple 2D cultures, which would include the complexity of an actual living organism with its homeostasis and active immune system for a better understanding of the interaction between cancer cells and the environment, and for better drug testing. Thus, the first genetically engineered mouse model (GEMM) of colorectal cancer was introduced in 1990 – the APC<sup>Min</sup> mouse model (Min – multiple intestinal neoplasia) [Moser A.R. et al., 1990]. It's a mouse model harboring a heterozygous mutation in the APC gene in codon 850 (APC protein has 2843 amino acids) which results in a truncating mutation and APC protein malfunction. It was obtained via a N-Ethyl-N-nitrosourea (ENU) mutagenesis

screen. Homozygous APC deletion is embryonic lethal, and it is therefore not possible to obtain homozygous APC<sup>Min</sup> mice. For the phenotypic transformation of intestinal cells to malignant polyps, a spontaneous loss-of-function (LOF) mutation in the second allele has to occur, and usually, APC<sup>Min</sup> mice start to develop polyps at the age of 2-3 months with up to 50 polyps developed through life [Moser A.R. et al., 1990; Zeineldin M. et al., 2013]. Tumor development in these mice is driven by APC LOF and this model represents a major cohort of adenomatous pathway tumors (60-70% of CRC cases). The APC<sup>Min</sup> mouse model is a great tool for studying the initial steps of tumor development which for the first time allowed identification and manipulation of genes involved in tumor development *in vivo* [Zhang L. et al., 2017]. Unfortunately, there are several limitations of that model: first of all, polyps develop at a relatively young age which doesn't reflect the tumor biology in humans where cancer arises mostly in older patients; second, the majority of the tumors develop in the small intestinal tract, while the majority of human cancers develop in the large intestine; third, the LOH mutation occurs spontaneously in the second allele of APC which could lead to some level of heterogeneity between the tumors. Moreover, one cannot exclude the possibility of mutations in other genes which leads to APC silencing. There is no possibility to know upfront the location and the nature of the second mutation unless a deep analysis of each tumor is performed. Fourth, these tumors typically do not progress to more advanced adenocarcinomas which are observed in the majority of human cases, the model thereby only reflects early stages of tumor development. Finally, the average lifespan of APC<sup>Min</sup> mice is 4 months which makes it tedious to maintain the mouse line.

The next breakthrough in designing adequate mouse models for studying CRC was the implementation of the Cre-loxP system [Lakso M. et al., 1992; Buertin F. et al., 2020] combined with the tamoxifen-inducible Cre-ERT2 system [Indra A.K. et al., 1999]. Several Cre reporters were introduced including Fabp1-Cre, Villin-Cre, and Cdx2-Cre [Saam J.R. et al., 1999; Madison B.B. et al., 2002; El Marjou F. et al., 2004; Hinoi T. et al., 2007; Feng Y. et al., 2013]. Villin and Fabp1 expression is not restricted to the large intestine: while Villin is predominantly detected in the small intestine with lower expression levels also detected in the large intestine, Fabp1 is also expressed in the distal small intestine. Cdx2 has a more specific pattern of expression – it expresses in large intestine and therefore is a more relevant reporter for studying CRC tumorigenesis. Combining such alleles with relevant loss-of-function (LOF) or gain-of-function (GOF) genetic models resulted in powerful tools for studying CRC. The most common mutations (and pathways they belong to) in colon cancer are APC (Wnt), TP53 (Apoptosis), KRAS (MAPK), SMAD4



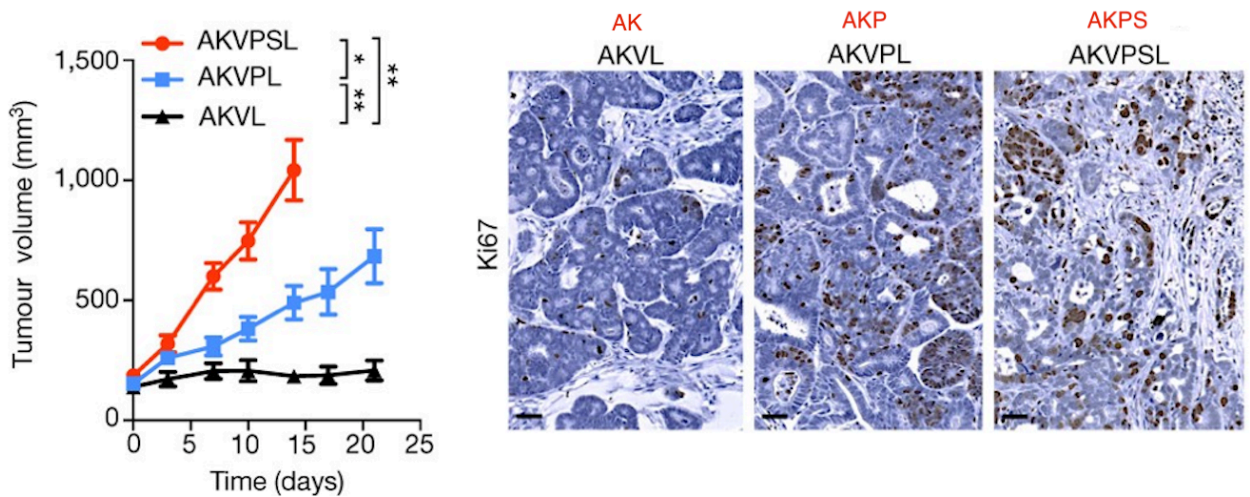
(TGF- $\beta$ ), PIK3CA (PI3K/AKT). According to the adenoma-carcinoma theory of tumor development by Fearon and Vogelstein [Fearon, E. R. et al., 1990] the acquiring of mutations in tumors appear in the following sequence – APC, KRAS, P53; and later - mutations altering TGF-  $\beta$  and PI3K/AKT signaling, such as SMAD4 and PIK3CA, leading to enhanced chemoresistance and metastatic properties of tumor cells. Therefore, several inducible mouse models were created to unravel the consequences of those transformations. The  $Apc^{Min/+};Kras^{LSLG12D/+};Vil1^{Cre}$  model was introduced to mimic early stages of polyps development in the colon. Indeed, due to the activating KRAS mutation, mice harboring these alleles (unlike  $APC^{Min}$  model) developed multiple polyps in the colon mimicking the first steps of tumor initiation [Enquist, I. B. et al., 2014]. More advanced models included p53 inactivating mutations which are usually acquired by tumors at the carcinoma stage, resulting in the APC KRAS p53 phenotype which reflects more advanced stages of the adenomatous pathway [Martin S.E. et al., 2013; Ragusa S. et al., 2014]. Interestingly, a study comparing APC p53 (AP) model versus APC p53 KRAS (AKP) model, showed that adding the KRAS activating mutation increases MAPK and PI3K/mTOR signaling as accessed via p-MEK, p-ERK, p-AKT, and p-S6 stainings. At the same time, Wnt signaling was not altered but sustained at high levels in both models (distinguished by  $\beta$ -Catenin staining) [Martin S.E. et al., 2013]. Several studies used mouse-derived organoids (Fig. 11) to generate desirable mutation patterns and used subsequent organoid transplantation into a mouse (subcutaneous, orthotopic injections) for studying CRC development [Drost J. et al., 2015; de Sousa e Melo et al., 2017; Roper J. et al., 2017].



**Figure 11.** Mouse cancer organoids harboring the  $Apc^{fl/fl};Kras^{Lsl-G12D};Tp53^{fl/fl}$  (AKP) genotype.

Comparison of AKP and AKPS (AKP + Smad4) models showed that both models develop highly proliferative tumors but while AKPS tumors develop as invasive carcinomas, the majority of AKP tumors are adenomas [Drost J. et al., 2015], though other studies showed that AKP mutations can progress to carcinomas [Martin S.E. et al.,

2013; Dionellis S.V. et al., 2021]. The discrepancies between the studies can be explained by the timing of tumor growth. AKPS organoids have much higher levels of genomic instability compared to AKP and AK organoids. Subcutaneous AKPS tumors display exponential growth while AKP tumors have more moderate growth dynamics (Fig. 12).



**Figure 12.** Subcutaneous tumors harboring AK, AKP and AKPS phenotypes. Organoids were generated from  $Apc^{Min/+};Kras^{LSLG12D/+};Villin-Cre;Lgr5^{DTR/eGFP}$  colonic tumors and further transfected with gRNAs targeting p53 and Smad4. Left panel: tumor growth; right panel - ki67 staining [adjusted from de Sousa e Melo et al., 2017].

Moreover, assessment of Lgr5 expression indicated that AKP and AKPS tumors have similar stem cell content (around 15-20%) [de Sousa e Melo et al., 2017]. Authors showed that only the AKPS model developed spontaneous metastasis in the orthotopic setting in a time frame of 6 weeks after injection [de Sousa e Melo et al., 2017]. At the same time, other research including ours, showed that the AKP model (Cdx2 tamoxifen-inducible) developed metastasis after 25 weeks after low dose tamoxifen injection. This is in line with the lower growth rate of the AKP model which probably causes delayed metastatic potential.

The AKP models reflect the adenomatous pathway of CRC development, which is the most common type of CRC accounting for 70% of all CRC cases. Other models of CRC including BRAF mutations, MMR-D, or MSI status were successfully implemented [Carragher L.A. et al., 2010; Coffee E.M. et al., 2013; Rad R. et al., 2013] and a detailed description can be found elsewhere [Buertin F. et al., 2020].

## 1.7 Colorectal cancer patient-derived organoids

Patient-derived organoids can be regarded as a special type of CRC modeling platform. For decades, research aimed to provide pivotal information for decision-making for patient treatment, including cancer patients. However, despite the extensive research

in the field we still know very little about the mechanisms, heterogeneity, metastatic process, and acquired resistance in colorectal cancer. As a consequence of lack of knowledge, only a limited number of drugs have passed clinical trials through past decades and have been approved by the FDA for use against metastatic CRC. For example, the first targeted therapy drug (Cetuximab, targeting EGFR signaling) was approved by the FDA only 15 years ago, and its effect was restricted only to a cohort of KRAS wt patients with moderate effect on overall survival [Cutsem E.V. et al., 2009]. This fact clearly illustrates that the scientific community is far away from fully understanding the mechanisms of this heterogeneous disease. With that, patient-derived organoids (PDOs) can provide important information about inter- and intra-patient heterogeneity and will allow to analyze the biology and drug-sensitivity of genetically diverse tumors.

It has been shown that PDOs (and PDXs models) recapitulate the histopathological features and mutational spectrum of original tumors [Cho Y.B. et al., 2014; van de Wetering M. et al., 2015; Bruun J. et al., 2020; Vlachogiannis G. et al., 2018; Mullinis S.C. et al., 2019; Fujii M. et al., 2016; Roerink F.S. et al., 2018; Ooft N.S. et al., 2019] and therefore provide representative models to study CRC, especially for precision medicine decision-making. The establishment of PDOs is very efficient and reaches 90% while the efficiency of PDX establishment is around 70% [Mullinis S.C. et al., 2019; van de Wetering M. et al., 2015]. The research utilizing PDOs unraveled a high level of inter-patient heterogeneity including morphological heterogeneity, genetic heterogeneity, and heterogeneity in terms of drug response, which may be related to the different spectra of acquired mutations for each individual tumor [van de Wetering M. et al., 2015; Fujii M. et al., 2016; Vlachogiannis G. et al., 2018]. The application of PDOs seems to be a useful tool for predicting tumor response to standard drug treatments [Ooft N.S. et al., 2019; Narasimhan V. et al., 2020; Vlachogiannis G. et al., 2018] which would facilitate and speed up the therapy decision-making process.

One of the advantages of organoids for clinical decision-making is that researchers may be able to model therapy resistance, identify and investigate resistant clones and predict the best second line therapy for each individual patient, or drug combinations that would most successfully treat the whole tumor or metastatic lesions. Perhaps the use of more personalized approaches emerging in recent years will be a breakthrough in the treatment of metastatic CRC. Using patient-derived organoids (PDOs) is clearly a beneficial strategy.

## 1.8 Screenings in colorectal cancer

PDOs, PDXs models, mouse-derived colon cancer organoids and 2D cell lines were used in different formats for high-throughput screening. Here I collected the most recent advances in reporter-based high-throughput techniques, novel imaging-based assays, and experiments on drug testing for personalized decision making.

### 1.8.1 Reporter-based assays in cell lines

Several high-throughput drug-screening platforms were recently established to find drugs targeting CRC. The advantage of the 2D cell line assays is that typically they are capable to screen thousands of drugs based on the readout of specific reporter activity.

Zhan et al. looked for drugs modulating Wnt signaling in CRC using three cell lines harboring mutations in KRAS, APC and  $\beta$ -catenin— HCT116, SW480, and DLD1 [Zhan T. et al., 2019]. They used the TCF-Wnt luciferase reporter to identify small molecules affecting Wnt signaling from a cohort of 2399 compounds. Unexpectedly, they discovered that MEK inhibitors activate Wnt signaling by downregulating AXIN1, a rate-limiting enzyme of the  $\beta$ -catenin destruction complex. Downregulation of AXIN1 mediated by ERG1 (known downstream target of MAPK signaling) led to activation of Wnt signaling. MEK inhibition leads to induction of stem cell phenotypes in CRC which is supported by recent studies on plasticity in CRC [Shimokawa, M. et al., 2017; de Sousa e Melo et al., 2017; Fumagalli A. et al., 2020].

Another study presented a high-throughput screening platform for identifying CRC targeting drugs by implementing a three-reporter system simultaneously monitoring the transcriptional activity of TCF/LEF, NF- $\kappa$ B, and NRF2 (transcription factor regulating antioxidant genes). Wnt signaling (TCF/LEF reporter) is responsible for tumor proliferation, NF- $\kappa$ B is a sensor of inflammation which is usually highly activated in tumors and oxidative stress response is an important survival mechanism in CRC. Thus, authors looked for the drugs inhibiting TCF/LEF and NF- $\kappa$ B reporters, or drugs activating NRF2 expression as a measure of increased oxidative stress. They screened the library of 1280 drugs and identified 8 hits, one of which is itraconazole, a drug which authors presume is involved in cholesterol trafficking. Moreover, the authors demonstrated that itraconazole suppresses the formation of polyps in APC<sup>Min</sup> mice [Miyamoto S. et al., 2020].

Another study implemented an alkaline phosphatase (ALP) reporter system to identify differentiation-inducing drugs in colorectal cancer [Plotnikov A. et al., 2020]. ALP is a differentiation marker in enterocytes in the normal intestine and is highly expressed in well-differentiated tumors projecting better survival. At the same time, ALP is lowly expressed in more aggressive, poorly-differentiated CRCs. The authors analyzed 5700

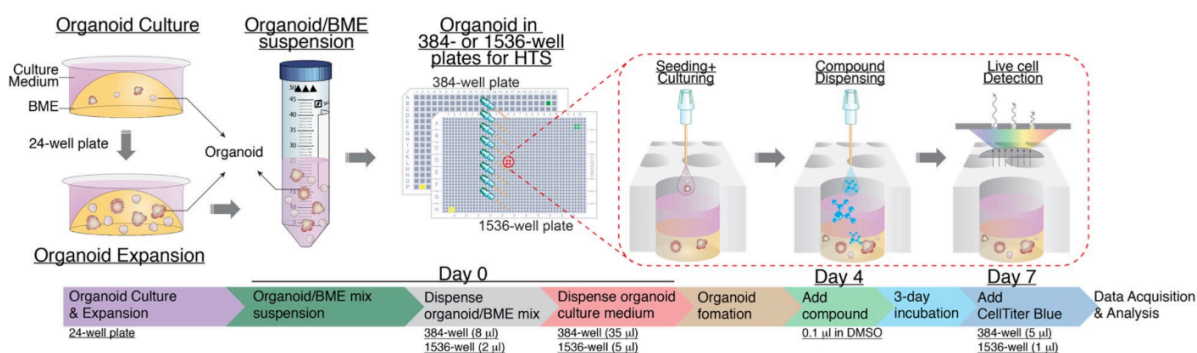
drugs for ALP activity in two cancer cell lines HT29 and HCT116 and one normal colon epithelial cell line CCD-841 (with already high levels of ALP activity) for comparison. They found that MS023, an inhibitor of the histone methylase PRMT1, significantly upregulated ALP activity in cancer cell lines but didn't change ALP levels in normal epithelial cell lines. The authors confirmed the differentiation signature induced by PRMT1 inhibitors using RNA-seq and showed its high efficacy against CRC in HT29 xenograft mouse model. It remains to be seen if this drug can be used *in vivo* in more complex CRC models.

### 1.8.2 High-throughput screenings in 3D cultures

The next step in generating more representative screening platforms for CRC targeting drugs was the implementation of 3D organoid- or spheroid-based screenings.

Folkesson E. et al. reported a high-throughput screening in 3D spheroids cultures of several cell lines and compared those readouts to results obtained from 2D screening. Moreover, they were interested in analyzing the effects of 21 clinically relevant drug combinations in 2D and 3D cultures. They used an ATP-based viability screen in HCT116, HT29, and SW620 cell lines or spheroids. The authors discovered that drug synergy more frequently works in cell lines than spheroid cultures. They also found out that 3D spheroids are generally more sensitive to MEK inhibitors and drug combinations containing MEK inhibitors [Folkesson E. et al., 2020].

Similarly, Du Y. et al established ultra-high-throughput screening in organoids based on cell viability using CellTiter Blue. They screened 2036 drugs in a APC<sup>-/-</sup>; KRAS<sup>G12D</sup> human colon organoid model in 384 well plate format. The technical difficulty of this type of screening is that organoids are growing in solidified drops of gel matrixes (such as Cultrex, basement membrane extract gel). Further, they derived optimized conditions for organoid preparation for screening in the 1536 well format (Fig. 13) [Du Y. et al., 2020].



**Figure 13.** Experimental set-up of high-throughput drug screening in colon cancer organoids [taken from Du Y. et al., 2020].

The proposed high-throughput miniaturized model could be a useful tool in *ex vivo* analysis of patient tumor organoids for therapy decision-making.

### 1.8.3 High-content imaging-based screenings

The high-throughput platforms described above have clear advantages for the screening of many thousands of conditions, but due to the implementation of only one to three readouts, the information obtained with those assays requires additional orthogonal experiments which would allow looking at mechanistic insights of drug action. A step further in developing a high-content screening platforms was achieved by implementing the microscopy-based analysis of treated 3D organoids. Works by the Liberali lab and the Lutolf lab are of particular interest [Lukonin I. et al., 2020; Brandenburg N. et al., 2020].

The Liberali lab established a microscopy-based platform for phenotypic characterization of intestinal organoids upon drug treatment. They generated multivariate profiles of hundreds of organoids based on their morphology and intensity and distribution of DAPI (DNA staining), aldolase B (enterocytes), and lysozyme (Paneth cells) signals accessed by immunofluorescent staining. They distinguished 15 phenotypic classes of organoids and screened 2789 drugs observing 301 drugs that strongly induced one of the monitored phenotypes. They found that an antagonist of retinoid X receptor (RXR)- $\alpha$  enriched organoids in a regenerative state. They confirmed the regenerative capabilities of RXR- $\alpha$  antagonist by performing *in vivo* experiments in irradiation-induced colitis regeneration [Lukonin I. et al., 2020].

Recently, the Lutolf lab developed a high-throughput automated system for drug screenings consisting of engineered cell culture microwells of 3D spheroids [Brandenburg N. et al., 2020]. Engineered 96-wells contain around a hundred individually separated spheroids located each at the bottom of a microwell. These 3D cultures are grown from a cell suspension in 2% matrigel or diluted synthetic hydrogel which in theory allows their use for downstream analysis. The system is automated for media change and drug dispersion. High reproducibility of growth allowed authors to perform a proof of concept high-content imaging-based screening on HCT116 derived spheroids. Using calcein-AM ethidium homodimer-1 fluorescent labeling for live and dead cells they were able to identify many HCT116 targeting drugs.

### 1.8.4 CRISPR screenings in organoids

Organoids were used successfully in CRISPR screening approaches. Several recent studies applied large libraries of pooled gRNAs in human colon organoids

harboring APC<sup>-/-</sup> or APC<sup>-/-</sup>;KRAS<sup>G12D</sup> genotypes [Michels E.B. et al., 2020; Ringel T. et al., 2020]. Two studies focused on resistance to TGF-β stimulation, an effect seen in CRC during later stages of tumor development.

Michels E.B. et al. used pooled CRISPR-Cas9 screens in APC<sup>-/-</sup>;KRAS<sup>G12D</sup> human colon organoids and identified TGFBR1, TGFBR2, and SMAD2 as tumor suppressors. Further, they applied a library of transduced organoids in xenotransplantation experiments and looked for clone enrichment after 12 weeks of tumor growth. Interestingly, they found TGFBR2 as the strongest hit confirming the strong tumor-suppressive role of TGF-β in the tumor microenvironment. Other hits included TP53, ATM-2, SMAD4, STK11, and SMARCA, known drivers of CRC tumorigenesis. At the same time, the analysis showed an enrichment of control gRNAs in surviving tumor cells indicating a high rate of false-positive or presence of passenger gRNAs in tumor cells.

Ringel T. et al. modified their approach in analyzing single organoids instead of using pooled screening followed by bulk RNA-seq. This improvement led to more robust results eliminating the noise generated during the selection procedure. In *in vitro* screens authors looked for resistance to TGF-β treatment in WT organoids and were able to identify known TGF-β pathway players TGFBR1, TGFBR2, and SMAD4 as positive hits which accounted for 96% of all hits. As TGF-β associated mutations arise at later steps of CRC development the authors decided to analyze the synergy between TGF-β signaling mutations and APC<sup>-/-</sup> and APC<sup>-/-</sup>; TP53 phenotypes. Interestingly, they found that only 65% and 19% of gRNAs were targeting known TGF-β components, correspondingly, in APC<sup>-/-</sup> and APC<sup>-/-</sup>; TP53 organoids. The rest of the genes were not related to TGF-β which highlights higher potency of APC, TP53 mutated tumor cells in overcoming TGF-β resistance, or the existence of synergy with alternative pathways which could overcome that resistance. Among those genes were known chromatin modifiers ARID1A, ARID2, SMARCA4, and KEAP1, also highly mutated genes in CRC [Ringel T. et al., 2020].

The application of organoids in CRISPR screening highlights the possibilities that organoid technology offers to study CRC. It is already now a useful tool in understanding the biology of cancer development and could be especially helpful for the identification of resistance mechanisms.

### 1.8.5 Screenings in patient-derived organoids

The number of research projects utilizing PDOs for CRC treatment has been growing rapidly in the last 10 years. CRC is a very heterogeneous disease with various mutation combinations presented in epithelial tumor cells. This leads to the existence of several types of CRCs which have different physiology, occurrence, survival time, and response rates to particular drug treatments. Eventually, almost every patient can have a unique combination of acquired mutations which makes it very difficult to make predictions on treatment responses based on conventional CRC models. In that sense, PDOs can provide an excellent opportunity for modeling personalized treatments for such patients.

The majority of screening approaches in PDOs use only cell viability as a main readout in their experiments. Boehnke K. et al. reported a high-throughput set-up for drug screening in PDOs. They developed a protocol for tumor cell expansion and seeding in 384-well plate format which allows screening for many drug treatment conditions. Authors used ATP consumption-based screening for accessing cell viability measured by relative luminescence units (RLU) [Boehnke K. et al., 2016].

Other studies concentrated their effort in collecting and characterizing large biobanks of CRC PDOs and tried to distinguish the response rates to the drug treatments between the different cohorts of PDOs based on their mutational status. The Clevers lab performed a proof of concept study with 20 established PDOs. The most common mutations in those organoids were APC, TP53, KRAS, PIK3CA, and SMAD4, PDOs were separated as hypermutated or non-hypermutated according to the number of acquired mutations. The authors screened a small library of 83 compounds to determine PDOs-specific responses. Interestingly, the clustering of the PDOs-based IC50 values predict the existence of 3 clusters, but there was no overlap with the previously identified hypermutation status. Further molecular dissection of the potential correlation between oncogenic mutation and drug sensitivity discovered only a few, already known aspects about the CRC tumors: TP53 WT PDOs were sensitive to the MDM2 inhibitor nutlin-3a; KRAS mutant PDOs were resistant to anti-EGFR treatment [van de Wetering et al., 2015]. Another study by Schuette M. et al. reported a collection of 35 PDOs and performed expression profiling of those tumor cells. They also analyzed gene expression changes in response to several drug treatments including 5-FU, cetuximab, afatinib, and AZD8931. They reported a new gene expression signature that better predicts sensitivity to EGFR treatment compare to KRAS/RAF mutational status [Schuette M. et al., 2017]. Another study used a small cohort of PDOs with either KRAS WT or KRAS mutant status for accessing the response of tumor cells to the combinations of drug therapies targeting



RAS pathway [Verissimo et al., 2016]. They showed that simultaneous inhibition of EGFR, MEK, and ERK pathways doesn't cause tumor cell death but rather induces their quiescent state. This finding was confirmed in a mouse xenotransplant model using pan-HER/MEK inhibitors. Similarly, Vlachogiannis G. et al. performed a small drug screen on a panel of 19 PDO samples with 55 drugs from clinical trials [Vlachogiannis G. et al., 2018]. Given the small cohort of samples, the authors could not link the responses of the PDOs to drug treatments and their mutational status.

The idea of linking mutational status of the organoids to their responses to drug treatment is a promising approach in finding predictive biomarkers for therapy decision-making. As explained above, several attempts with proof of concept studies were carried out in this direction. One of the limitations in these studies is the small sample size, which doesn't allow the discovery of significant links for mutation - drug response correlations. Moreover, many of those studies used cell viability as a main readout for performing comparisons and clustering. Considering that the main goal of CRC therapy is to find a successful treatment for a given patient, using cell viability readouts might be useful, at the same time it doesn't necessarily lead to the discovery of predictive responses due to complex interconnections of mutational burden. Moreover, depending on the experimental setup this type of experiments cannot always explain the resistance mechanisms of a given drug treatment. Instead, usage of gene expression signatures, which more closely reflect tumor state, might be more successful in predicting PDO responses to drug treatments [Schuette et al., 2017].

### 1.8.6 RNA-seq based screening platforms

Currently, there are no RNA-seq based platforms for performing high-throughput screenings in organoids due to technical difficulties (mainly organoids being trapped in solidifying drops of matrigel) and the high cost of RNA-sequencing. Though several groups established bulk RNA-seq based screening methods for cell lines [Ye C. et al., 2018; Bush E.C. et al., 2017], the high cost of whole transcriptome sequencing is not compatible with high-throughput assays on a large collection of samples. Several targeted approaches were suggested [Simon J.M. et al., 2019; Teder H. et al., 2018] but were also not compatible with high-throughput drug screening in organoids.

## 1.9 Prospects of drug screening methods in organoids

For the last decades, only a limited number of drugs for CRC treatment successfully passed various stages of clinical trials and were approved by the FDA. At the same time,

many screens in CRC cell lines, tumor-originated spheroids, and PDOs were carried out. There may be several reasons for this discrepancy. First of all, the use of insufficient models e.g. 2D cell lines for predicting tumor response to drug treatment. Tumors have a much more complex architecture with the presence of many cell types and active signaling pathways which cannot be modelled by planar 2D cell lines. Second, the majority of screening platforms rely on the activity of a single reporter system (such as various viability assays, or Wnt signaling activity, or differentiation reporters) in predicting the outcomes in patients. Those assays fail to elucidate resistance mechanisms upon drug treatment and therefore cannot suggest corresponding drug combinations which would successfully overcome resistance. Of course, there are far more reasons why the therapy could fail, but the majority of the cases linked either to drug toxicity or its inefficiency. We believe that understanding of the mechanisms which drive the resistance would help solving both problems: first of all, it would allow to find a more suitable efficient therapy by combining several drugs targeting different pathways, secondly, drug combinations could allow the use of lower doses of corresponding drugs which may reduce the drug toxicity. Thus, currently, there is a need for such drug screening platforms which would quantify and evaluate complex cellular phenotypes and cellular alterations upon drug treatment. One of the possibilities is to use targeted RNA-seq based high-throughput platforms as they are compatible with the screening in organoids and enable identification of the mechanism of drugs action.

## 1.10 Objectives of the project

### 1.10.1 First part of the thesis

The main goal of the first project was the development of a high-throughput RNA-seq based platform for drug discovery in colorectal cancer. Application of targeted RNA-seq and optimization of the library preparation protocols should overcome the challenges linked to the high cost of RNA-seq profiling of thousands of samples. The large amount of gene expression data available for intestinal cell types, organoids, and colorectal tumors should be advantageous in designing highly specific gene panels for targeted analysis.

Further, we aimed to discover differentiation-inducing drugs for CRC therapy using our targeted RNA-seq based platform and a library of FDA-approved drugs.

### 1.10.2 Second part of the thesis

In the second part of the thesis and in collaboration with the group of Prof. Halazonetis, we investigated tumor development during the early stages of CRC development. According to the adenoma-carcinoma multistep carcinogenesis process proposed by Fearon and Vogelstein tumor cells acquire mutations in a subsequent manner during tumor development. At the same time, it's known that the number of mutations detected in most tumors is too high taking into account the average rate of acquisition of mutations of normal cells. Therefore, there was the hypothesis that the transformation process in tumor cells leads to a higher capability of such cells to acquire mutations at an enhanced rate. In this study, we investigated the role of p53 and KRAS mutations for the frequency of mutations in the APC<sup>lof</sup>;KRAS<sup>G12D</sup>;TP53<sup>lof</sup> mouse model of CRC.

# Chapter 2: Results

## 2.1 Overview

This thesis results chapter consists of two parts. Both parts are published articles. In both parts of the projects, we study colorectal cancer biology. While in the first project we established the platform that allows us to identify and track differentiation trends induced by drug treatments in colorectal cancer organoids, in the second project we look more into colorectal cancer development: how cancer driver genes (p53 and KRAS) affect the rate of mutation acquisition by tumors.

In the first project, we developed a high-throughput high-content targeted RNA-seq based platform (TORNADO-seq) for monitoring the expression of large gene signatures in intestinal organoids which allowed us to quantify and evaluate complex cellular phenotypes and cellular alterations upon drug treatment. This approach allows us to identify differentiation-inducing drugs. We applied TORNADO-seq in drug screening in wt and cancer ( $APC^{lof};KRAS^{G12D};TP53^{lof}$ ) intestinal organoids.

In the second part of the thesis we investigated the mechanisms underlying the accumulation of genetic alterations in CRC tumors. We performed exome sequencing of single cell-derived organoids obtained from tumors with either  $APC^{lof}$  (A)  $APC^{lof};KRAS^{G12D}$  (AK) or  $APC^{lof};KRAS^{G12D};TP53^{lof}$  (AKP) mutation genotypes that reflect early stages of CRC development. Using low-dose tamoxifen injections and  $Cre^{ERT2}$  inducible system we were able to generate colonic tumors developing over long period of time (more than 5-6 months). We analyzed the mutational burden of those tumors and corresponding liver metastasis and compared it to adenoma-staged, more benign, early tumors.

2.2. The first part of the thesis results is presented as a published paper:

**“High-content, targeted RNA-seq screening in organoids for drug discovery in colorectal cancer.”**

Cell Reports, 2021. 20;35(3):109026. doi: 10.1016/j.celrep.2021.109026.

**Authors:**

**Maxim Norkin**<sup>1</sup>, Paloma Ordóñez-Morán<sup>2</sup>, Joerg Huelsken<sup>3</sup>.

**Authors' affiliation:**

<sup>1</sup>Swiss Institute for Experimental Cancer Research (ISREC), École Polytechnique Fédérale de Lausanne-(EPFL-SV), 1015 Lausanne, Switzerland.

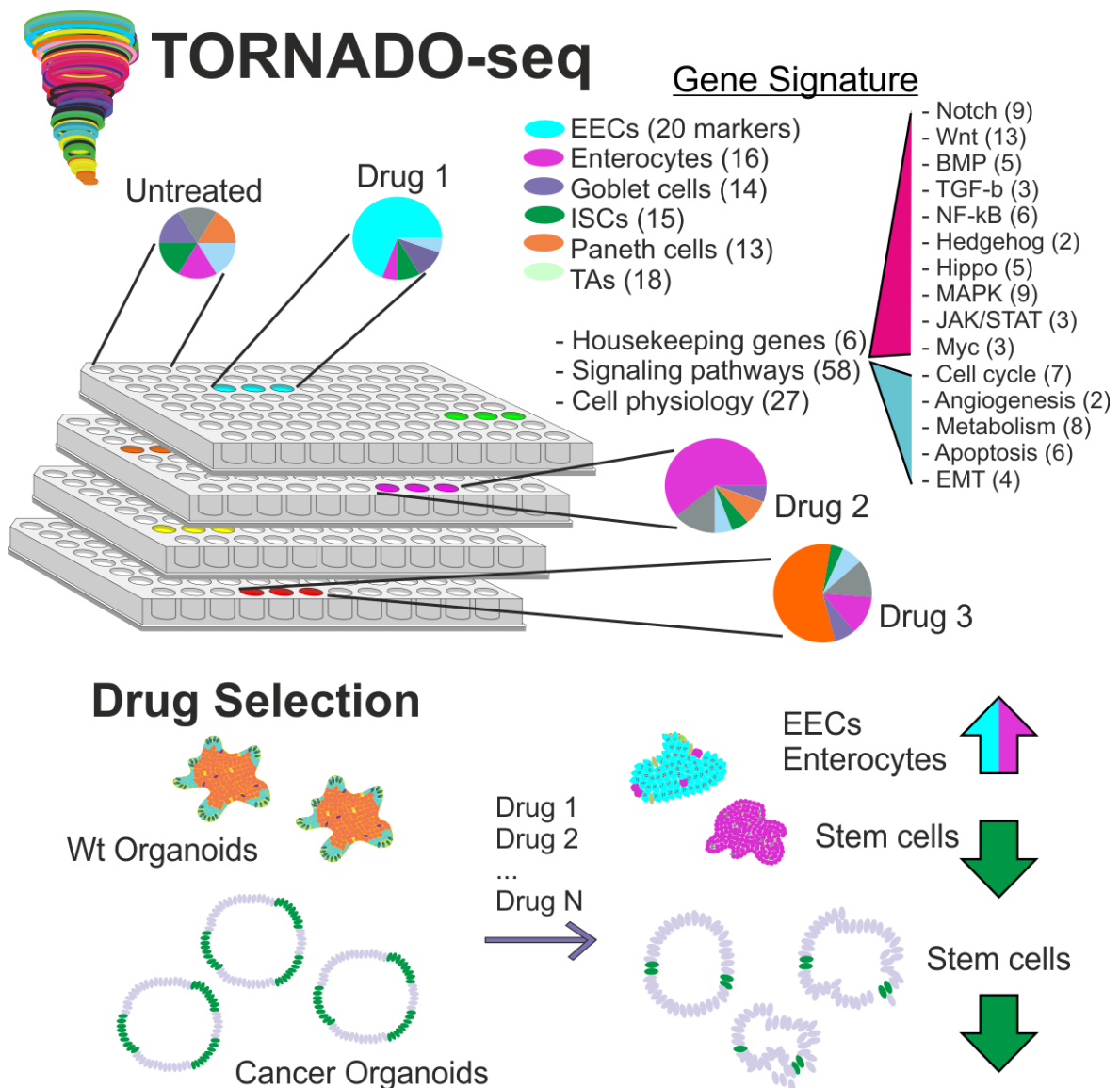
<sup>2</sup>Department of Cancer and Stem Cells, School of Medicine, Biodiscovery Institute, University of Nottingham, Nottingham NG7 2RD, UK.

<sup>3</sup>Swiss Institute for Experimental Cancer Research (ISREC), École Polytechnique Fédérale de Lausanne-(EPFL-SV), 1015 Lausanne, Switzerland. Electronic address: joerg.huelsken@epfl.ch.

**Author contributions:**

J.H. conceived the project. M.N. and J.H. designed the research and M.N. performed the experiments, established the bioinformatics pipelines, and performed the computational analysis, with contributions from J.H. P.O.-M. helped with organoid culture establishment and advised during the project. M.N. and J.H. prepared the manuscript, with input from P.O.-M.

## 2.2.1 Graphical abstract



### In brief

Norkin et al. develop high-throughput, high-content organoid screening for the systematic, large-scale, and quantitative analysis of complex biological systems involving many cell types and alternative activation states. Targeted RNA-seq analysis of signature gene sets allows the classification of responses to large collections of drugs or other perturbagens at minimal cost.

### 2.2.2 Summary

Organoids allow to recapitulate intestinal homeostasis and cancerogenesis *in vitro*, however, RNAseq-based methods for drug screens are missing. We developed TORNADO-seq—a high-throughput, high-content drug discovery platform that for the first time uses targeted RNAseq to monitor the expression of large gene signatures for the detailed evaluation of cellular phenotypes in organoids. TORNADO-seq is a fast, highly-reproducible, time- and cost-effective (5\$ per sample) method that can probe cell mixtures and their differentiation state in the intestinal system. We applied this method to isolate drugs that enrich for differentiated cell phenotypes and used them for colorectal cancer treatment. Indeed, these drugs are highly efficacious against cancer compared to wild type organoids. Further, TORNADO-seq facilitated in-depth insight on the mode of action of these drugs. Our technology can easily be adapted to many other biological systems and will allow for a more systematic, large-scale, and quantitative approach to study their biology.

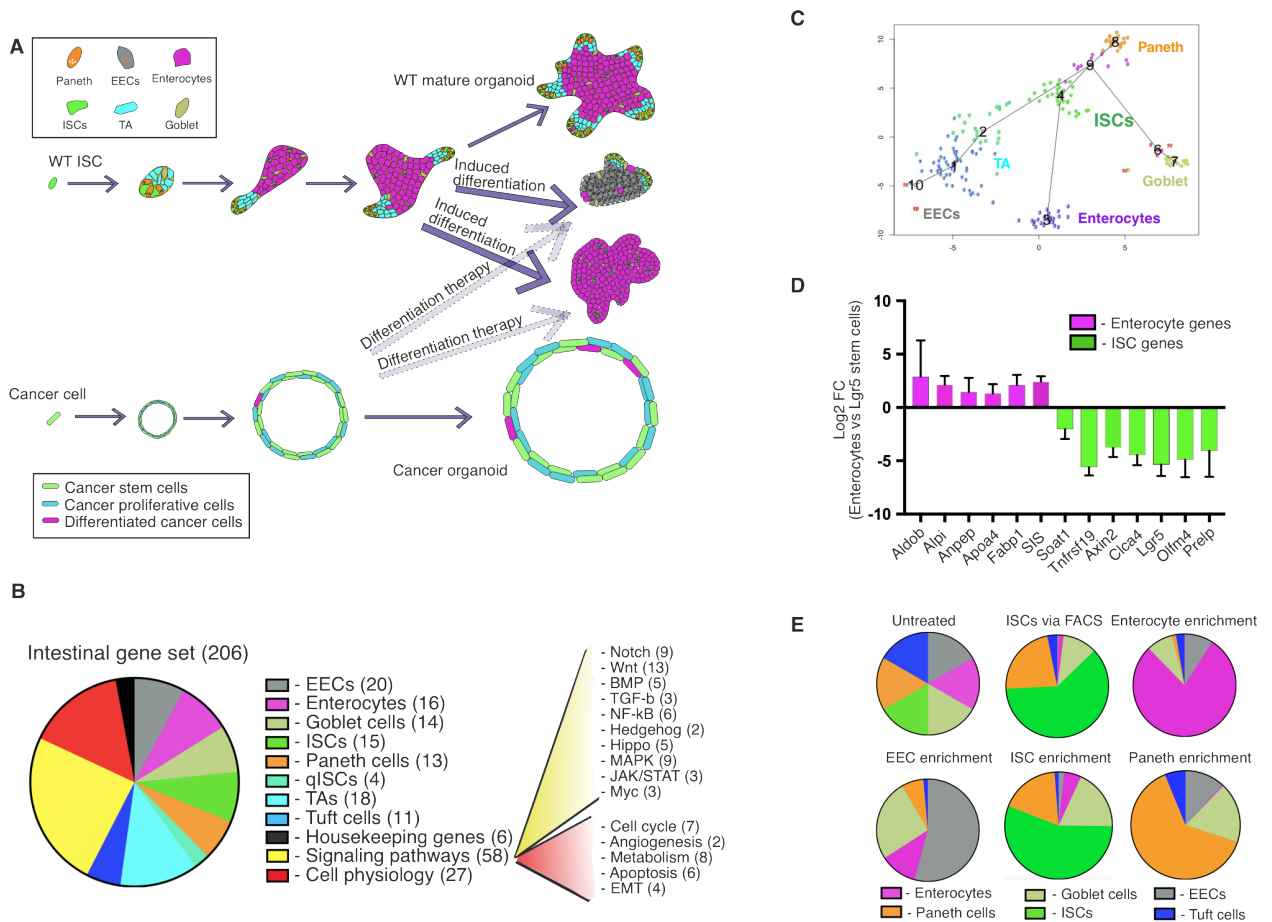
### 2.2.3 Introduction

Colorectal cancer (CRC) is among the three most commonly diagnosed cancers and one of the major causes of cancer-related morbidity and mortality (Hagggar and Boushey, 2009). In particular late-stage and metastatic tumours remain challenging to treat. Novel drug-based therapies are urgently needed, since despite many attempts, successful surgery still contributes most to patient outcome. CRC tumours were shown to retain partially differentiated, enterocyte- and goblet-like phenotypes, as well as stem-like epithelial phenotypes, indicating existence of rudimentary differentiation programs (Li et al., 2017). In particular, the frequency of stem-like cells has been linked to tumour maintenance, progression, and metastasis (Fumagalli et al., 2020). We have recently established differentiation therapy as a novel modality for CRC, which targets these stem-like cells and induces their differentiation, resulting in tumour regression in patient-derived xenotransplants and mouse models of intestinal cancer (**Figure 1A**) (Ordóñez-Morán et al., 2015). Intestinal epithelial differentiation comprises a complex program which involves at least 8 different cellular phenotypes and has been modelled recently using organoid systems (Yin et al., 2014). Intestinal organoids are derived from intestinal stem cells (ISCs) which expand and differentiate in an extracellular matrix to contain all cell types present in the gut epithelium. This highly representative model has been used to study development, mechanisms of differentiation, and rare epithelial phenotypes of the intestine (Yin et al., 2014; Basak et al., 2017). For example, it has been shown that inhibition of Wnt and activation of Notch signalling induces absorptive lineage differentiation, while inhibition of Wnt, Notch, and EGFR signalling induces enteroendocrine cell (ECC) fate (Basak et al., 2017). Further, organoids can be derived from CRC patient samples (patient derived organoids, PDOs), and have been used for personalized therapy approaches with a limited number of drugs (10-60 PDOs and 1-83 compounds; Gao et al., 2015; van de Wetering et al., 2015; Verissimo et al., 2016; Schütte et al., 2017; Tashiro et al., 2017). These assays typically measured cell growth/death, which provided only limited information on the biology of drug action. Drug screens with CRC cell lines often covered a larger number of drugs, however, were still limited by a small number of measured parameters such as the activity of one to three pathway reporters (Miyamoto et al., 2019; Zhan et al., 2019)

Here, we developed a novel form of high-content, high-throughput screening in organoids in order to identify small molecule drugs able to induce differentiation of intestinal wild type and cancer cells. This is achieved by RNA expression analysis



quantifying a large number of cell-type specific genes followed by deconvolution methods to infer cell type composition.



**Figure 1. Differentiation therapy and the establishment of a representative intestinal gene set.** **A.** Schematic representation of the differentiation process in wt organoids and differentiation therapy in cancer. **B.** Composition of the selected 206 gene set for the assay. **C.** Clustering of single intestinal epithelial cells based on the expression of our 206 gene signature into the different intestinal populations using the StemID2 algorithm (Grun et al., 2015). Cluster numbers are depicted in black; each dot corresponds to one cell. Cluster 1 – Enterocyte progenitors, cluster 2 – TA, cluster 3 Enterocytes, cluster 4 – ISCs, cluster 5 -EECs (not depicted), cluster 6 – goblet cells, cluster 7 – goblet cells, cluster 8 Paneth cells, cluster 9 – Paneth cells, cluster 10 – EECs. **D.** qPCR data on organoids enriched for stem cells or absorptive enterocytes. Log<sub>2</sub> FC with SD values are shown. Number of replicates n=3. **E.** qPCR data on wt organoid samples enriched for certain cell populations represented as pie charts. Each sector of a pie represents one cell population which was detected in the sample based on expression of 5-10 specific marker genes. Untreated organoids are shown as a pie with equal percentage of each cell population (artificial cell composition is depicted). Fold change of each population in the sample composition is converted to the percentages represented as sectors in the pies. Organoid culture conditions: Enterocytes - IWP-2 and valproic acid (VPA); ISCs - CHIR and VPA; Paneth - DAPT; EECs - DAPT, IWP-2 and Gefitinib, 2 days treatment.

As a high-content technology, mRNA expression analysis offers several advantages such as precision, scalability, and sensitivity over other methods such as

those that are antibody or reporter-based. Until now, no NGS-based screens in organoids were developed due to the technical limitations imposed by organoids growing in solidified drops of matrigel and the cost of classical RNA-sequencing assays. Further, eukaryotic transcriptomes exhibit a wide dynamic range of gene expression levels with a minority of highly expressed genes comprising the majority of RNA molecules within a cell. Classical RNA sequencing (RNA-seq) insufficiently samples weakly expressed transcripts, resulting in sparse sequence coverage and uncertain quantification. By using oligonucleotide primers to restrict analysis to selected genes, targeted RNA-seq improves sequencing read coverage, allowing sensitive and reliable gene expression measurement over an extended dynamic range. Furthermore, in combination with multiplex library preparation, the increased efficiency of targeted RNA-seq can also reduce costs by increasing the number of conditions analysed in parallel.

For the analysis of complex phenotypes such as multi-lineage differentiation, there is a need for high-content assays which not only enable drug discovery, but also provide insight into biological mechanisms. Here we developed a high-content assay to measure gene expression profiles that provide information on the frequency of all possible intestinal cellular phenotypes, major signalling pathways, and general cell physiology, detecting 206 carefully selected genes in either wild type (wt) or cancer intestinal organoids. Our assay is cost-efficient and allows the analysis of thousands of treated samples in one sequencing reaction, while achieving high efficiency and reproducibility. Our technology can be easily adapted to answer many other biological questions and will allow a more systematic, large-scale, and quantitative approach in a number of fields such as developmental biology, physiology, pharmacology, personalized medicine, and others, which involve dynamic, multicellular *in vitro* systems.

## 2.2.4 Results

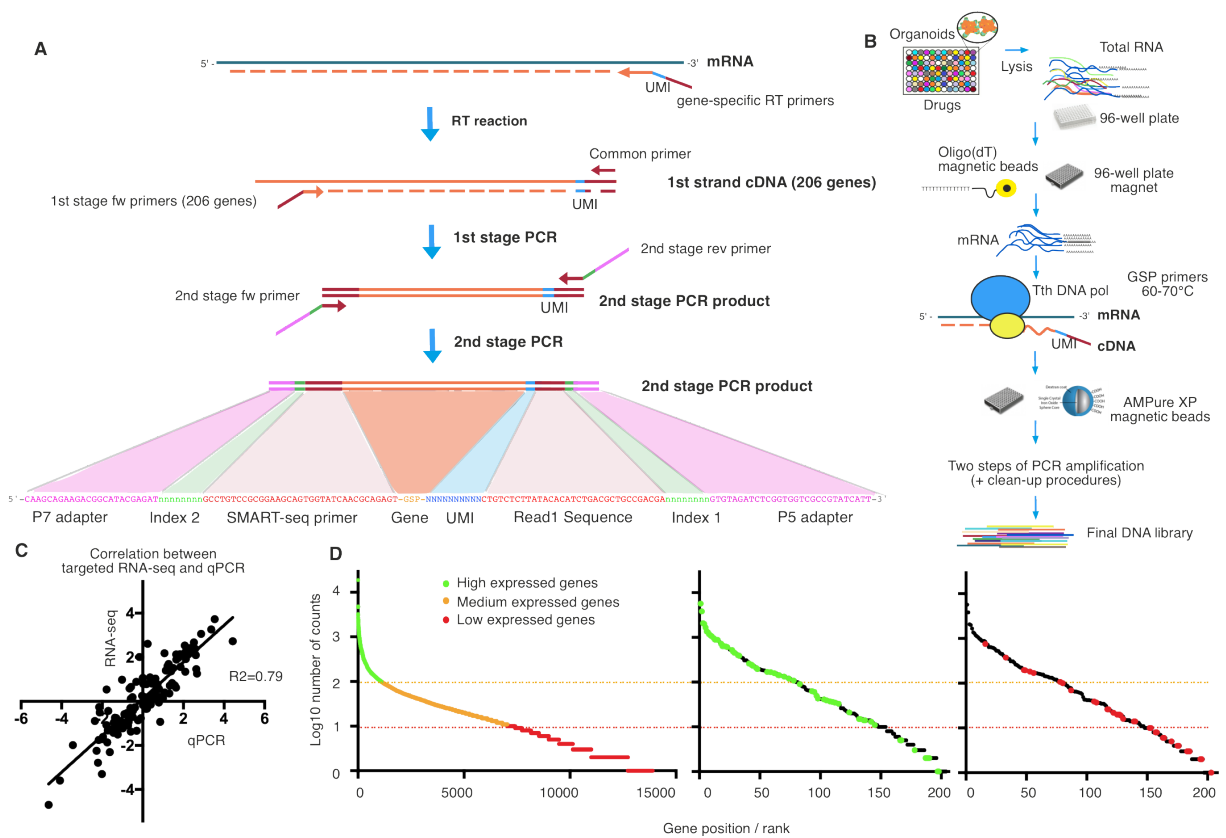
### 2.2.4.1 Screen design and validation

In order to cover the different intestinal cell phenotypes, we selected 111 marker genes for these eight populations: intestinal stem cells (ISC), transient amplifying cells (TA), absorptive enterocytes (E), enteroendocrine cells (EEC), goblet cells, Paneth cells (P), tuft cells, and quiescent ISCs (qISCs) (**Figure 1B**). These marker genes were selected from published data sets of bulk RNA-seq and single-cell RNA-seq of primary intestinal cell suspensions and organoid cultures which were further enriched by FACS or *in vitro* differentiation protocols (please see **Supp. Table 1** for the full list of studies;

Basak et al., 2017; Grun et al., 2015; Grun et al., 2016). We further included 58 genes reporting on major signalling pathways (Notch, Wnt, Hedgehog, Hippo, BMP/TGF- $\beta$ , NF- $\kappa$ B, and growth factor signalling), 27 genes of general cell physiology (cell cycle, metabolism, angiogenesis, apoptosis, epithelial-mesenchymal transition), and 6 housekeeping genes expressed in the whole intestine. The complete list of 206 genes was selected based on uniqueness, expression level, reproducibility, and multi-set intersection (**Supp. Table 1**). The clustering of single cells into the different intestinal populations based on the expression of only this gene list produced comparable results as the full transcriptome with 90% overlap in classifications of cell identity using the StemID (Grun et al., 2015; Grun et al., 2016) or the Seurat (Butler et al., 2018) algorithm confirming the performance of the selected gene set (**Figure 1C, Figure S1A**).

Primers were designed using standard primer design guidelines considering secondary structures and primer dimers using established software for multiplex PCR design (for details please see Methods). To validate primers for our gene list, we performed qPCRs on different cell populations of the murine small intestine. Cells were either sorted by FACS (Lgr5-GFP<sup>+</sup> ISCs) or induced to differentiate into E, ISC, P or EEC lineages in organoid culture using published media conditions (Yin et al., 2014; Basak et al., 2017). Our results showed robust upregulation of lineage-specific markers in the respective populations: for example, the stem cell signature was highly enriched in Lgr5<sup>+</sup> ISCs but lost in absorptive enterocytes; and vice versa, the enterocyte-specific signature was enriched in organoids induced to differentiate into absorptive enterocytes and barely detectable in ISCs (**Figure 1D**). Such validation was performed for all major lineages (**Figure 1E**).

In order to implement this gene set for high-throughput screening we developed and validated a targeted RNA-seq method (**Figure 2A, 2B**) that we termed TORNADO-seq (**Targeted ORgANOiD sequencing**). Briefly, we isolated mRNA of treated organoids from one well of a 96-well plate (about 10,000 cells) using oligo(dT) magnetic beads and synthesized cDNA in a reverse transcription (RT) reaction using a mixture of 206 gene-specific reverse primers carrying a unique molecular identifier (UMI) and a common part needed for subsequent amplification steps. The obtained cDNA was amplified in a 1<sup>st</sup> stage PCR with one common reverse primer and a mixture of 206 gene-specific forward primers (containing a second common part) (**Figure 2A**). Obtained products were barcoded in a 2<sup>nd</sup> stage PCR and pooled for library preparation and sequencing (**Figure 2A**).



**Figure 2. Schematic representation of library preparation for targeted RNA sequencing.**

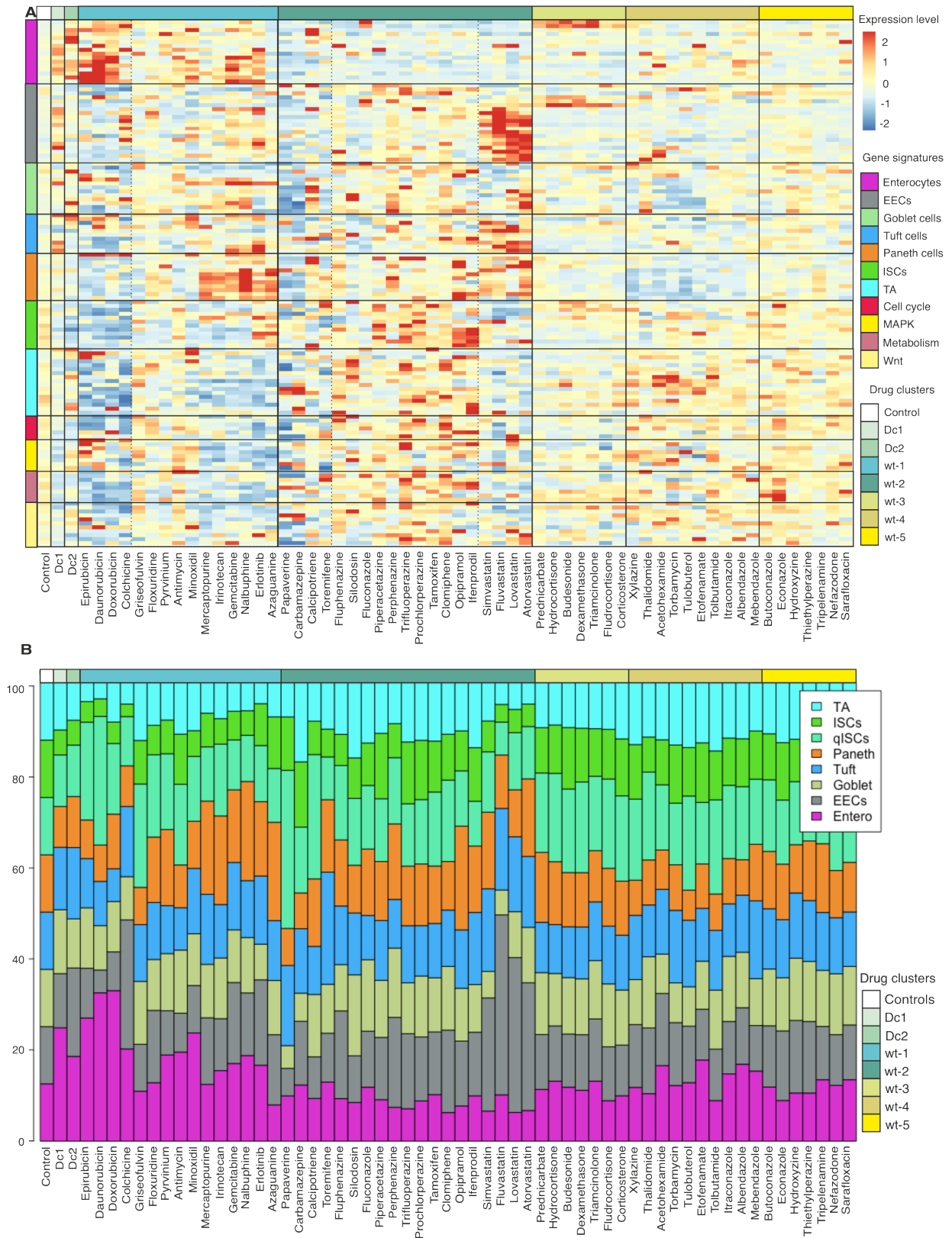
**A.** cDNA is synthesized from polyA mRNA in an RT reaction using a mixture of 206 gene specific reverse primers (GSP) carrying a UMI and a common sequence (part of Illumina Read1 sequence) needed for further amplification steps. The obtained cDNAs are further amplified in a 1<sup>st</sup> stage PCR with one common primer and the mixture of 206 gene specific forward primers (carrying a second common part derived from SMART-seq). The obtained products are barcoded in a 2<sup>nd</sup> stage PCR using a set of forward and reverse primers (carrying index1 and index2, and P5 and P7 adapters for Illumina sequencing). **B.** Schematic view of high-throughput processing of drug treated intestinal organoids. Organoids were seeded in a 96 well plate format in Matrigel, drug-treated, then lysed. mRNA was isolated from lysates using oligo-dT magnetic beads in a 96 well plate format. Tth DNA polymerase was used for the RT reaction, and cDNA is purified with AMPure XP magnetic beads. After two rounds of PCR amplification, an additional clean-up procedure with AMPure XP magnetic beads is performed, and final libraries are pooled and purified twice on 2% agarose gel (not shown). **C.** Pearson's correlation of mRNA quantification between qPCR and targeted RNA-seq. 30 random genes were chosen for the comparison. Log<sub>2</sub> FC values are represented. **D.** Left: Conventional RNA-seq data of untreated organoids from repository GSM2358985. 3,000,000 reads distribution, 14,000 genes represented. Genes from our 206 gene signature with high, mid, and low expression are represented by green, orange, and red dots, respectively. Right: Targeted RNA-seq of untreated organoids. 40,000 reads distribution. High (green) and low (red) expressed genes (colours as quantified in 2D) are depicted with their expression rank quantified by our targeted RNA-seq method. The number of reads was normalized by the combined number of all 206 selected genes. The number of reads between conventional and targeted RNA-seq was normalized by the combined expression of the 206 selected genes.

After analysis of the initial targeted RNA-seq results, primers showing inadequate or unspecific amplification were re-designed and substituted by new pairs. In addition, primer concentrations were adapted so that amplicons of high abundance were diluted enabling enrichment of sequencing read coverage for weakly expressed genes and more precise quantification. In total, three iterations of primer re-design and concentration adjustment were performed to optimize the assay. The final setup provides excellent reproducibility showing a Pearson's correlation coefficient of  $r > 0.95$  based on raw read counts of untreated biological replicates (**Figure S1B**). Comparison of RNA-seq and qPCR-based quantification of gene expression revealed a high correlation confirming the accuracy of our method with a Pearson's correlation coefficient of  $r = 0.79$  (**Figure 2C**). We sequenced more than 3,000 samples with an average sequencing depth of 70,000 reads per sample. The average read coverage was 370 reads per gene with 50% of genes having counts above 100 reads per gene and more than 80% of genes having counts above 10 reads per gene (**Figure 2D**). Average mapping efficiency to the targeted gene signature was 70% and reached 95% as a function of mRNA quantity. Duplication rate as assessed by UMIs varied between genes and equalled 2.2 transcripts per UMI on average. Due to the improved uniformity of read distributions by TORNADO-seq compared to conventional RNA-seq technology, we can employ moderate sequencing depth (50,000 reads per sample) allowing the analysis of more samples in parallel and reducing the total cost of our assay (including reagents for cell culture, sequencing library preparation, and sequencing cost) to 5 USD per sample.

#### **2.2.4.2 Identification of differentiation-inducing drugs in wt intestinal organoids**

In order to identify small molecule drugs that induce intestinal epithelial differentiation in organoids from wt mice, we assayed 320 compounds from a library of FDA-approved drugs. Our screen identified 56 drugs as potential hits using a threshold of at least five significantly altered, differentially expressed (DE) genes ( $|\log_2FC| > 1$ ,  $p_{adj} < 0.05$ ) based on obtained gene expression profiles (**Figure 3A**). The screen exhibited good reproducibility showing a Pearson's correlation coefficient of  $r = 0.93$  based on raw read counts across all drugs (**Figure S1C**) and  $r = 0.65$  for top drug hits using fold change (FC) values over untreated samples. Among those potential hits were substances already utilized for treatment of colon cancer (itraconazole; Buczacki et al., 2018), pyrvinium (Wiegering et al., 2014; Li et al., 2014), tubulin inhibitors (bendazoles, colchicine), and cytotoxic drugs known to affect general cell physiology (anti-

metabolites: gemcitabine, azaguanine, mercaptopurine, floxuridine; topoisomerase inhibitors; and cytotoxic antibiotics: anthracyclines and antimycins). Interestingly, our screen identified several novel drug candidates such as antipsychotic phenothiazines, cholesterol-lowering statins, anti-mycotic conazoles, selective estrogen-receptor modulators (SERM), glucocorticoids, and antihistamines. Supervised clustering based on cell type-specific genes identified common gene expression alterations that resulted in five distinct clusters of drugs (**Figure 3A**). All cytotoxic compounds (cluster wt-1) suppressed ISC, TA, and cell cycle signatures as expected. At the same time, the more potent anthracyclines (epirubicin, doxorubicin, daunorubicin) exhibited a very specific pattern and led to a relative increase of absorptive enterocyte-specific genes. Loss of ISCs and increased enterocytic differentiation was confirmed by FABP1 immunofluorescence staining and cytometry for Lgr5-GFP (**Figure 4A, 4C**). This is likely an indirect effect since these compounds deplete proliferative cells (**Figure 4C**) causing a relative accumulation of differentiated cells, most prominently absorptive enterocytes as the most frequent differentiated cell type in organoids. Indeed, when analysing the microscopy images that were taken before harvesting the organoids for RNA isolation, cell death was twice as frequent for cytotoxic compared to non-toxic compounds (**Figure S2C, S2D**). We excluded the possibility that dead cells affected the measured RNA expression changes (**Figure S2A, S2B**). Importantly and in contrast to the cytotoxic drugs of cluster wt-1, the remaining clusters (wt-2 – wt-5) induced different responses. The vast majority of drugs in cluster wt-2 (statins, SERMs, and phenothiazines) decreased enterocyte frequencies, upregulated markers of other differentiated cell phenotypes such as EECs, goblet, or tuft cells, and decreased proliferation-related genes to a lesser extent. In particular, statins and phenothiazines were characterised by a high increase in EEC markers. Cluster wt-3 contained glucocorticoids which induced an inflammatory signature evidenced by upregulation of *Apoa4*, *Cdkn1a*, *Pdlim2*, *Prpa1*, and *NFKBIA*, and downregulation of *Clca4*, *Sis*, *Cdc25c*, and *Cck*. This more global pattern in response to glucocorticoids was also observed in the CMAP database, which supports the validity of our assay (**Figure S3A**). Drug cluster wt-4 showed robust downregulation of Paneth cells, while drugs from cluster wt-5, which included conazoles, altered only few differentiation-associated genes. Phenotypic changes reported in the screen were validated by qPCR for some of these clusters, confirming the robustness of our method (**Figure S3B**).



**Figure 3. TORNADO-seq identifies differentiation-inducing drugs in wt intestinal organoids. A.** Supervised clustering of wt treated samples based on the expression of cell type specific genes. Log<sub>10</sub>[counts + 1] values are represented. Expression level is computed as row Z-score values. Drug cocktails (dc) dc1 and dc2 enrich for absorptive enterocytes and EECs (details in methods part). **B.** Calculated cell type composition represented as stacked bar charts. Each sector of a bar represents one cell population which was detected in the sample based on

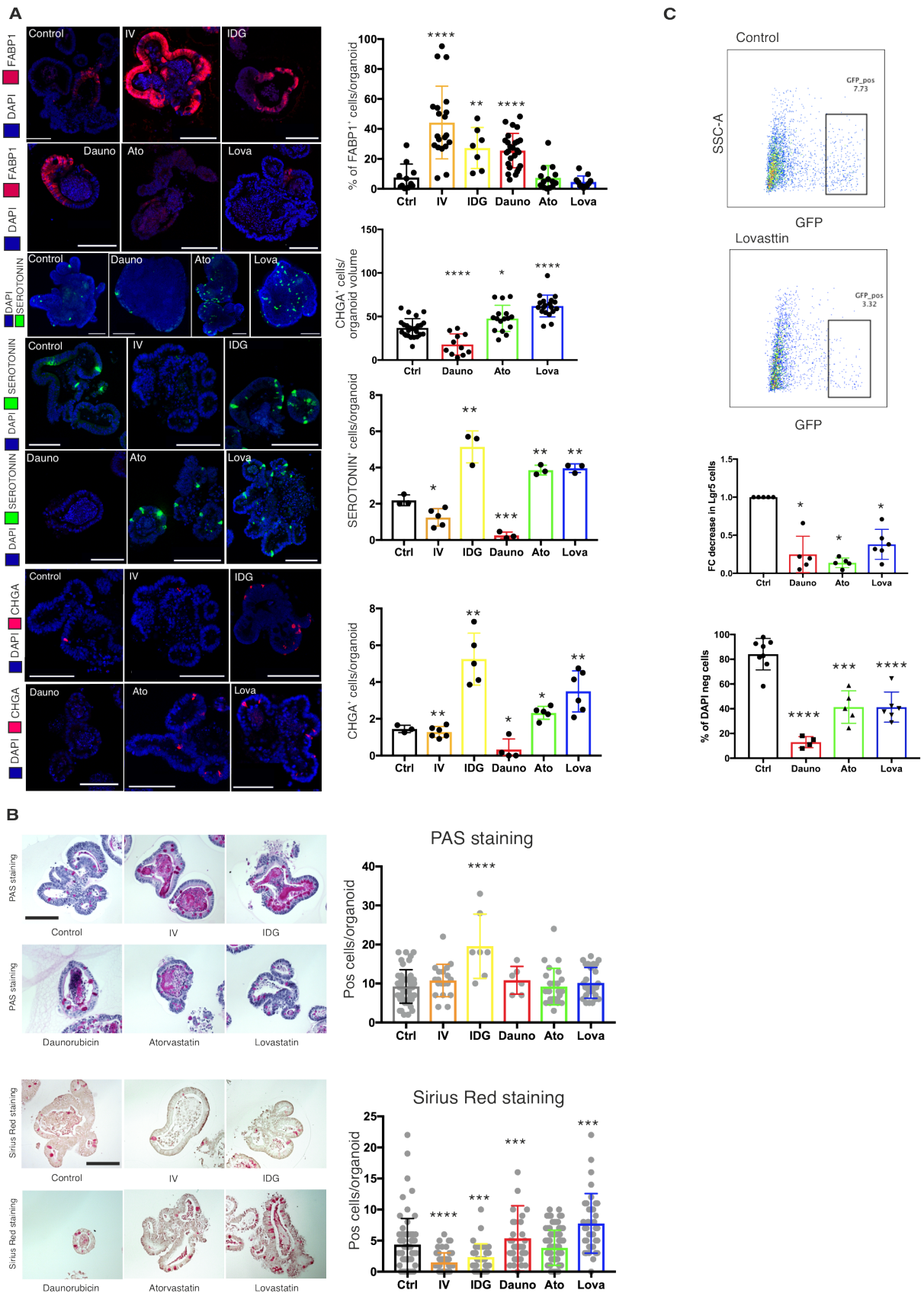
expression of specific marker genes. The untreated control depicts a cell type distribution in wt organoids where each population is of equal size to facilitate comparisons. Calculated relative fold change of each population upon drug treatment was converted to bar sectors. The drug order and gene expression information are identical to that in Figure 3A. Drug cluster colour bar represented at the top of the figure.

The majority of potential drug hits was further evaluated based on the upregulation or downregulation of at least five differentiation or stem cell genes, respectively ( $|\log_2FC| > 1$ ,  $p_{adj} < 0.05$ ), which resulted in 27 drugs classified as differentiation-inducing. Further, cell type enrichment analysis revealed that some of these drugs induced differentiation to a specific cell type. We identified lovastatin, atorvastatin, perphenazine and trifluoperazine to specifically induce EEC enrichment as well as Paneth cell markers. All four drugs do not upregulate absorptive enterocyte specific genes, which is in contrast to another EEC induction cocktail (**Figure 3A:DC2**) suggested earlier (Basak et al., 2017) which has broader effects on intestinal differentiation including induction of enterocytes and goblet cells (**Figure 4A, 4B**). Moreover, we also identified drugs which specifically deplete certain cell types: for example, silodosin depleted enterocytes, while tulobuterol, xylazine, and other drugs from cluster wt-4 depleted Paneth cell markers (**Figure 3A**). Cell type enrichments for all drugs are quantified and summarized in **Figure 3B** and **Supp. Table 2**. For selected drugs, treatment effects were validated by orthogonal assays (**Figure 4**). We used immunofluorescence analysis of Fabp1 to quantify enterocytes and of ChromograninA and Serotonin to detect enteroendocrine cells. Further, we employed flow cytometry to measure intestinal stem cells via the Lgr5-GFP allele and cell death via DAPI. Finally, we used the histology stains Sirius Red and Periodic acid-Schiff (PAS) to identify Paneth and goblet cells, respectively. These assays confirmed changes in cell lineage frequency in line with what we had measured using our targeted RNA-seq approach. For example, statins increased secretory lineages such as enteroendocrine and Paneth cells and decreased stem cell frequency while Daunorubicin was increasing enterocytes and decreasing stem and enteroendocrine cell frequencies.

#### **2.2.4.3 Comprehensive analysis of drug-induced differentiation patterns reveals connections between cellular phenotypes and signalling pathways in wt organoids**

One of the advantages of TORNADO-seq is the possibility to discover potential crosstalk between the various signalling pathways and cell types by identifying frequent co-regulation of genes over many treatment conditions.





**Figure 4. Differentiation phenotypes are confirmed by orthogonal assays. A.** Immunofluorescence stainings with the indicated antibodies against FABP1, ChromograninA (ChrA) and Serotonin of organoids cultured under the following conditions: IDG (IWP-2, DAPT, Gefeteneb) and IV (IWP-2, VPA) are known drug cocktails (Yin et al., 2014; Basak et al., 2017), Dauno – daunorubicin, Ato – Atorvastatin, and Lova – Lovastatin are drugs identified in our

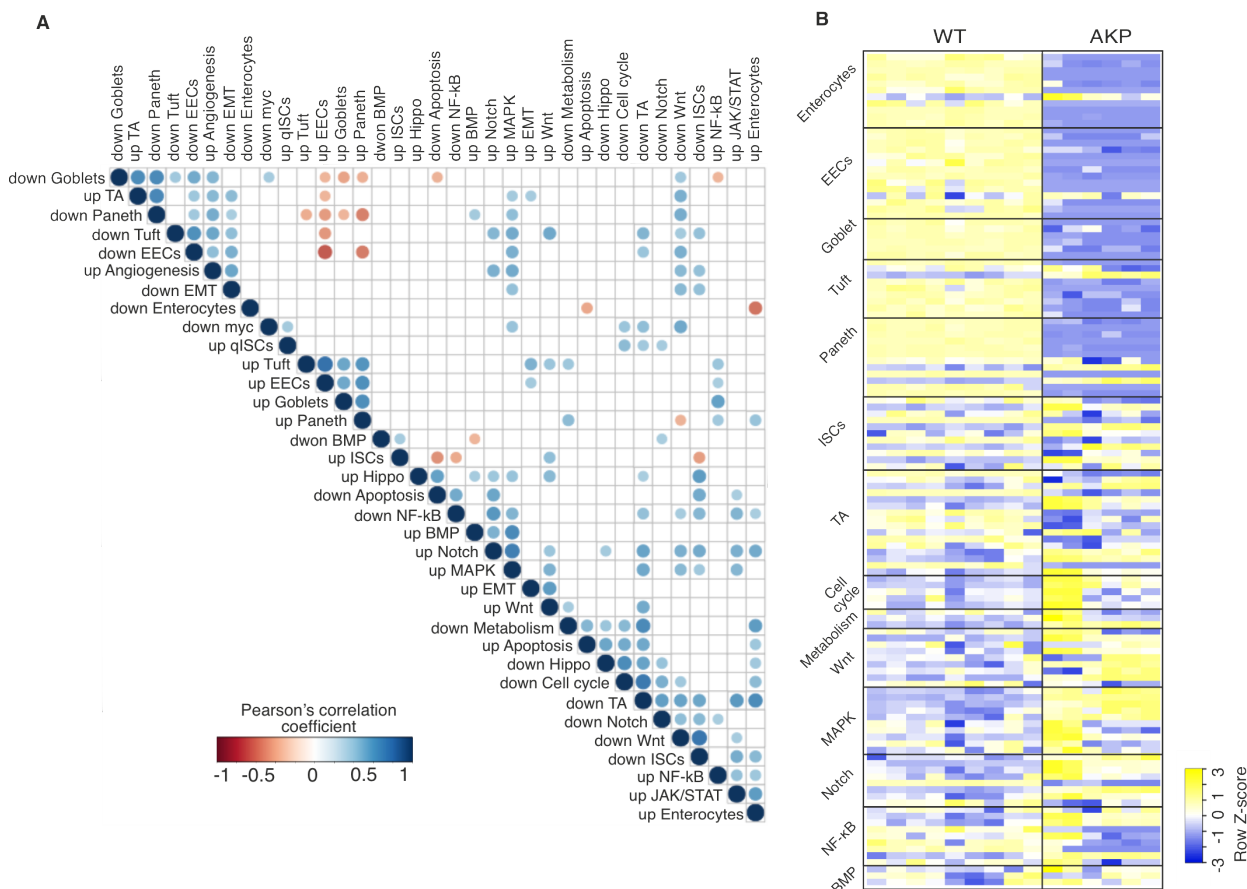
screen. Organoids were treated at 10  $\mu$ M for 2 days in ENR medium. Scale bars, 100  $\mu$ m. **B.** Periodic acid-Schiff (PAS) and Sirius Red stainings for the detection of goblet and Paneth cells of organoids cultured under multiple conditions. IDG (IWP-2, DAPT, Gefeteneb) and IV (IWP-2, VPA) are known drug cocktails (Yin et al., 2014; Basak et al., 2017); Dauno – daunorubicin, Ato – Atorvastatin, and Lova – Lovastatin are drugs identified in our screen. Organoids were treated at 10 $\mu$ M for 2 days in ENR medium. Scale bars, 100  $\mu$ m. **C.** FLOW cytometry of organoids treated as above using the Lgr5-GFP allele to measure ISCs and DAPI to measure alive cells.

We analysed these possible connections using correlograms (**Figure 5A**) to directly show such gene-gene correlations (**Figure S4**). We observed that absorptive enterocyte differentiation strongly correlates with downregulation of ISCs and TAs as well as proliferative signatures. Differentiation into EEC, tuft, goblet, and Paneth cells is highly interconnected, in line with their placement in the secretory lineage with generation from a common, secretory lineage progenitor cell. The same applies to signalling pathways: Wnt, Myc, and cell cycle genes show high correlation, consistent with c-Myc as a known downstream target of the Wnt pathway and the role of this pathway in driving intestinal proliferation and cancer development. As well-known organizers of the stem cell niche in the intestine, Notch and Wnt pathways also displayed a high correlation. Further, Wnt/ $\beta$ -catenin signalling correlated with epithelial-mesenchymal transition (EMT) as expected (Basu et al., 2018; Kim et al., 2019). Interestingly, a decrease in Paneth cell markers was correlated with increased proliferation and Notch signalling (**Figure 5A**). This may be linked to the ability of Paneth cells to de-differentiate and acquire stem-like properties upon injury (Schmitt et al., 2018). Moreover, Notch inhibition is well known to promote Paneth cell expansion (Yin et al., 2014). *In vivo*, the intestinal epithelium is characterized by low oxygen concentrations and a decreasing oxygen gradient towards the lumen (Karhausen et al., 2004). Our analysis identified VEGF, a known marker of hypoxia, to anti-correlate with several differentiated lineages including enterocytes. It remains to be seen to what extent oxygen concentrations may contribute to the intestinal lineage differentiation profile. In essence, TORNADO-seq identifies both known and novel interactions between cell states and signalling pathways, suggesting that this method can be an advantageous discovery tool for hypothesis generation in a variety of biological systems.

#### 2.2.4.4 Evaluating differentiation-inducing drugs in cancer organoids

We compared the profiles of untreated, wt, and transformed APC<sup>lof</sup>:KRAS<sup>G12</sup>:TP53<sup>lof</sup> (AKP) organoids which were obtained from advanced mouse intestinal tumours (**Figure 5B**). As expected, in comparison to wt organoids, we observed an upregulation of proliferative signatures and a downregulation of

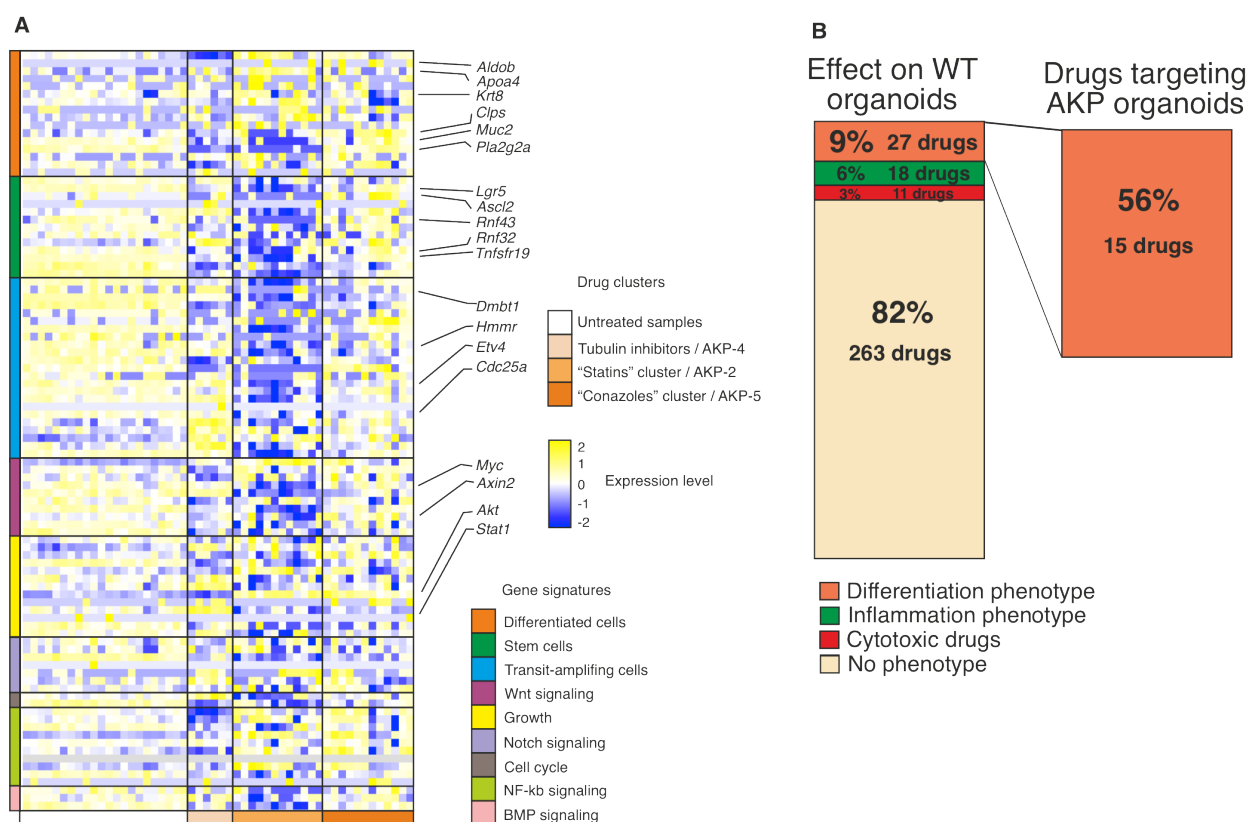
differentiation markers in AKP organoids, which is a hallmark of advanced tumour progression. Of note, wt and cancer organoids exhibit different morphology: while wt organoids grow as budding structures, the cancer organoids form spheres (**Figure 1A**).



**Figure 5. TORNADO-seq identifies connections between cellular phenotypes and signalling pathways in organoids. A.** Correlogram showing the correlations between various pathways and cell types in wt organoids. Correlations are computed from the gene expression profiles (Figure 1B) of organoids treated with drugs from Figure. 3A (30 drugs with highest number of DE genes, see Method section) with `cor` function in R with assay p-values equal to 0.05. **B.** Expression profiles of untreated wt and AKP organoids for 130 highly expressed genes in both systems from our 206 gene set grouped by cell type or pathway/function.  $\log_{10}[\text{counts} + 1]$  values are represented. Expression level is computed as row Z-score values.

We next performed TORNADO-seq on drug treated AKP cancer organoids selecting only those drugs which had scored as potential hits in wt organoids. We found over half (16 out of 27) of the drugs inducing differentiation in the wt also affect AKP organoids (using the same criteria as for the wt: deregulation of at least five stem or differentiation-related genes) by inducing loss of stem cell signatures concomitant with an upregulation of differentiation markers (**Figure 6B, Supp. Table 3**). Importantly, when we evaluated the full set of 320 drugs on AKP cancer organoids by morphological scoring, we obtained identical results: only drugs inducing differentiation in wt organoids

targeted cancer organoids causing prominent growth arrest and cell death while all other drugs were ineffective. This suggests that differentiation induction may be a major factor determining drug activity.



**Figure 6. Differentiation-inducing drugs target AKP cancer organoids. A.** Clustering of drug-treated AKP cancer organoids based on their gene expression profiles (97 genes differentially expressed between the samples are depicted). For each drug, both replicates are displayed on the heatmap. Log10[counts + 1] values are represented. Expression level is computed as row Z-score values. **B.** Summary of drug effects on wt and AKP organoids. Left panel represents the phenotypes of all drug-treated wt organoids. Right panel represents the effect of the drugs that induce differentiation in the wt system on AKP cancer organoids. Drugs targeting AKP organoids are defined as those that downregulate/upregulate 5 or more stem cell/differentiation genes ( $(|\log_2FC| > 1, \text{padj} < 0.05)$ ).

Supervised clustering (excluding the drugs of cluster "wt-1" which largely contained cytotoxic drugs), revealed three drug clusters which resembled the clusters for wt organoids: one cluster containing tubulin inhibitors – "AKP-4"(cluster 4 in the wt), a cluster containing statins — "AKP-2" (cluster 2 in the wt) and a cluster containing conazoles — "AKP-5" (cluster 5 in the wt; **Figure 6A**). The "AKP-2" cluster revealed the strongest phenotype displaying a profound decline in proliferation, reduction of ISC and Wnt signatures, and a decrease of lipid biosynthesis genes, whereas stress-response genes (*Nupr1*, *NF-kB2*) and *MAPK* signalling increased. Importantly, statins also increased expression of differentiation genes such as tuft cell lineage markers *Krt8* and

Krt18. The “AKP-5” cluster induced higher expression of differentiation markers especially for goblet and Paneth cell specific genes (Muc2, Pla2g2a, Clps) and downregulated a variety of proliferation and growth-related genes. Tubulin inhibitors showed a distinct signature with a strong decrease of inflammation response genes and a moderate decrease in ISC and Wnt signalling genes. Drugs of the glucocorticoid cluster identified in the wt screens (wt-3) did not affect cancer organoids. For selected drugs, we analysed time dependence of the response in cancer organoids. This revealed early and late responses (**Figure S5**): a transient increase of stress responses only at day 1 of treatment involving Nupr1, Nfkb, Birc3, and NFKBIA, while Wnt signalling decreased and differentiation markers increased gradually from day 0 to day 2. Overall, this analysis revealed a striking overlap in the response of wt and cancer organoids suggesting similar modes of action of the drugs in either model.

#### **2.2.4.5 Treatment responses of cancer organoids exhibit distinct signatures**

As we have already established, untreated wt and AKP organoids display different expression profiles (**Figure 5B**). Correlation and differential expression analysis for treatment-responsive AKP samples (**Figure S6**) showed a strong reduction of many ISC genes while goblet-specific genes (Muc2, Agr2, Clps) were upregulated. Evaluation of gene-gene correlations using gene ontology revealed four differentially affected group of genes: proliferation/metabolic genes, G1/S phase genes, cell death/apoptosis genes, and inflammation/stress response genes (GO term p-values 2E-7, 2E-10, 2E-6, 2E-5, respectively; **Figure S6**). Interestingly, many genes from the first three groups were co-regulated in wt organoids (**Figure S4**: green circle), but in AKP organoids were differentially affected and each cluster of drugs – statins, conazoles, tubulin inhibitors, and cytotoxic compounds altered a specific combination of these group of genes. While proliferation/metabolic genes were downregulated in all drug clusters, the G1/S phase genes were only reduced by statins. Further, while statins, conazoles, and cytotoxic drugs increased inflammation genes, these were downregulated by tubulin inhibitors. Apoptosis genes were only activated by cytotoxic drugs (**Figure S7A**). In line with previous reports, the proliferation signature strongly anti-correlated with inflammation (Schmitt et al., 2018). Thus, using gene expression profiles of treated organoids, we were able to identify novel biological responses not only in wt but also in cancer organoids.

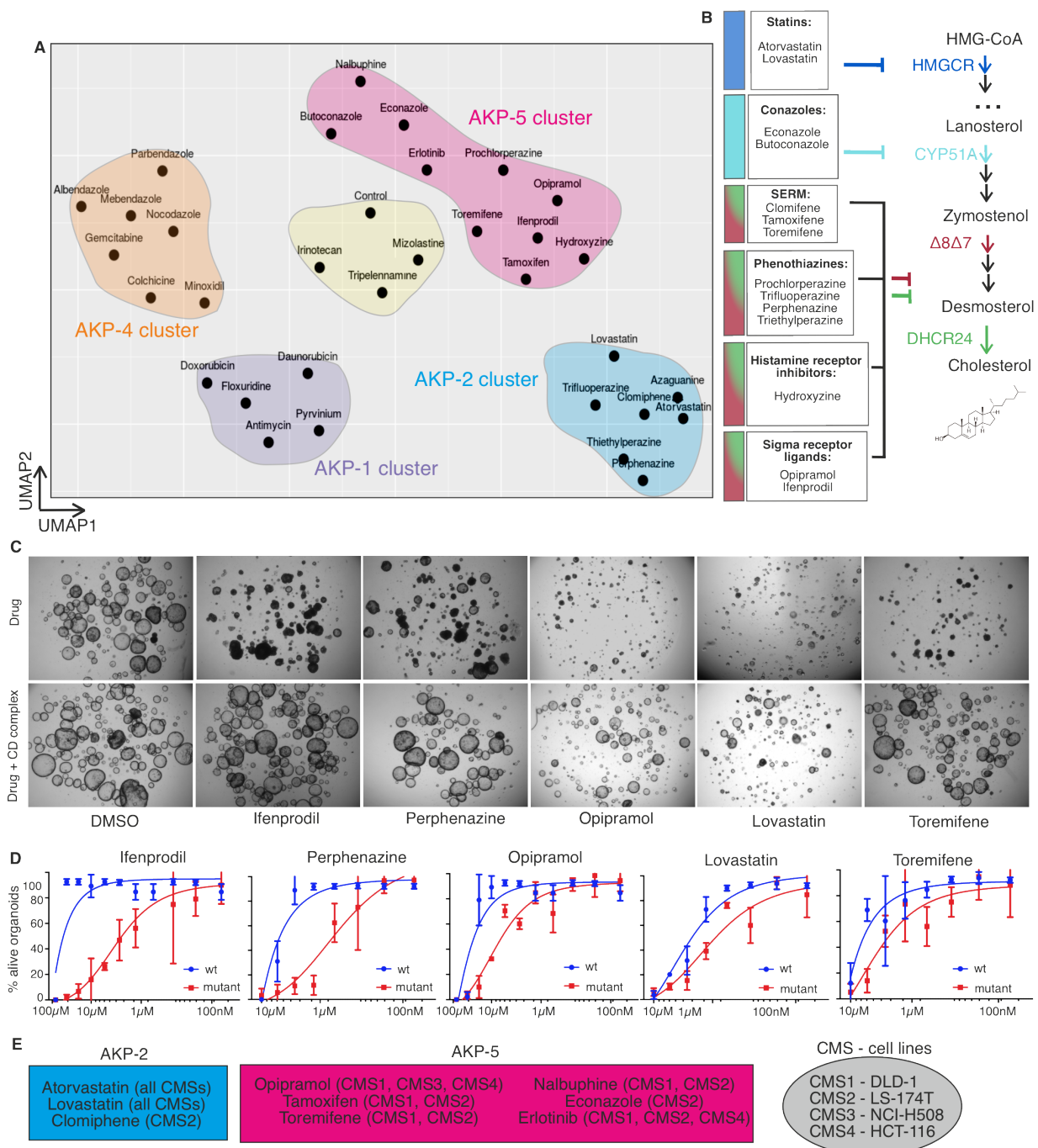
#### 2.2.4.6 Drug classification helps uncover the mechanism of action against cancer organoids

We performed clustering of treated wt and AKP samples based on gene expression profiles using non-negative matrix factorization (NMF) and Uniform Manifold Approximation and Projection for dimension reduction (UMAP) (**Figure S7B**). Despite the biological differences between wt and AKP organoids, the majority of drugs targeting AKP organoids clustered similarly to the wt setting, which indicates a common mode of action (MOA) in both systems. Interestingly, some of the co-clustering drugs were not known to have a common MOA. For example, phenothiazines, SERM inhibitors, and antihistamines produced similar responses as statins and conazoles (“AKP-2”, “AKP-5”), which are well known cholesterol lowering agents (**Figure 7A**). We therefore explored phenothiazines, SERM inhibitors, and antihistamines in more detail to elucidate if their MOA may be related to the cholesterol biosynthesis pathway. A literature search indicated that these drugs, among many other functions, can also inhibit enzymes in the cholesterol biosynthesis pathway or may alter cholesterol trafficking (**Figure 7B**) (Adams et al., 2003; Korade et al., 2016; Wages et al., 2018; Shim et al., 2015; Moebius et al., 1997). To verify this idea, we performed rescue experiments by adding cholesterol-cyclodextrin complex to drug-treated organoids. This indeed overcame differentiation phenotypes produced by statins, phenothiazines, SERMs and the sigma receptor ligands Ifenprodil and Opipramol (**Figure 7C**). Of note, this cholesterol-induced rescue was weakest for statins and increased drug concentrations overcame cholesterol-mediated rescues. A role of cholesterol for intestinal homeostasis and cancer is supported by recent findings (Wang et al., 2018; Voorneveld et al., 2017; Miyamoto et al., 2019). Importantly, several of these drugs (ifenprodil, opipramol, perphenazine, toremifene) exhibited higher potency against AKP compared to wt organoids suggesting a favourable therapeutic window for future application (**Figure 7D**). This example shows that our high-content analysis of phenotypes over a large number of drugs can help to identify MOAs and relevant pathways which may not correspond to the expected MOA for certain drugs.

We finally evaluated drugs of the AKP-2 (“Statins”) and AKP-5 (“Conazoles”) clusters on human cell lines. We chose four cell lines representing different consensus molecular subtypes (CMS) of colorectal cancer (Guinney et al., 2015): DLD-1 – CMS1, LS-174T- CMS2, NCI-H508 – CMS3 and HCT-116 – CMS4. More than half (10 out of 17) of the drugs identified to target AKP organoids also targeted human CRC cell lines



(Figure 7E, Supp. Table 4). While phenothiazines were not effective against human cell lines growing in culture medium with 10% FCS (which contains cholesterol), the statins Atorvastatin and Lovastatin, more potent inhibitors of cholesterol biosynthesis, showed the highest potency. Also for the human cell lines, statin effects could be rescued by further addition of cholesterol (data not shown) in support of their MOA. Thus, our screen identified several interesting candidates, statins, opipramol and toremifene for further *in vivo* validation as novel drug candidates for CRC therapy.



**Figure 7. The mode of action of cancer organoid targeting drugs is linked to the cholesterol biosynthesis pathway. A.** UMAP clustering of drug treated AKP organoids. Colour codes signify different clusters. UMAP is based on the expression of the most DE genes among

AKP treated samples. Cluster wt-3 is not present as glucocorticoids didn't affect cancer organoids. **B.** Schematic representation of cholesterol biosynthesis pathway and drugs from "Statins" and "Conazoles" clusters targeting different enzymes of this pathway according to a literature search (reference list in the main text). **C.** Phase-contrast images of rescue of drug effects with cholesterol. CD is cholesterol:methyl- $\beta$ -cyclodextrin (M $\beta$ CD) complex (Christian et al., 1997). **D.** Drug response titration curves comparing wt (blue) and AKP mutant (red) treated organoids after 4 days of treatment **E.** Colorectal cancer cell lines are targeted by drugs from AKP-2 (blue box) and AKP-5 (violet box) clusters. Right ellipse (grey) shows the assignment of these cell lines to a particular CMS.

## 2.2.5. Discussion

Here we describe TORNADO-seq as a novel high-content approach for high-throughput drug discovery in organoids. While the majority of drug screens are based on few parameters such as the activity of signalling pathways or the measurement of cell viability, they lack rich data required for the analysis of complex phenotypes and the MOA of drugs. We demonstrate here that TORNADO-seq is an excellent tool to overcome these limitations, for the discovery of new biological read-outs, and for the analysis of multicellular phenotypes. To our knowledge, this is the first NGS-based HTS platform developed for organoid cultures.

TORNADO-seq compares favourably with other high-content methods with respect to availability, cost, and applicability to organoid-based screens. TORNADO-seq shows high-reproducibility, requires only 6h to perform, is cost-efficient (5 USD per sample including culture, library preparation, and sequencing cost), and doesn't require any specialized equipment. Due to the implemented UMI-counting and sequencing of amplicons, TORNADO-seq lacks PCR- and ligation-based bias and misdetections. Primer design is straightforward, needs minor optimization, and only requires moderate effort. Several high-throughput NGS methods have been established previously which are inapplicable to organoid-based screens. Genome-wide RNA-seq methods require high sequencing depth for many of the weakly expressed genes relevant for our system which increases assay cost (Ye et al., 2018; Bush et al., 2017). Other targeted approaches are based on ligation reactions between acceptor and donor probes (Simon et al., 2019; Teder et al., 2018) and require tedious optimization and troubleshooting while suffering from PCR amplification bias and difficulties in detecting weakly expressed targets (**Supp. Table 5** for comparison).

Analysis of treated wt organoids revealed known and new aspects of intestinal epithelium organization and signalling networks. Cytotoxic compounds predominantly enrich for absorptive enterocytes, which is likely related to the fast induction of cell death



and therefore loss of proliferating stem and progenitor cells leaving behind enterocytes as the main component of intestinal organoids. Indeed, a higher frequency of enterocytes was observed in daunorubicin treated organoids by immunofluorescence analysis. In contrast, non-toxic drugs often eliminate enterocytes and enrich other cell types (e.g. EECs or goblet cells) indicating specific differentiation induction. We have also noticed a delayed phenotype with the non-toxic drugs compared to the swift action of cytotoxic drugs, which is in line with the observation that differentiation may act over a longer timescale. We were able to identify drugs specifically enriching organoids for a certain cell type. In particular, lovastatin and trifluoperazine specifically induced EEC enrichment more precisely than previously proposed drug cocktails (Basak et al., 2017) (**Supp. Table 2**). Observed phenotypes were detected by both, bulk gene expression changes and single cell analysis using immunofluorescence, FACS and histology. Indeed, higher expression of EEC marker genes was caused by an increased frequency of EECs.

While we can not fully exclude the possibility of passive enrichment of particular cell types by the corresponding drug treatments, it seems that non-toxic drugs indeed induce specific differentiation. While cytotoxic compounds most probably induce a simple shift in the cellular composition of organoids, statins seem to induce the enhanced differentiation into EECs. Indeed, cytotoxic compounds predominantly enrich organoids for absorptive enterocytes, which is likely related to the fast induction of cell death and therefore loss of proliferating stem and progenitor cells leaving behind enterocytes as the main cell type present in intestinal organoids. At the same time, statins specifically upregulate only EEC markers, which was also shown by additional orthogonal experiments (Immunofluorescence) where corresponding single cells were counted. Additional time course and *in vivo* lineage tracing experiments could be useful in further assessment of the mechanism of EECs enrichment.

Importantly, we identified many novel drug candidates targeting colon cancer organoids which were not described previously for CRC. Among them are antipsychotic phenothiazines, cholesterol-lowering statins, anti-mycotic conazoles, SERMs, glucocorticoids, and anti-histamines. Based on the obtained gene expression profiles we were able to propose and confirm MOAs for some of these drugs which we found to act by targeting the cholesterol pathway. Several of these drugs (ifenprodil, opipramol, perphenazine, toremifene) showed a beneficial therapeutic window targeting preferentially cancer compared to wt organoids. Finally, statins, opipramol and toremifene also targeted human CRC cell lines warranting future *in vivo* validation.

In line with our initial hypothesis, differentiation induction seems to be a common mechanism for drugs targeting cancer organoids, as all of these drugs also induced differentiation in wt organoids. Most of these drugs trigger stem cell loss in cancer organoids which may provide additional benefits for therapy since the stem-like population of cancer cells has been shown to have important functions in tumour maintenance, therapy resistance and metastasis.

TORNADO-seq is a tool which can be easily optimised and adapted for mouse or human *in vitro* systems such as organoids of various tissue types, which are increasingly being developed (Schutgens and Clevers, 2020), or other differentiation-capable systems such as embryonic stem cells or induced pluripotent stem cells. It can be further combined with other perturbators such as CRISPR or expression library technologies. TORNADO-seq is therefore a promising tool in high-throughput drug discovery and translational, personalized medicine approaches, as well as for basic questions in biology that address mechanisms of development or homeostasis of multicellular systems.

## **2.2.6 Acknowledgments**

This work was supported by grants from the Carigest and the ISREC foundation as well as by SNSF grants. We thank T. Petrova (UNIL, Lausanne) and F. Radtke (EPFL Lausanne) for mouse strains and B. Mangeat (EPFL, Lausanne) for advice on library structure development, P. Dessen (EPFL, Lausanne) for growth factor production, A. Madurga Alonso (EPFL, Lausanne) for help in setting up FACS stains, L. Spisak (EPFL, Lausanne) for help in setting up IF experiments, and C.M. Young (EPFL, Lausanne) for comments on the manuscript. We thank the EPFL core facilities including GECG, BSF, HCF, BIOP, and FCCF for their help during various parts of the project.

## **Author contributions**

J.H. conceived the project. M.N. and J.H. designed research and M.N. performed experiments, established the bioinformatics pipelines and performed computational analysis with contributions from J.H.. P.O.M. helped with organoid culture establishment and advised during the project. M.N. and J.H. prepared the manuscript with input from P.O.M..

## **Declaration of Interests**

The authors declare no competing financial interests.

## 2.2.7 STAR Methods

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Joerg Huelsken ([joerg.huelsken@epfl.ch](mailto:joerg.huelsken@epfl.ch)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

Gene expression profiles generated in this paper with TORNADO-seq has been deposited in the Gene Expression Omnibus (GEO) under accession number GEO: GSE157167. The files and the scripts for processing the raw data are deposited at <https://github.com/MaximNorkin91/Tornado-seq-protocol/>.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

All experiments with mice were authorized by the Canton of Vaud (license VD3396) and were performed according to accepted guidelines for animal handling.

### Mice

All mice were kept on a 12-hour light/dark cycle in individually ventilated cages. We used  $Lgr5^{GFP-IRES-CreER/+}$  or wt mice for generation of normal intestinal organoids. The cancer organoids were generated from  $APC^{lof}:KRAS^{G12}:TP53^{lof}$  mice. The  $Apc^{fl/fl}$  mice (Shibata et al., 1997),  $Kras^{Lsl-G12D}$  mice (Jackson et al., 2001) and  $Tp53^{fl/fl}$  mice (Marino et al., 2000) were combined with  $Cdx2^{CreERT2/+}$  mice (Feng et al., 2013) to obtain  $APC^{lof}:KRAS^{G12}:TP53^{lof}$  intestinal tumours. All mice were in a C57BL/6 background. Healthy 8-10 weeks old male and female mice were used in the study. Mice used in this study had no previous history of drug administration, surgery or behavioural testing.

### Cell lines

DLD-1 and NCI-H508 were cultured in RPMI 1640 media (61870010, Thermo Fisher Scientific). HCT-116 and LS-174T were cultured in DMEM-F12 media (31331-028, Gibco). All cells were supplemented with 10% FBS (F7524, Sigma-Aldrich) and 1x pen strep (15140-122, Gibco). Cells were passaged using Trypsin-EDTA (25300-054,

Gibco) every 2-3 days or when they reached 90% confluence. All cells were maintained in a 37°C, 5% CO<sub>2</sub> atmosphere.

## METHOD DETAILS

### Organoid culture and drug screening

Organoid cultures were derived either from *Lgr5*<sup>GFP-IRES-CreER/+</sup> or wt mice as described previously (Sato et al., 2009). Cancer organoids were obtained by activating Cre-ERT2 in *APC*<sup>lof</sup>;*KRAS*<sup>G12</sup>;*TP53*<sup>lof</sup> mice. Cre-ERT2 was activated by a single intraperitoneal injection of tamoxifen dissolved in peanut oil (0.6 mg/ml) at a dose of 3 mg/kg. Six-month-old tumours were processed as follows: cut into small 2 mm pieces, washed thoroughly in PBS-EDTA at 4 °C, then homogenized with a teflon pestle in 1.5 ml tubes. Tissue homogenates were trypsinized in Trypsin-EDTA for 3-4 min and quickly pipetted up and down, approximately 100-200 times, using 200 µl tips to disrupt any cell aggregates. After centrifugation, the pellets were resuspended in ENR media, filtered through 70 µm cell strainers (BD Bioscience) and single cell suspensions were mixed with cold Matrigel (Corning) and plated in 96-well plates. Organoids were cultured in 10 cm dishes in EGF, Noggin, and R-Spondin (ENR)-containing media for wt and AKP organoids. For the screening, organoids were physically disrupted by pipetting (50-100 times) until small clusters of cells or crypts were obtained and seeded at a concentration of 50 clusters/crypts per well of a 96-well plate. A selection of 320 drugs from a library of FDA-approved drugs was provided by the Biomolecular Screening Facility at EPFL. The drugs were applied 24h after organoid seeding at a final concentration of 10 µM (0.1% DMSO in ENR medium), organoids were treated for 2-4 days and harvested in 300 µl of lysis buffer (ref. S1550S, NEB). Two to four biological replicates were used in the screening. Chosen timing of the whole assay is long enough to be able to detect differentiation and short enough to avoid organoid starvation. The AKP model was selected as it reflects progression to a more advanced CRC stage. All wells of a 96-well plate were imaged using bright field microscopy (DMI4000b inverted microscope, Leica) prior to harvesting. Organoids were treated with published drug cocktails for 2 days (10 µM DAPT or 2 µM IWP-2, 1.5 mM VPA (DC1 – drug cocktail 1) or 3 µM CHIR99021, 1.5 mM VPA or 2 µM IWP-2, 10 µM DAPT, 5 µM Gefitinib (DC2 – drug cocktail 2)) according to previously established protocols for the primer validation experiments (Yin et al., 2014; Basak et al., 2017). These cocktails were used as positive controls in each

separate 96-well experiment, along with untreated samples. Rescue experiments were performed with cholesterol:methyl- $\beta$ -cyclodextrin (M $\beta$ CD) complex (CD complex) at molar ratio of 1:8 with 242  $\mu$ g/ml cholesterol and 5mM M $\beta$ CD.

### **Targeted RNA-seq library construction**

#### ***mRNA isolation:***

All library preparation steps were performed in 96-well plates. mRNA was isolated with The New England Biolabs (NEB) Magnetic mRNA Isolation Kit (S1550S) and the isolation was performed according to the manufacturer's protocol. Briefly, the media was removed by inverting and flicking the plate. For lysing the organoids, 200  $\mu$ l of lysis buffer (S1550S, NEB) was added to each well using a multichannel pipette, Matrigel was disrupted by mechanical force using 200  $\mu$ l tips and a multichannel pipette. Lysis buffer containing organoids was transferred to special 1.2 ml low profile plates (AB-1127, Thermo Fisher Scientific) and additional 100-200  $\mu$ l of lysis buffer was added to each well. Organoids were disrupted by gently pipetting around 50 times. The 96-well plate was covered with aluminium sealing (6570, Corning) and put on a shaker at 300 rpm at RT for 10-15 min for efficient lysis. For mRNA binding, 22  $\mu$ l of Oligo d(T)<sub>25</sub> magnetic beads (S1550S, NEB) was used. The mRNA isolation was performed according to the manufacturer's protocol using 200  $\mu$ l of all wash buffers and 40-50  $\mu$ l of elution buffer. A magnet (AM10027, Thermo Fisher Scientific, magnetic stand - 96) was used for separating the magnetic beads during the washes and a second magnet (12331D, Thermo Fisher Scientific, DynaMag - 96 Side) was used for separating the magnetic beads during the elution step. The mRNA-bead suspension was transferred to a non-skirted PCR plate (AB0600, Thermo Fisher Scientific) and incubated in a PCR machine at 55°C for 140s prior to elution. Eluent was transferred to a V-bottom 96-well plate (6018323, 96-Well micro test plates, V-bottom, PP Ratiolab) and stored at -80°C.

#### ***RT reaction:***

RT reaction was carried out with Tth DNA-polymerase (11 480 022 001, Roche) according to the manufacturer's protocol in a non-skirted PCR plate (AB0600, Thermo Fisher Scientific) with 2.5 units of enzyme per reaction and 1  $\mu$ l of a 0.5  $\mu$ M mixture of gene specific reverse primers. Each reaction was pre-incubated at 72°C for 5 min without enzyme and then placed on ice. After the enzyme was added, the reaction was carried out at 72°C for 1 min and 60°C for 30 min. Obtained cDNA transcripts were cleaned with Agencourt AMPure magnetic beads (A63881, Beckman Coulter Inc.)

according to the manufacturer's protocol at a 1:1.8 ratio of cDNA synthesis mixture:AMPure beads and eluted with 40 µl of elution buffer.

### ***Amplification (1<sup>st</sup> and 2<sup>nd</sup> stage PCR):***

Cleaned cDNA was used as a template for 1<sup>st</sup> stage PCR and amplified with GoTaq G2 HS Polymerase (M7423, Promega). 2 µl of purified cDNA, 1 µl of 5 µM 1<sup>st</sup> stage common primer and 1 µl of a 0.5 µM mixture of 206 forward primers were used (**Supp. Table 6**). The following PCR program was used - **Supp. Table 7**. Obtained products were purified with AMPure magnetic beads as described above. 2<sup>nd</sup> stage PCR was carried out with GoTaq G2 HS Polymerase, 1 µl of each 5 µM Fw and Rev 2<sup>nd</sup> stage indexing primers, 2 µl of purified 1<sup>st</sup> stage PCR product, and amplified for 15-20 cycles with the following program: 94°C - 20 sec, 68°C – 15 sec, 72°C – 10 sec. Separately amplified libraries were pooled together by plates (96 samples) and purified twice on 2% agarose gels using QIAquick Gel Extraction Kit (28704, Qiagen) according to the manufacturer's protocol.

### ***Targeted RNA-sequencing and data processing***

Pooled libraries were sequenced at an average depth of 70,000 reads per sample. Sequencing was performed on a NextSeq500 platform using a single-end (Read1 Illumina sequencing primer), standard depth 75 nt run at the Gene Expression Core Facility (EPFL, Lausanne, Switzerland). Custom index1 sequencing primer (PAGE purified) was added to the machine: 5'- act ctg cgt tga tac cac tgc ttc cgc gga cag gc - 3'. Raw reads were first demultiplexed by standard Illumina protocols. The demultiplexed reads were processed with a custom script. Briefly, the first 10 bases of the each read contained UMI sequences and were deduplicated with UMI-tools. Deduplicated reads were processed with the Bowtie (ver. 1.2.2.)/samtools (ver. 1.2.7) pipeline and were aligned to a custom genome containing only the 206 target sequences. Aligned reads were counted with Cufflinks (ver. 2.2.1). The files and scripts for processing the raw data are deposited at <https://github.com/MaximNorkin91/Tornado-seq-protocol/>. Some samples were excluded from the analysis due to technical errors using a threshold of 1000 total raw reads per sample. Obtained raw count tables were pre-processed for the analysis with DESeq2 package (ver. 1.24.0) in R. Batch effect correction was performed between 96-well plates: each sample was downscaled to 50000 reads and average gene expression per plate was calculated; scaling factor for each gene was calculated by dividing

average gene expression value in one plate over average gene expression value in another plate. Normalization across the samples was performed based on the sample size (total counts per sample). Potential drug hits for the wt screen were determined as drugs altering more than 5 differentially expressed (DE) genes ( $|\log_2FC| > 1$ ,  $padj < 0.05$ ). The average Pearson's correlation for all compound hits using FC values was  $r=0.65$ . Drugs inducing differentiation in wt or transformed organoids were defined as downregulating/upregulating 5 or more stem cell/differentiation genes ( $|\log_2FC| > 1$ ,  $padj < 0.05$ ) upon drug treatment. Heatmaps were obtained with pheatmap (ver. 1.0.12.) or ComplexHeatmap (ver. 2.3.1.) packages. Clustering was performed with umap (ver. 0.2.3.1.) and nmf (0.21.0) packages. Mysort toolbox in R was used to obtain cell type compositions. Correlograms were obtained with the corrplot (ver. 0.84) package. Figure 5a was obtained via the corrplot function in R by analysing the 30 strongest (with more than 10 DE genes) drug hits. All genes with  $|\log_2FC| > 1$  were used for obtaining specific up- and down-regulated signatures. Figure 7a was obtained via the umap package in R using 40 most DE genes (significantly altered in at least 5 analysed samples) between selected drugs for treated AKP organoids. Figure S7B was obtained with nmf package in R using the same criteria as in Figure 7a, for both wt and cancer treated organoids.

### **Analysis of published single-cell sequencing data**

Single-cell RNA-sequencing data was obtained from GSE76408 and analysed with StemID2 and Seurat packages in R according to published guidelines. For the analysis, 240 genes of interest were extracted from the file "transcript\_counts\_intestine\_5days\_YFP.xls". Cells were filtered based on a 100 reads threshold for the sum of these 240 gene counts; cells with fewer reads were discarded. Based on the results of the analysis and subsequent sequencing results 206 out of 240 genes were selected.

### **FACS**

Lgr5<sup>+</sup> stem cells were sorted for the primer validation experiments. Intestinal crypts were isolated from 2-3 months old Lgr5<sup>GFPiresCreER/+</sup> mice as described previously (Sato et al., 2009). After several washes, the crypt cells were incubated with Trypsin-EDTA at 37°C for 5 min. Dead cells were excluded by DAPI staining. The GFP<sup>high</sup> (Lgr5<sup>+</sup>) cells were sorted on a BD FACSAria™ Fusion at FCCF EPFL.

FITC AnnexinV (640906, Biolegend) and DAPI (D9542, Sigma-Aldrich) were used for detection of apoptotic and dead cells in human cell lines and were analyzed on the

Attune NxT sorter. Cells in 96-well plates were trypsinized, washed with PBS and stained with AnnexinV antibody for 15 min in the dark (1:20 dilution), DAPI was added right before FACS analysis at 0.1 µg/ml. The results of this analysis are shown in **Suppl. Table 4**; drugs which scored for inducing increased cell death were considered as relevant.

### **Whole mount immunofluorescence stainings**

Whole mount stainings were performed as described (Dekkers et al., 2019) with few adjustments. In short, organoids were extracted from the wells by mechanical disruption of matrigel, sedimented by gravity to remove dead cells, and fixed for 1h in 4% PFA on ice. Tips precoated with 1% (wt/vol) BSA in PBS were used for handling. Fixed organoids were carefully washed with PBT buffer (0.1% (vol/vol) Tween in PBS) several times by aspirating the supernatant and adding 1ml of a fresh PBT solution, followed by an incubation for 10 min. Organoids were blocked in OWB buffer (0.1% (vol/vol) Triton X-100 and 0.2% (wt/vol) of BSA in PBS) for 15 min on ice and stained with primary antibodies (1:100-1:200 dilution) overnight at 4°C. Excess antibody was removed by washing several times with OWB buffer (3x30 minutes). Organoids were further incubated with secondary antibodies (1:500 dilution) for 4 hours at 4°C. After several washes organoids were resuspended in fructose-glycerol solution (pre-warmed to RT) and incubated for 20 min before embedding on slides as described previously. Imaging was performed on an inverted Leica SP8 confocal microscope and three-dimensional reconstructed confocal images were assembled in Fiji.

### **Immunofluorescence, Immunohistochemistry**

Organoids were extracted from at least 10 wells of a 48 well plate by mechanical disruption of the matrigel, sedimented by gravity to remove dead cells, and fixed overnight at in 4% PFA on ice. Tips precoated with 1% (wt/vol) BSA in PBS were used for handling. Fixed organoids were carefully washed with PBT buffer (0.1% (vol/vol) Tween in PBS), mixed with pre-warmed Histogel (HG-4000-012, Thermo Fisher Scientific) and embedded in a final volume of 40 µl placed on the lid of a 10cm dish to solidify (4°C, 5 min). The Histogel drop containing the organoids was placed in the cassette and processed for dehydration and paraffin embedding according to standard procedures. Sections of 5-7 µm thickness were cut using a rotary microtome (Hyrax M25 V2) and dried at 60°C for 1h. After rehydration, antigen retrieval was performed by heat induced epitope retrieval (HIER) using pH 6.0 citrate buffer. Slides were washed with



PBS (3x 5min). Organoid sections were surrounded with a hydrophobic barrier using a barrier pen and each section was incubated with a few drops of 1% BSA in PBS for blocking. Blocking solution was removed by tapping each slide several times. Sections were incubated with primary antibodies (1:100 – 1:400 diluted in 1% BSA in PBS) at 4°C overnight in a humidity chamber. Slides were washed in PBS (3x 5min) and incubated with the secondary antibody (1:1000 diluted in 1% BSA in PBS) for 1h at room temperature. Slides were washed in PBS (3x 5min) and mounted in water-based Fluoromount media (0100-01, SoutherBiotech). Dehydration, rehydration procedures and antigen retrieval were performed at HCF, EPFL. Imaging was performed on a confocal microscope (Leica SP8 Inverted).

PAS staining was performed at HCF, EFPL, using standard protocols. Sirius Red staining was performed on rehydrated sections: 10 min Hematoxyline, 2 min tap water, 30 sec (several dips) in 70% ethanol, 2h in Alkaline Sirius Red solution, rinsed with tap water, dried at room temperature and mounted in xylene based mounting media. Imaging was performed on an upright microscope (Leica DM5500 B).

The following primary and secondary antibodies were used: mouse anti-L-FABP (sc-374537, Santa Cruz), rat Anti-Serotonin (ab6336, Abcam), mouse anti-ChrA (sc-393941, Santa Cruz), donkey anti-mouse Alexa 568 (A-10037, ThermoFisher Scientific), goat anti-mouse Alexa 488 (A-11029, ThermoFisher Scientific), donkey anti-rat Alexa 488 (A-21208, ThermoFisher Scientific), goat anti-rat Alexa 568 (A-11077, ThermoFisher Scientific).

### **qPCR validation**

qPCRs during the validation step were performed with 50-100 organoids. RNA extraction was performed by two methods: with magnetic beads as described above and with a column-based method (74104, RNeasy Mini Kit, Qiagen) for comparison of the results. Both RNA isolation procedures were performed according to the manufacturer's protocol. RT reaction from RNA samples isolated with the column-based method was performed with 50-500 ng of total RNA per reaction using either Superscript II (18064014, Thermo Fisher Scientific) or Tth DNA polymerase for comparison of the results. PowerUp SYBR Green Master Mix (A25742, Thermo Fisher Scientific) was used for qPCR amplification with 10 µl total volume per reaction and 0.5 µM final primer concentration. qPCR was performed in a StepOnePlus thermocycler (Applied Biosystems) and relative gene expression was determined by the comparative CT method based on expression of 6 housekeeping genes. A list of the primer sequences

is provided in **Supp. Table 6**. Primers for qPCR comprise only the gene-specific part of the full-length primers for library construction from **Supp. Table 6**.

## 2.2.8 QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical tests were executed using GraphPad Prism (Prism 7 for Mac OS X) software or the statistical software R (version 4.0.3). Mean with SD values are reported in the figures where qPCR validation, FACS or immunostaining quantifications were made. Differences between variables of control and treated samples were assessed by 2-tailed Student's t test.

2-4 biological replicates were used in organoid drug screening. Potential drug hits for the wt screen were determined as drugs altering more than 5 differentially expressed (DE) genes ( $|\log_2FC| > 1$ ,  $padj < 0.05$ ). Drugs inducing differentiation in wt or transformed organoids were defined as downregulating/upregulating 5 or more stem cell/differentiation genes ( $|\log_2FC| > 1$ ,  $padj < 0.05$ ) upon drug treatment.

## 2.2.9 KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
mouse anti-L-FABP	Santa Cruz	sc-374537
rat anti-Serotonin	Abcam	ab6336
mouse anti-ChrA	Santa Cruz	sc-393941
donkey anti-mouse IgG Alexa 568	ThermoFisher Scientific	A-10037
goat anti-mouse IgG Alexa 488	ThermoFisher Scientific	A-11029
goat anti-rat IgG Alexa 568	ThermoFisher Scientific	A-11077
donkey anti-rat IgG Alexa 488	ThermoFisher Scientific	A-21208
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
human EGF Recombinant Protein	ThermoFisher Scientific	PHG0313
mouse R-spondin1 fusion to mouse Fc	Was produced and purified using a construct obtained from Dr. Calvin Kuo, Stanford University; (!!! INVALID CITATION !!!)	mRspo1_mFc

mouse Noggin	cDNA was synthesized according to NM_008711.2 followed by this linker and His tag (IEGRGGGSGGG SGGGSPGHHHH HHHH) produced and purified.	mNoggin_His
IWP-2	Stemgent	130-105-335
Gefetinib	Santa Cruz Biotechnology	sc-202166
CHIR99021	Sigma-Aldrich	SML1046-5MG
Valproic acid	Sigma-Aldrich	V0033000
Y-27632	Tocris	1254
DAPT	Sigma-Aldrich	D5942
Daunorubicin	Cayman chemical	CAY-14159-5
Atorvastatin	Sigma-Aldrich	PHR1422
Lovastatin	Adipogen Life Sciences	AG-CN2-0051-M025
Methyl-beta-cyclodextrin, average Mw 1310	Acros Organics	ACR37711-0050
Cholesterol	Adipogen Life Sciences	CDX-C0249-G025
DAPI for nucleic acid staining	Sigma-Aldrich	D9542
FITC Annexin V	Biolegend	640906
Peanut oil	Sigma-Aldrich	P2144
Tamoxifen	Sigma-Aldrich	T5648
<b>Critical Commercial Assays</b>		
Tth DNA-polymerase	Roche	11 480 022 001
Superscript II	Thermo Fisher Scientific	18064014
GoTaq G2 HS Polymerase	Promega	M7423
Oligo d(T) <sub>25</sub> magnetic beads	New England Biolabs	S1550S
Agencourt AMPure magnetic beads	Beckman Coulter Inc.	A63881
Magnetic mRNA Isolation Kit	New England Biolabs	S1550S
PowerUp SYBR Green Master Mix	Thermo Fisher Scientific	A25742
QIAquick Gel Extraction Kit	Qiagen	28704
RNeasy Mini Kit	Qiagen	74104
<b>Deposited Data</b>		
Gene expression profiles of treated organoids produced by TORNADO-seq	This paper	GSE157167
<b>Experimental Models: Cell Lines</b>		
HCT-116	ATCC	CCL-247
DLD-1	ATCC	CCL-221
NCI-H508	ATCC	CCL-253
LS-174T	ATCC	CCL-188
<b>Experimental Models: Organisms/Strains</b>		
Lgr5 <sup>GFP-IRES-CreER+</sup> , C57BL/6J,	Jackson Laboratory	Stock No.: 008875

WT, C57BL/6J	Jackson Laboratory	Stock No.: 000664
Cdx2 <sup>CreERT2/+</sup> :APC <sup>lof</sup> :KRAS <sup>G12</sup> :TP53 <sup>lof</sup> , C57BL/6J	Jackson Laboratory	Stock No.; 022390, 009045, 008179, 008462
<b>Oligonucleotides</b>		
Rev primers for RT reaction, see supp. table 6	This paper	NA
Fw primers for 1 <sup>st</sup> stage PCR, see supp. table 6	This paper	NA
Common rev primer for 1 <sup>st</sup> stage PCR, see supp. table 6	This paper	NA
Primers for 2 <sup>nd</sup> stage PCR (i5 index), see supp. table 6	This paper	NA
Primers for 2 <sup>nd</sup> stage PCR (i5 index), see supp. table 6	This paper	NA
<b>Software and Algorithms</b>		
FlowJo v10.7.1	FlowJo, LLC	<a href="https://www.flowjo.com/solutions/flowjo">https://www.flowjo.com/solutions/flowjo</a> ; RRID: SCR_008520
Oligoanalyzer	Integrated DNA Technologies	<a href="https://eu.idtdna.com/calc/analyzer">https://eu.idtdna.com/calc/analyzer</a>
ImageJ	(Schindelin et al., 2012)	<a href="https://fiji.sc/">Fiji</a> , RRID:SCR_002285
Primer Pooler v1.41	(Brown et al., 2017)	<a href="http://ssb22.user.srcf.net/pooler/">http://ssb22.user.srcf.net/pooler/</a>
BLAT	Jim Kent	<a href="https://genome.ucsc.edu">https://genome.ucsc.edu</a>
UMI-tools	(Smith et al., 2017)	<a href="https://github.com/CGATOxford/UMI-tools">https://github.com/CGATOxford/UMI-tools</a>
Bowtie v1.2.2.	(Langmead et al., 2009)	<a href="http://bowtie-bio.sourceforge.net/index.shtml">http://bowtie-bio.sourceforge.net/index.shtml</a>
Samtools v1.2.7	(Li et al., 2009; Li, 2011)	<a href="http://samtools.sourceforge.net">http://samtools.sourceforge.net</a>
Cufflinks v2.2.1	(Trapnell et al., 2010)	<a href="https://github.com/cole-trapnell-lab/cufflinks">https://github.com/cole-trapnell-lab/cufflinks</a>
DESeq2	(Love et al., 2014)	10.1186/s13059-014-0550-8
pheatmap v1.0.12	(Kolde, 2019)	<a href="https://github.com/raivokolde/pheatmap">https://github.com/raivokolde/pheatmap</a>
ComplexHeatmap v2.3.1.	(Gu et al., 2016)	10.18129/B9.bioc.ComplexHeatmap
umap v0.2.3.1.	(McInnes, 2020)	<a href="https://cran.r-project.org/web/packages/umap/">https://cran.r-project.org/web/packages/umap/</a>
Nmf v0.21.0	(Gaujoux and Seoighe, 2010)	<a href="https://cran.r-project.org/web/packages/NMF/">https://cran.r-project.org/web/packages/NMF/</a>

mysort	(Chen et al., 2018)	<a href="https://testtoolshe.d.g2.bx.psu.edu/view/moneycat/mysort/e3afe097e80a">https://testtoolshe.d.g2.bx.psu.edu/view/moneycat/mysort/e3afe097e80a</a>
corrplot v0.84	(Wei, 2017)	<a href="https://cran.r-project.org/web/packages/corrplot/">https://cran.r-project.org/web/packages/corrplot/</a>
Other		
Non-skirted PCR 96-well plate, for mRNA elution	Thermo Fisher Scientific	AB0600
V-bottom 96-well plate, for mRNA storage	PP Ratiolab	6018323
DynaMag - 96 Side magnet, for mRNA elution	Thermo Fisher Scientific	12331D
Magnetic Stand-96, for mRNA isolation	Thermo Fisher Scientific	AM10027
Low profile 96-well plates, for mRNA isolation	Thermo Fisher Scientific	AB-1127
96-well plates, for organoid growth	Grenier Bio-One	655090
Aluminium sealing	Corning	6570
Histogel	Thermo Fisher Scientific	HG-4000-012
Matrigel	Corning	356231

## 2.2.10 References

- Adams C.M., Goldstein J.L. Brown M.S. **Cholesterol-induced conformational change in SCAP enhanced by Insig proteins and mimicked by cationic amphiphiles.** *Proc. Natl. Acad. Sci. USA.* 2003; **100**: 10647-10652
- Basak O., Beumer J., Wiebrands K., Seno H., van Oudenaarden A. Clevers H. **Induced Quiescence of Lgr5+ Stem Cells in Intestinal Organoids Enables Differentiation of Hormone-Producing Enteroendocrine Cells.** *Cell Stem Cell* 2017; **20**: 177-190
- Basu S., Cheriyaundath S. Ben-Ze'ev A. **Cell-cell adhesion: linking Wnt/beta-catenin signaling with partial EMT and stemness traits in tumorigenesis.** *F1000Res* 2018; **7**: 1488
- Brown S.S., Chen Y.W., Wang M., Clipson A., Ochoa E. Du M.Q. **PrimerPooler: automated primer pooling to prepare library for targeted sequencing.** *Biol Methods Protoc* 2017; **2**: bpx006
- Buczacki S.J.A., Popova S., Biggs E., Koukorava C., Buzzelli J., Vermeulen L., Hazelwood L., Francies H., Garnett M.J. Winton D.J. **Itraconazole targets cell cycle heterogeneity in colorectal cancer.** *The Journal of experimental medicine* 2018; **215**: 1891-1912
- Bush E.C., Ray F., Alvarez M.J., Realubit R., Li H., Karan C., Califano A. Sims P.A. **PLATE-Seq for genome-wide regulatory network analysis of high-throughput screens.** *Nat Commun* 2017; **8**: 105
- Butler A., Hoffman P., Smibert P., Papalexi E. Satija R. **Integrating single-cell transcriptomic data across different conditions, technologies, and species.** *Nat Biotechnol* 2018; **36**: 411-420
- Chen S.H., Kuo W.Y., Su S.Y., Chung W.C., Ho J.M., Lu H.H. Lin C.Y. **A gene profiling deconvolution approach to estimating immune cell composition from complex tissues.** *BMC Bioinformatics* 2018; **19**: 154
- Christian A.E., Haynes M.P., Phillips M.C. Rothblat G.H. **Use of cyclodextrins for manipulating cellular cholesterol content.** *J Lipid Res* 1997; **38**: 2264-2272
- Dekkers J.F., Alieva M., Wellens L.M., Ariese H.C.R., Jamieson P.R., Vonk A.M., Amatngalim G.D., Hu H., Oost K.C., Snippert H.J.G., *et al.* **High-resolution 3D imaging of fixed and cleared organoids.** *Nature protocols* 2019; **14**: 1756-1771
- Feng Y., Sentani K., Wiese A., Sands E., Green M., Bommer G.T., Cho K.R. Fearon E.R. **Sox9 induction, ectopic Paneth cells, and mitotic spindle axis defects in mouse colon adenomatous epithelium arising from conditional biallelic Apc inactivation.** *Am J Pathol* 2013; **183**: 493-503
- Fumagalli A., Oost K.C., Kester L., Morgner J., Bornes L., Bruens L., Spaargaren L., Azkanaz M., Schelfhorst T., Beerling E., *et al.* **Plasticity of Lgr5-Negative Cancer Cells Drives Metastasis in Colorectal Cancer.** *Cell Stem Cell* 2020; **26**: 569-578
- Gao H., Korn J.M., Ferretti S., Monahan J.E., Wang Y., Singh M., Zhang C., Schnell C., Yang G., Zhang Y., *et al.* **High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response.** *Nat Med* 2015; **21**: 1318-1325
- Gaujoux R. Seoighe C. **A flexible R package for nonnegative matrix factorization.** *BMC Bioinformatics* 2010; **11**: 367

- Grun D., Lyubimova A., Kester L., Wiebrands K., Basak O., Sasaki N., Clevers H. van Oudenaarden A. **Single-cell messenger RNA sequencing reveals rare intestinal cell types.** *Nature* 2015; **525**: 251-255
- Grun D., Muraro M.J., Boisset J.C., Wiebrands K., Lyubimova A., Dharmadhikari G., van den Born M., van Es J., Jansen E., Clevers H., *et al.* **De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data.** *Cell Stem Cell* 2016; **19**: 266-277
- Gu Z., Eils R. Schlesner M. **Complex heatmaps reveal patterns and correlations in multidimensional genomic data.** *Bioinformatics* 2016; **32**: 2847-2849
- Guinney J., Dienstmann R., Wang X., de Reynies A., Schlicker A., Soneson C., Marisa L., Roepman P., Nyamundanda G., Angelino P., *et al.* **The consensus molecular subtypes of colorectal cancer.** *Nat Med* 2015; **21**: 1350-1356
- Hagggar F. Boushey R. **Colorectal Cancer Epidemiology: Incidence, Mortality, Survival, and Risk Factors.** *Clin Colon Rectal Surg* 2009; **22**: 191-197
- Jackson E.L., Willis N., Mercer K., Bronson R.T., Crowley D., Montoya R., Jacks T. Tuveson D.A. **Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras.** *Genes Dev* 2001; **15**: 3243-3248
- Karhausen J., Furuta G.T., Tomaszewski J.E., Johnson R.S., Colgan S.P. Haase V.H. **Epithelial hypoxia-inducible factor-1 is protective in murine experimental colitis.** *J Clin Invest* 2004; **114**: 1098-1106
- Kim W.K., Kwon Y., Jang M., Park M., Kim J., Cho S., Jang D.G., Lee W.B., Jung S.H., Choi H.J., *et al.* **beta-catenin activation down-regulates cell-cell junction-related genes and induces epithelial-to-mesenchymal transition in colorectal cancers.** *Sci Rep* 2019; **9**: 18440
- Kolde R. **pheatmap: Implementation of heatmaps that offers more control over dimensions and appearance.** *R package version 1.0.12* 2019: <https://cran.r-project.org/web/packages/pheatmap/index.html>
- Korade Z., Kim H.Y., Tallman K.A., Liu W., Koczok K., Balogh I., Xu L., Mirnics K. Porter N.A. **The Effect of Small Molecules on Sterol Homeostasis: Measuring 7-Dehydrocholesterol in Dhcr7-Deficient Neuro2a Cells and Human Fibroblasts.** *J Med Chem* 2016; **59**: 1102-1115
- Langmead B., Trapnell C., Pop M. Salzberg S.L. **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009; **10**: R25
- Li B., Flaveny C.A., Giambelli C., Fei D.L., Han L., Hang B.I., Bai F., Pei X.H., Nose V., Burlingame O., *et al.* **Repurposing the FDA-approved pinworm drug pyriminidyl as a novel chemotherapeutic agent for intestinal polyposis.** *PLoS One* 2014; **9**: e101969
- Li H. **A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data.** *Bioinformatics* 2011; **27**: 2987-2993
- Li H., Courtois E.T., Sengupta D., Tan Y., Chen K.H., Goh J.J.L., Kong S.L., Chua C., Hon L.K., Tan W.S., *et al.* **Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors.** *Nat Genet* 2017; **49**: 708-718
- Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G. Durbin R. **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics* 2009; **25**: 2078-2079

- Love M.I., Huber W. Anders S. **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** *Genome Biol* 2014; **15**: 550
- Marino S., Vooijs M., van Der Gulden H., Jonkers J. Berns A. **Induction of medulloblastomas in p53-null mutant mice by somatic inactivation of Rb in the external granular layer cells of the cerebellum.** *Genes Dev* 2000; **14**: 994-1004
- McInnes L.H., J.; Melville, J. **UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.** 2020: arXiv:1802.03426
- Miyamoto S., Narita T., Komiya M., Fujii G., Hamoya T., Nakanishi R., Tamura S., Kurokawa Y., Takahashi M. Mutoh M. **Novel screening system revealed that intracellular cholesterol trafficking can be a good target for colon cancer prevention.** *Sci Rep* 2019; **9**: 6192
- Moebius F.F., Striessnig J. Glossmann H. **The mysteries of sigma receptors: new family members reveal a role in cholesterol synthesis.** *Trends Pharmacol Sci* 1997; **18**: 67-70
- Ordóñez-Morán P., Dafflon C., Imajo M., Nishida E. Huelsken J. **HOXA5 Counteracts Stem Cell Traits by Inhibiting Wnt Signaling in Colorectal Cancer.** *Cancer Cell* 2015; **28**: 815-829
- Sato T., Vries R.G., Snippert H.J., van de Wetering M., Barker N., Stange D.E., van Es J.H., Abo A., Kujala P., Peters P.J., *et al.* **Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche.** *Nature* 2009; **459**: 262-265
- Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., Preibisch S., Rueden C., Saalfeld S., Schmid B., *et al.* **Fiji: an open-source platform for biological-image analysis.** *Nat Methods* 2012; **9**: 676-682
- Schmitt M., Schewe M., Sacchetti A., Feijtel D., van de Geer W.S., Teeuwssen M., Sleddens H.F., Joosten R., van Royen M.E., van de Werken H.J.G., *et al.* **Paneth Cells Respond to Inflammation and Contribute to Tissue Regeneration by Acquiring Stem-like Features through SCF/c-Kit Signaling.** *Cell Rep* 2018; **24**: 2312-2328
- Schutgens F. Clevers H. **Human Organoids: Tools for Understanding Biology and Treating Diseases.** *Annu Rev Pathol* 2020; **15**: 211-234
- Schütte M., Risch T., Abdavi-Azar N., Boehnke K., Schumacher D., Keil M., Yildirim R., Jandrasits C., Borodina T., Amstislavskiy V., *et al.* **Molecular dissection of colorectal cancer in pre-clinical models identifies biomarkers predicting sensitivity to EGFR inhibitors.** *Nat Commun* 2017; **8**: 14262
- Shibata H., Toyama K., Shioya H., Ito M., Hirota M., Hasegawa S., Matsumoto H., Takano H., Akiyama T., Toyoshima K., *et al.* **Rapid colorectal adenoma formation initiated by conditional targeting of the Apc gene.** *Science* 1997; **278**: 120-123
- Shim J.S., Li R.J., Lv J., Head S.A., Yang E.J. Liu J.O. **Inhibition of angiogenesis by selective estrogen receptor modulators through blockade of cholesterol trafficking rather than estrogen receptor antagonism.** *Cancer Letters* 2015; **362**: 106-115
- Simon J.M., Paranjape S.R., Wolter J.M., Salazar G. Zylka M.J. **High-throughput screening and classification of chemicals and their effects on neuronal gene expression using RASL-seq.** *Sci Rep* 2019; **9**: 4529
- Smith T., Heger A. Sudbery I. **UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy.** *Genome research* 2017; **27**: 491-499



Tashiro T., Okuyama H., Endo H., Kawada K., Ashida Y., Ohue M., Sakai Y. Inoue M. **In vivo and ex vivo cetuximab sensitivity assay using three-dimensional primary culture system to stratify KRAS mutant colorectal cancer.** *PLoS One* 2017; **12**: e0174151

Teder H., Koel M., Paluoja P., Jatsenko T., Rekker K., Laisk-Podar T., Kukuskina V., Velthut-Meikas A., Fjodorova O., Peters M., *et al.* **TAC-seq: targeted DNA and RNA sequencing for precise biomarker molecule counting.** *NPJ Genom Med* 2018; **3**: 34

Trapnell C., Williams B.A., Pertea G., Mortazavi A., Kwan G., van Baren M.J., Salzberg S.L., Wold B.J. Pachter L. **Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.** *Nat Biotechnol* 2010; **28**: 511-515

van de Wetering M., Francies Hayley E., Francis Joshua M., Bounova G., Iorio F., Pronk A., van Houdt W., van Gorp J., Taylor-Weiner A., Kester L., *et al.* **Prospective Derivation of a Living Organoid Biobank of Colorectal Cancer Patients.** *Cell* 2015; **161**: 933-945

Verissimo C.S., Overmeer R.M., Ponsioen B., Drost J., Mertens S., Verlaan-Klink I., Gerwen B.v., van der Ven M., Wetering M.v.d., Egan D.A., *et al.* **Targeting mutant RAS in patient-derived colorectal cancer organoids by combinatorial drug screening.** *Elife* 2016; **5**: e18489

Vorneveld P.W., Reimers M.S., Bastiaannet E., Jacobs R.J., van Eijk R., Zanders M.M.J., Herings R.M.C., van Herk-Sukel M.P.P., Kodach L.L., van Wezel T., *et al.* **Statin Use After Diagnosis of Colon Cancer and Patient Survival.** *Gastroenterology* 2017; **153**: 470-479

Wages P.A., Kim H.H., Korade Z. Porter N.A. **Identification and characterization of prescription drugs that change levels of 7-dehydrocholesterol and desmosterol.** *J Lipid Res* 2018; **59**: 1916-1926

Wang B., Rong X., Palladino E.N.D., Wang J., Fogelman A.M., Martin M.G., Alrefai W.A., Ford D.A. Tontonoz P. **Phospholipid Remodeling and Cholesterol Availability Regulate Intestinal Stemness and Tumorigenesis.** *Cell Stem Cell* 2018; **22**: 206-220

Wei T.S., V.; Levy, M.; Xie, Y.; Jin, Y.; Zemla, J. **corrplot: Visualization of a Correlation Matrix.** 2017: <https://cran.r-project.org/web/packages/corrplot/index.html>

Wiegering A., Uthe F.W., Huttenrauch M., Muhling B., Linnebacher M., Krummenast F., Germer C.T., Thalheimer A. Otto C. **The impact of pyrvinium pamoate on colon cancer cell viability.** *Int J Colorectal Dis* 2014; **29**: 1189-1198

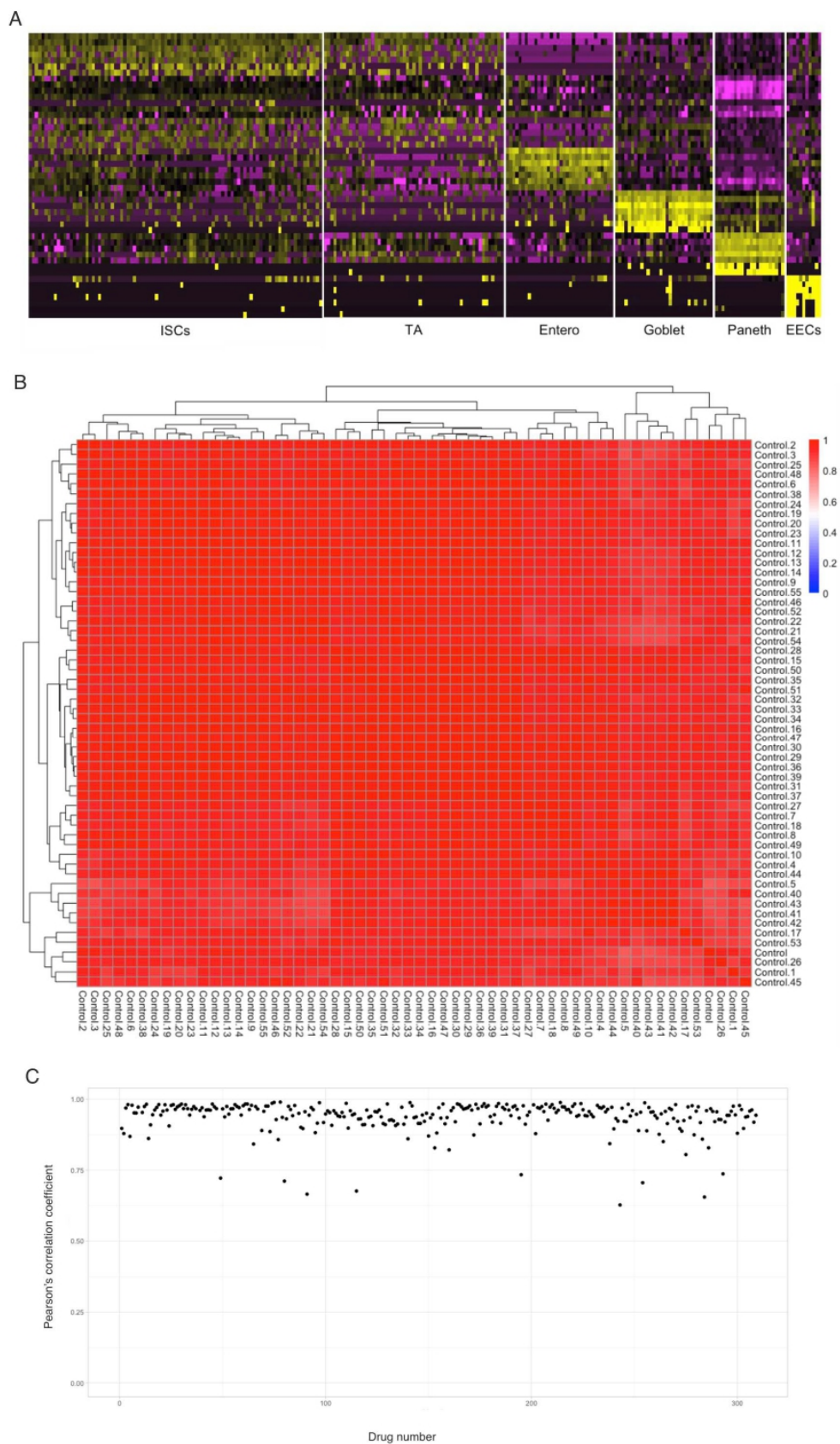
Ye C., Ho D.J., Neri M., Yang C., Kulkarni T., Randhawa R., Henault M., Mostacci N., Farmer P., Renner S., *et al.* **DRUG-seq for miniaturized high-throughput transcriptome profiling in drug discovery.** *Nat Commun* 2018; **9**: 4307

Yin X., Farin H.F., van Es J.H., Clevers H., Langer R. Karp J.M. **Niche-independent high-purity cultures of Lgr5+ intestinal stem cells and their progeny.** *Nat Methods* 2014; **11**: 106-112

Zhan T., Ambrosi G., Wandmacher A.M., Rauscher B., Betge J., Rindtorff N., Haussler R.S., Hinsenkamp I., Bamberg L., Hessling B., *et al.* **MEK inhibitors activate Wnt signalling and induce stem cell plasticity in colorectal cancer.** *Nat Commun* 2019; **10**: 2197

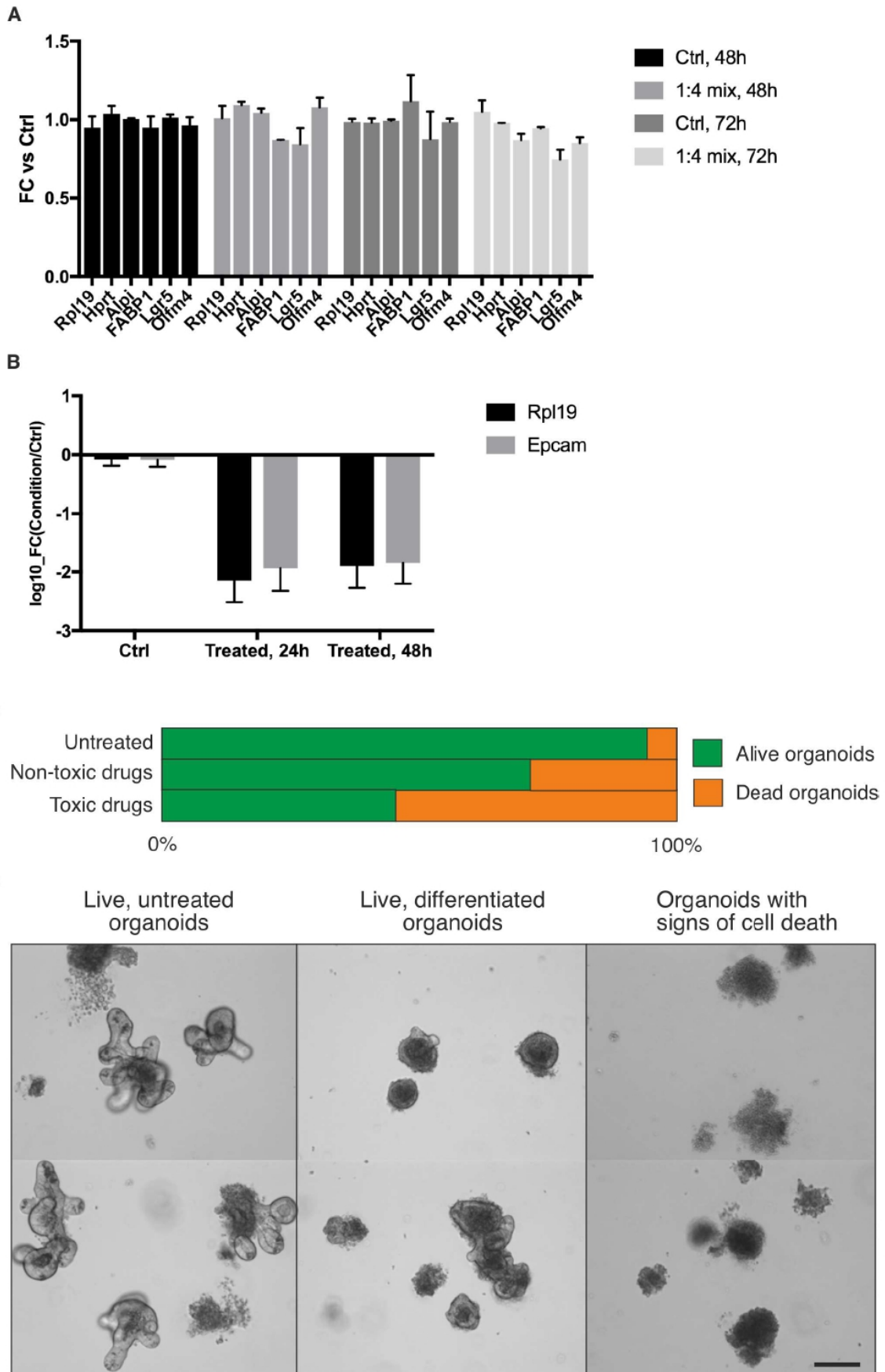
## 2.2.11 Supplementary figures and tables

### 2.2.11.1 Supplementary figures



**Figure S1. Validation of intestinal gene set and TORNAOD-seq accuracy and reproducibility. Related to Figure 1.**

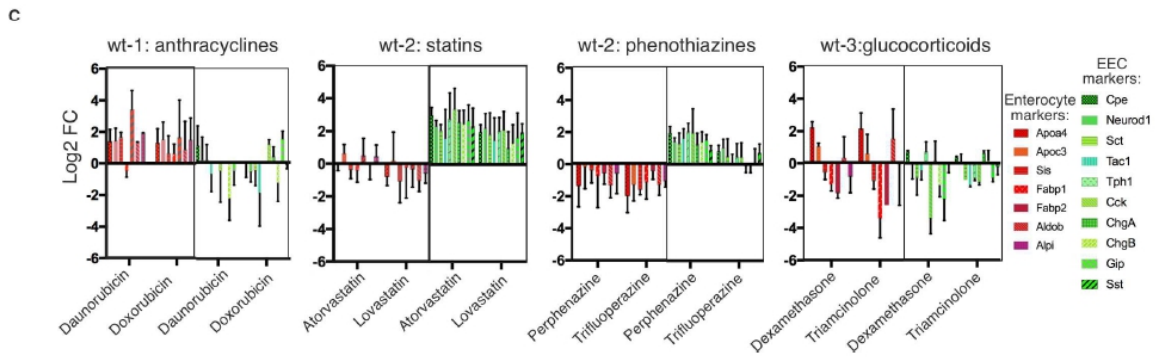
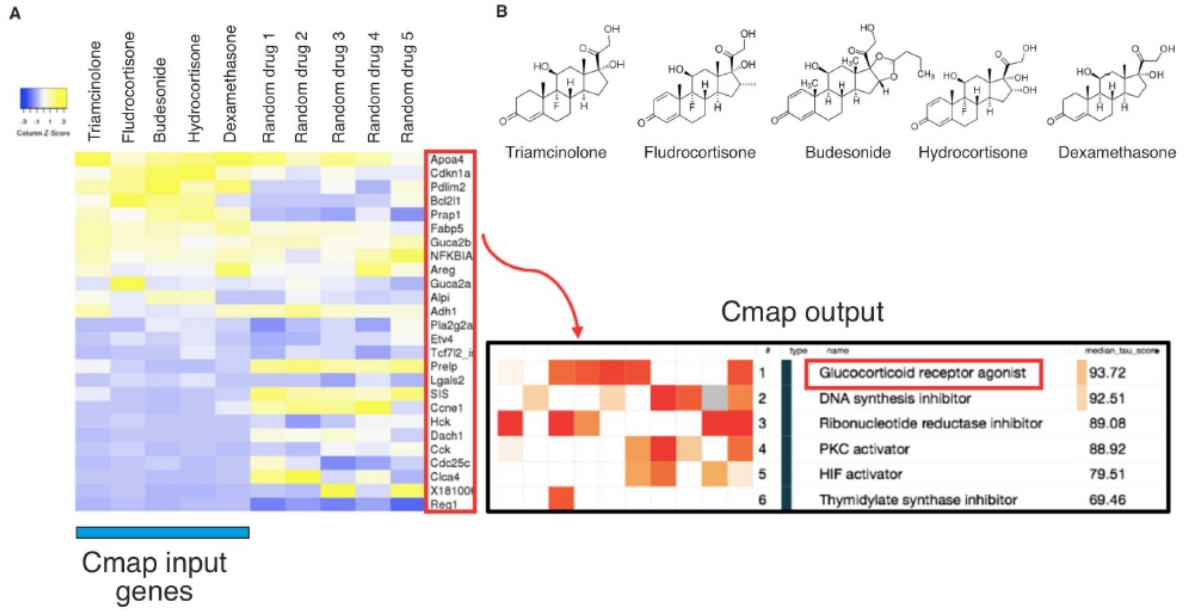
**A.** Clustering of single cells based on the expression of only our 206 gene signature into the different intestinal populations using the Seurat algorithm. Columns represent different single cells, rows represent cell type specific markers. Only six main cell types are presented. **B.** Pearson's correlation between untreated replicates of wt organoids as analysed by TORNAOD-seq. All samples were down-sampled to an average number of reads and batch correction was performed between the samples from different plates for the comparison. **C.** Average Pearson's correlation coefficients between replicates across all drug treated samples. For the comparison, all samples were down-sampled to an average number of reads and batch correction was performed between samples from different plates.



**Figure S2. Contribution of dead cells to the measured gene expression profiles.**

**Related to Figure 2.**

**A.** RNA quantification was performed by qPCR for the indicated genes either from untreated organoids (ctrl) or from a 1:4 mixture of untreated and 20 $\mu$ M daunorubicin-treated, dead organoids. Measured RNA levels were only altered up to 1.1-fold when comparing control and mixed samples. Average values of replicates are shown, error bars represent SD values. **B.** Comparison of RNA quantities isolated from either untreated (one 96 well) or 20 $\mu$ M daunorubicin treated organoids (four 96 wells) accessed by qPCR for the housekeeping genes Rpl19 and Epcam. Fold changes over untreated controls are represented. Average  $\log_{10}(\text{FC})$  values of replicates values are shown, error bars represent SD values. **C.** Percentage of wt organoid cultures showing signs of cell death either untreated or treated with drugs from cluster 1 (toxic drugs) or drugs of the other clusters (non-toxic drugs). Percentages are averaged from the percentage of organoids with signs of cell death per well over all samples from each group. **D.** Examples of different morphologies of wt organoids untreated or treated with drugs for 2 days as used in Figure S2C. Scale bar is 50  $\mu$ m. Top and bottom images show technical replicates. “Live, differentiated organoids” are the images of organoids treated with 10  $\mu$ M Ifenprodil, and “Organoids with signs of cell death” are organoids treated with 10  $\mu$ M Epirubicin.

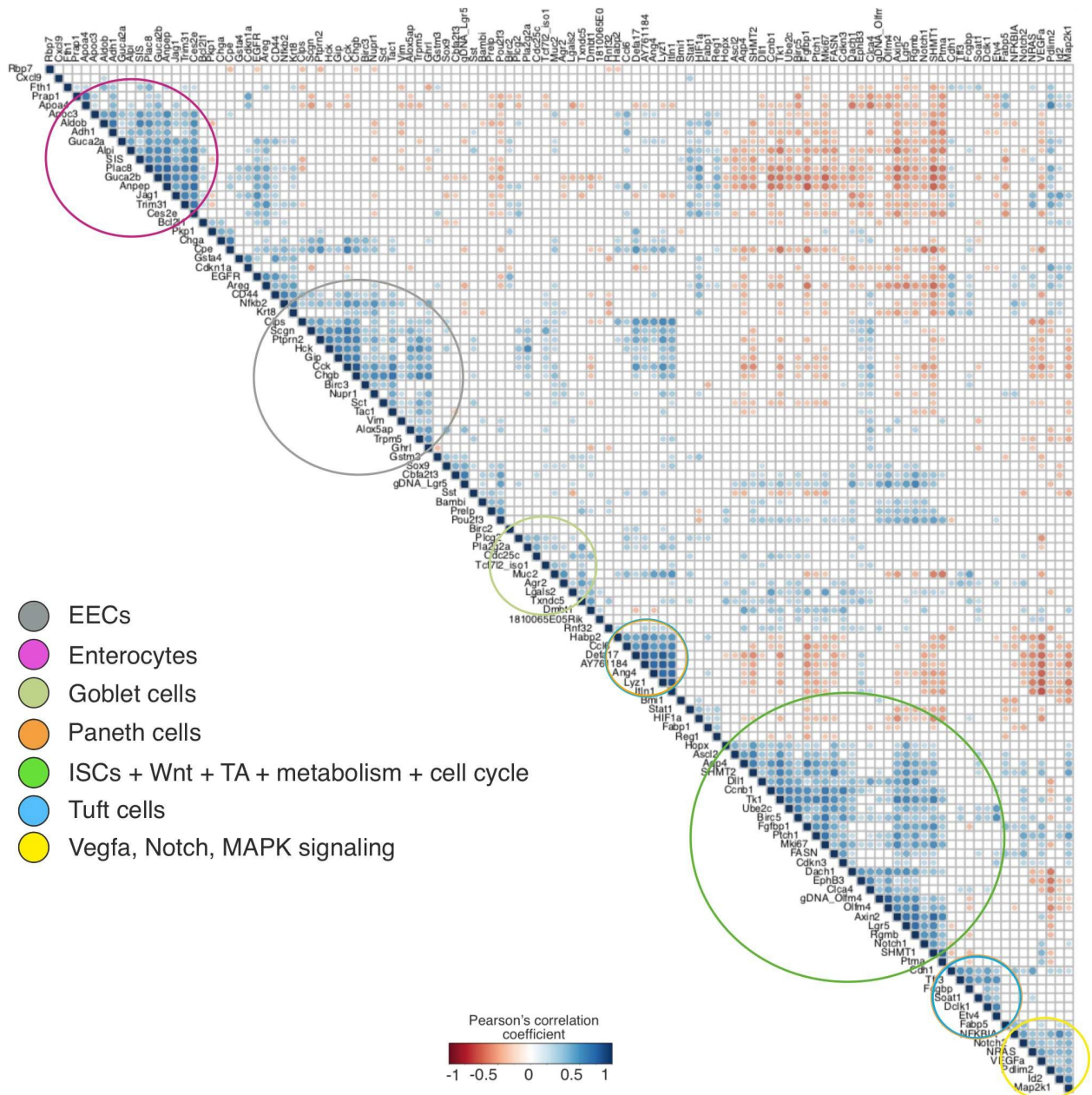


**Figure S3. Confirmation of TORNADO-seq validity and robustness.**

**Related to Figure 3.**

**A.** Comparison of targeted RNA-seq results with CMAP database (Subramanian et al., 2017). Left: Heatmap of gene expression FC values vs. untreated controls for glucocorticoid-treated wt organoids (first 5 columns) and randomly selected drugs acting as negative controls (last 5 columns). Colours display column Z-score values highlighting upregulated (yellow) and downregulated (blue) genes vs. untreated controls. Genes in the red box are consisting the glucocorticoid-response signature as obtained by TORNADO-seq which was used as input for CMAP. Right bottom: CMAP output showing that the submitted gene signature is recognized as being induced by glucocorticoid drugs. **B.** Chemical structure of glucocorticoids in a. **C.** qPCR confirmation of targeted RNA-seq profiles for 4 drug groups (statins, phenothiazines, glucocorticoids, anthracyclines) from 3 clusters (wt-1, wt-2, wt-3) using two representative drugs each. Shown are fold changes in enterocyte- and EEC-specific marker expression. Statins and phenothiazines (group 1 and 2) correspond to cluster wt-2 from **Figure 3A**, glucocorticoids (group 3) corresponds to cluster wt-3 from **Figure 3A** and anthracyclines (group 4) correspond to cluster wt-1. Red colours represent enterocyte markers, green represents EEC markers. Average log<sub>2</sub>(FC) values of replicates are shown, error bars represent SD values.



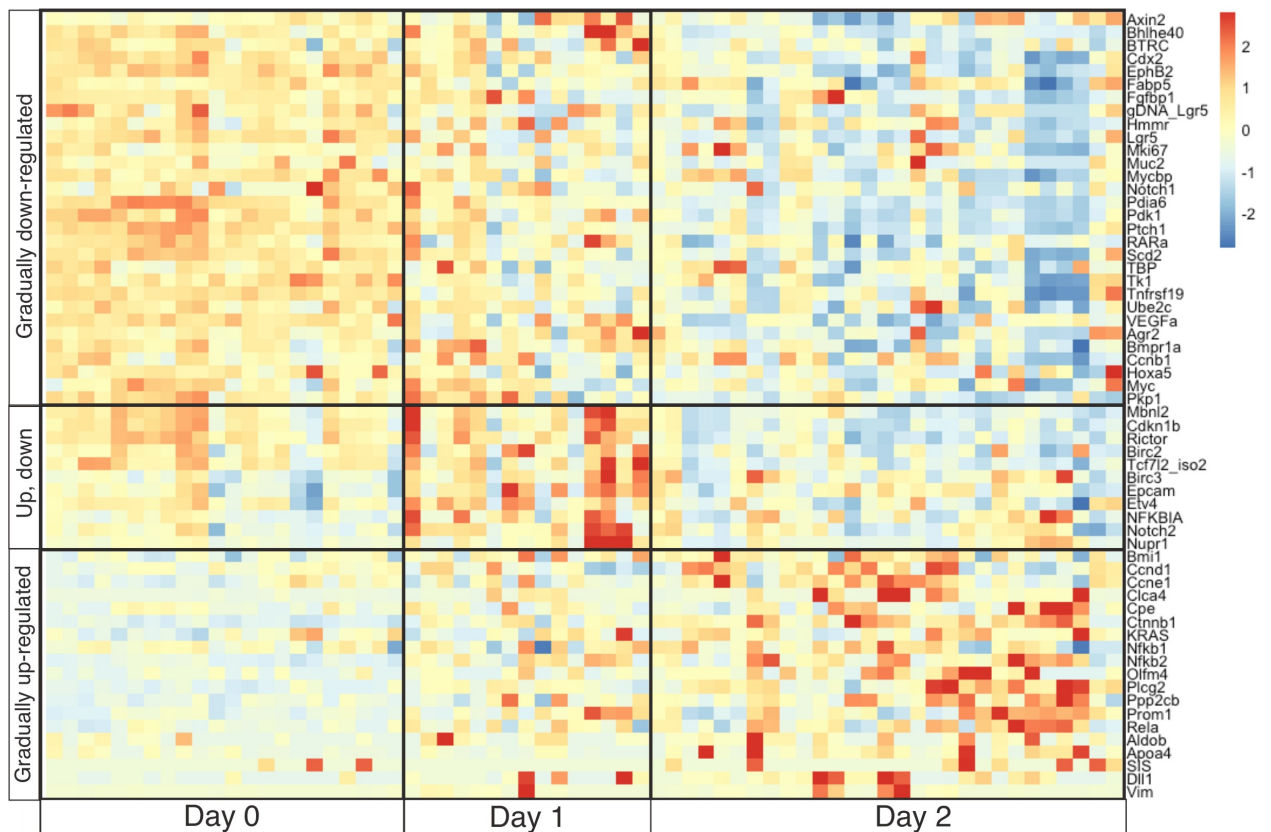


**Figure S4. Gene-gene correlations for co-regulated genes in drug-treated wt organoids.**

**Related to Figure 5.**

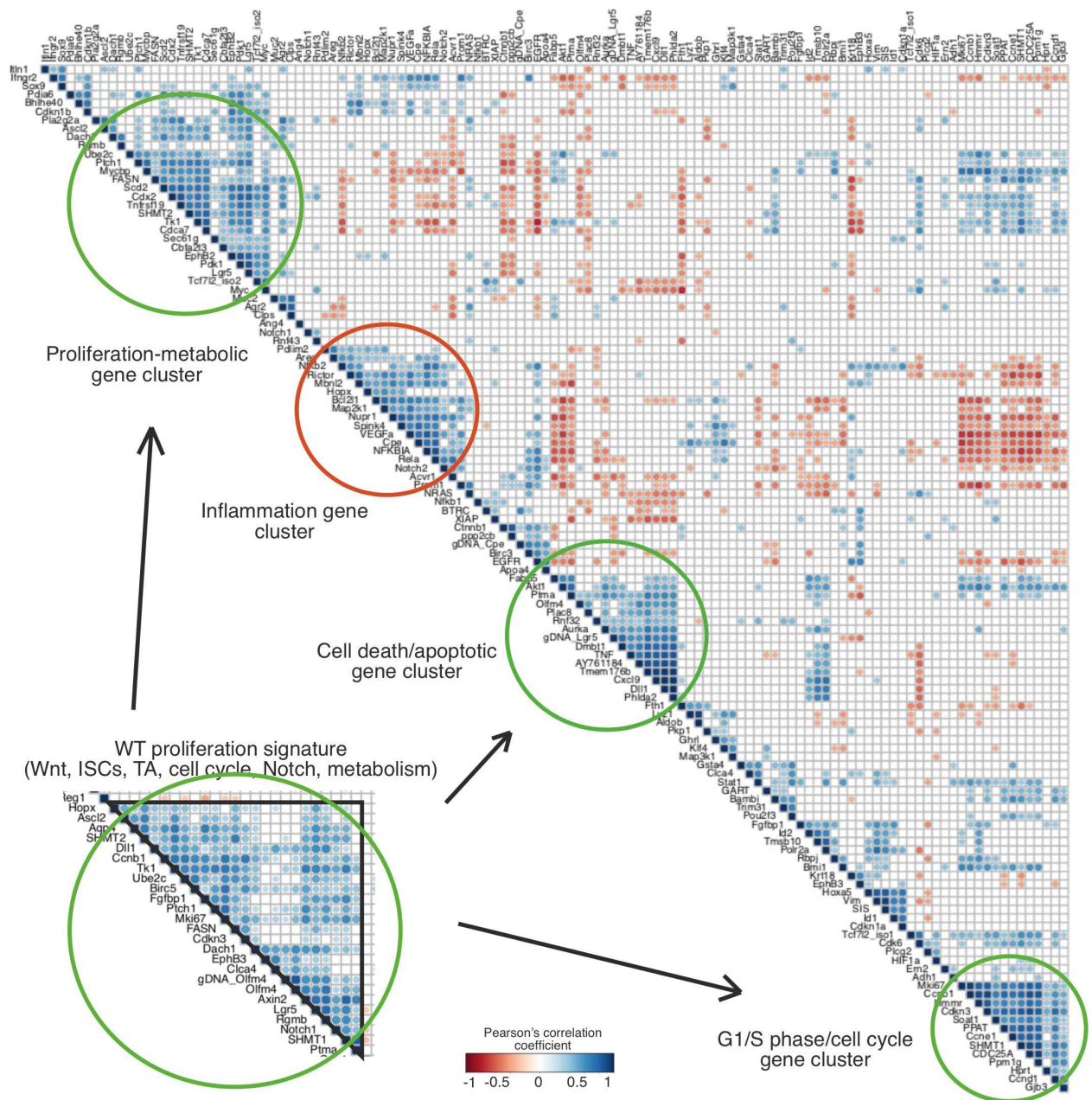
Circles highlight main cell type specific genes coregulated upon drug treatments with colours corresponding to their signature. The green circle depicts proliferation-related genes from the ISC, Wnt, TA, metabolism and cell cycle signatures. Aggregated, pathway correlations summarized in Figure 5A are based on this analysis. Tuft and goblet cell markers are highly co-regulated and highlighted by double circles of corresponding colours.





**Figure S5. Drug responses over time for AKP cancer organoids show the existence of early and late cell responses. Related to Figure 6.**

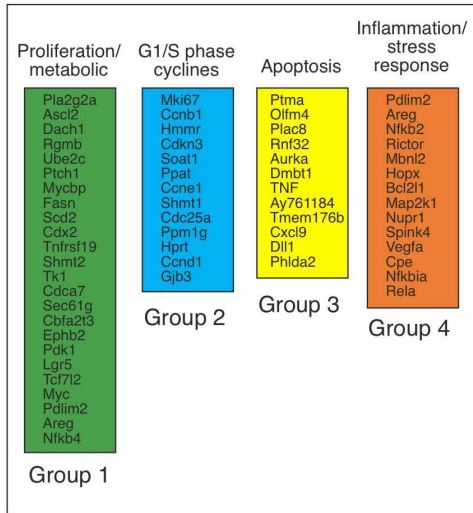
Heatmap displaying drug responses over time for AKP cancer organoids. Columns represent drug treated samples (day 1, 2) or untreated controls (day 0). The number of replicates for day 1 is half of that for day 2. The data are shown as  $\log_{10}[\text{counts} + 1]$  – expression value and displayed as Z- scores per gene (colour bar). “Up, down” refers to transient upregulation of gene expression at day1 and downregulation at day 2 compared to day 0 (untreated controls).



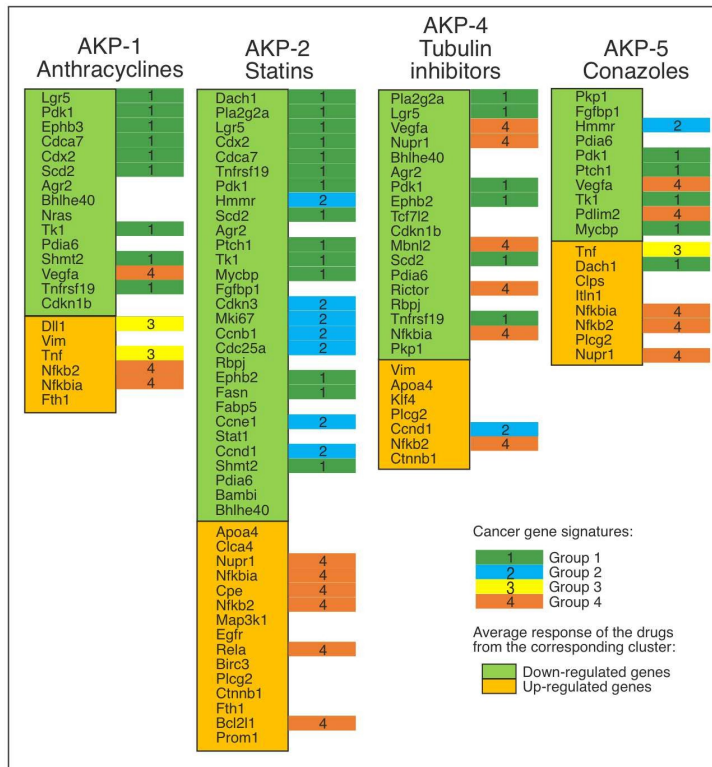
**Figure S6. Gene-gene correlations for co-regulated gene clusters in drug-treated AKP cancer organoids. Related to Figure 6.**

Gene-gene correlations for co-regulated gene clusters in drug-treated AKP cancer organoids. Circles highlight clusters of genes according to GO term enrichment. Green circles contain genes from signatures such as Wnt, ISCs, TA, metabolism, and cell cycle of wt treated organoids (corresponding part of Figure S4 shown in the lower left corner depicts the clustering of these genes in the wt system). The red circle contains genes from signatures such as NF- $\kappa$ B, apoptosis, and angiogenesis. Arrows indicate distribution of wt proliferation genes into 3 new independent clusters in AKP cancer system.

A

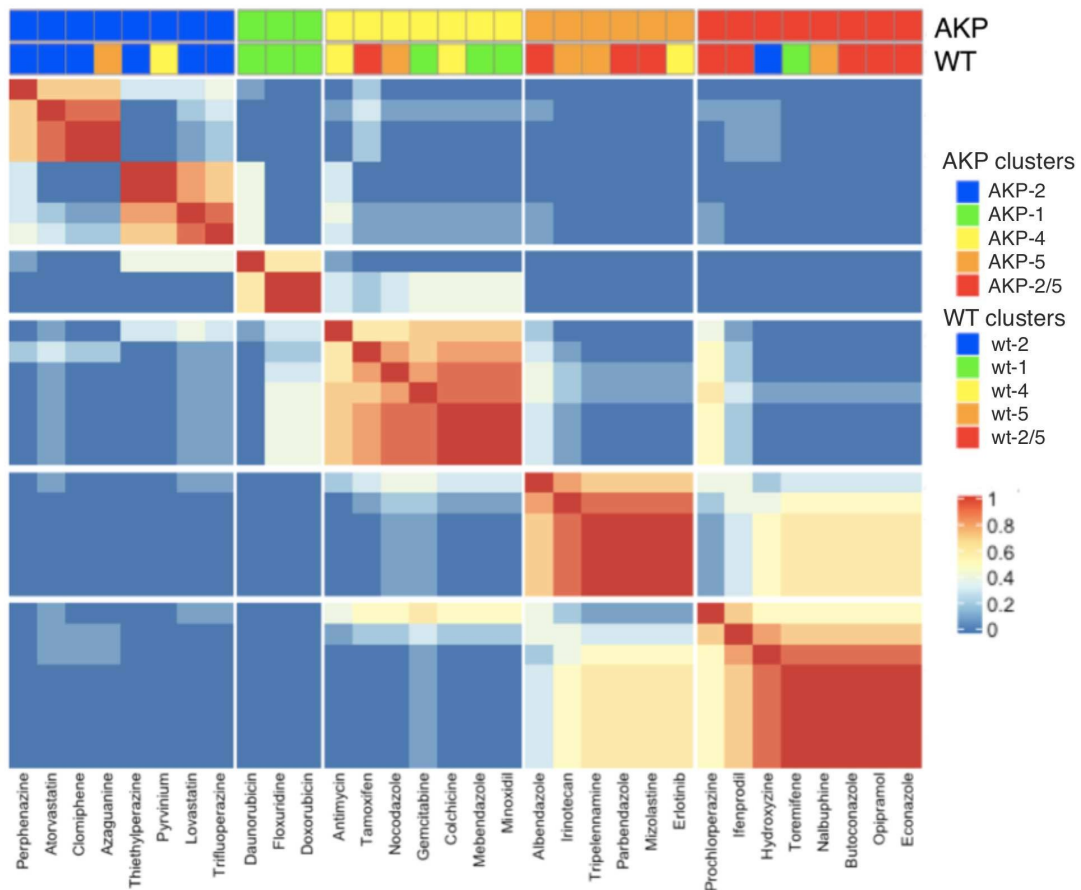


## Cancer gene signatures



## Drug cluster signatures

B



**Figure S7. Drug clusters caused different cancer gene signatures.**

**Related to Figures 6 and 7.**

**A.** Left box: Identified gene signatures in cancer organoids extracted from Figure S6. Right box: 4 drug groups: depicted downregulated and upregulated genes in light green and light orange, respectively. Beside each gene name is the number of the corresponding gene cluster from the left box (also depicted with the colour code). Up- and down-regulated genes are calculated as average per corresponding drug group. **B.** Clustering of drugs for treated wt and AKP samples. NMF heatmap for drug treated AKP organoids is displayed on the bottom panel. The colour code represents the correlation between the samples as calculated by NMF. Clustering is based on the expression of the most DE genes among AKP treated samples as in **Figure 6A**. See methods for details. On the top panel, clustering results for drug treated wt organoids are shown for comparison which were obtained by comparing NMF and UMAP clustering results.

## 2.2.11.2 Supplementary tables

Table S1: Assignments of all 206 genes to a particular cell type or biological pathway and full list of single-cell studies which were used to select our 206 gene signature set. Related to STAR Methods.

Gene	Signature_group	Gene	Signature_group
Vegfa	Angiogenesis	Scd2	Metabolism
Hif1a	Angiogenesis	Lars2	Metabolism
Birc3	Apoptosis	Tk1	Metabolism
Birc2	Apoptosis	Fasn	Metabolism
Xiap	Apoptosis	Gart	Metabolism
Bax	Apoptosis	Ppat	Metabolism
Bcl2l1	Apoptosis	Shmt2	Metabolism
Bcl2	Apoptosis	Shmt1	Metabolism
Id1	BMP	NFKBIA	nF kB
Id2	BMP	Nfkb2	nF kB
Id3	BMP	Nfkb1	nF kB
Bmpr1a	BMP	Rela	nF kB
Gdf15	BMP	Cxcl9	nF kB
Ccnb1	cell cycle	Tnf	NF kB
Cdkn3	cell cycle	Hes1	Notch
Cdc25c	cell cycle	Notch1	Notch
Cdc25a	cell cycle	Klf4	Notch
Cdkn1b	cell cycle	Atoh1	Notch
Ccnd1	cell cycle	Rbpj	Notch
Ccne1	cell cycle	Hes6	Notch
Chgb	EEC	Dll1	Notch
Cbfa2t3	EEC	Jag1	Notch
Rgs2	EEC	Notch2	Notch
Cdkn1a	EEC	Itln1	Paneth_cells
Chga	EEC	Ang4	Paneth_cells
Cck	EEC	Lyz1	Paneth_cells
Gip	EEC	Guca2b	Paneth_cells
Sst	EEC	Defa24	Paneth_cells
Tph1	EEC	Defa17	Paneth_cells
Rbbp4	EEC	Clps	Paneth_cells
Dach1	EEC	Pla2g2a	Paneth_cells
Tac1	EEC	Wnt3	Paneth_cells
Cpe	EEC	AY761184	Paneth_cells
Ghrl	EEC	Nupr1	Paneth_cells
Ptprn2	EEC	Habp2	Paneth_cells
Neurog3	EEC	Defa6	Paneth_cells
Fabp5	EEC	Hopx	qISCs
Sct	EEC	Pdlim2	qISCs
Neurod1	EEC	Bmi1	qISCs
Scgn	EEC	Lrig1	qISCs
Cdh1	EMT	Dmbt1	TA
Pkp1	EMT	Mki67	TA
Cd44	EMT	Ube2c	TA
Vim	EMT	Hmmr	TA
Reg1	Enterocytes	Gjb3	TA

Gene	Signature_group	Gene	Signature_group
Aldob	Enterocytes	Fgfbp1	TA
Ces2e	Enterocytes	Phlda2	TA
Prap1	Enterocytes	1810065E05Ri	TA
Alpi	Enterocytes	Ptma	TA
Gsta4	Enterocytes	Rbp7	TA
Gstm3	Enterocytes	Prox1	TA
Apoc3	Enterocytes	Prom1	TA
Plac8	Enterocytes	Etv4	TA
Fabp1	Enterocytes	Bhlhe40	TA
Sis	Enterocytes	Trim31	TA
Fabp2	Enterocytes	Pabpc1	TA
Apoa4	Enterocytes	Cox5a	TA
Anpep	Enterocytes	Ppm1g	TA
Fth1	Enterocytes	Acvr1	TGF-b
Mdh2	Enterocytes	Bambi	TGF-b
Tff3	Goblet_cells	Sox4	TGF-b
Muc2	Goblet_cells	Adh1	Tuft cells
Spink4	Goblet_cells	Krt8	Tuft cells
Fcgbp	Goblet_cells	Krt18	Tuft cells
Tmsb10	Goblet_cells	Aldh2	Tuft cells
Lgals2	Goblet_cells	Tmem176b	Tuft cells
Guca2a	Goblet_cells	Dclk1	Tuft cells
Agr2	Goblet_cells	Plcg2	Tuft cells
Txndc5	Goblet_cells	Pou2f3	Tuft cells
Sec61g	Goblet_cells	Alox5ap	Tuft cells
Pdia6	Goblet_cells	Trpm5	Tuft cells
Ccl6	Goblet_cells	Hck	Tuft cells
Ern2	Goblet_cells	Trp53	myc
Zg16	Goblet_cells	Myc	myc
Map2k1	MAPK/AKT	Mycbp	myc
Map3k1	MAPK/AKT	Ephb3	Wnt
Kras	MAPK/AKT	Ephb2	Wnt
Nras	MAPK/AKT	Axin2	Wnt
Rictor	MAPK/AKT	Btrc	Wnt
Egfr	MAPK/AKT	Porcn	Wnt
Pten	MAPK/AKT	Tcf7l2_iso1	Wnt
Akt1	MAPK/AKT	Tcf7l2_iso2	Wnt
Pdk1	MAPK/AKT	Cdx2	Wnt
Stat1	JAK/STAT	Hoxa5	Wnt
Ifngr2	JAK/STAT	Ctnnb1	Wnt
Socs3	JAK/STAT	Sox9	Wnt
Ihh	Hedgehog	Aldh1	Wnt
Ptch1	Hedgehog	Rara	Wnt
Areg	Hippo		
Birc5	Hippo		
Ppp2cb	Hippo		
Yap1	Hippo		
Aurka	Hippo		
Epcam	Houskeeping_genes		



Gene	Signature_group	Gene	Signature_group
<b>Tuba4a</b>	Houskeeping_genes		
<b>Hprt</b>	Houskeeping_genes		
<b>Polr2a</b>	Houskeeping_genes		
<b>Tbp</b>	Houskeeping_genes		
<b>Tubb2c</b>	Houskeeping_genes		
<b>gDNA_Cpe</b>	Genomic_DNA_conrols		
<b>gDNA_Lgr5</b>	Genomic_DNA_conrols		
<b>gDNA_Olfm4</b>	Genomic_DNA_conrols		
<b>gDNA_Dclk1</b>	Genomic_DNA_conrols		
<b>Clca4</b>	ISCs		
<b>Olfm4</b>	ISCs		
<b>Prelp</b>	ISCs		
<b>Slc12a2</b>	ISCs		
<b>Soat1</b>	ISCs		
<b>Lgr5</b>	ISCs		
<b>Ascl2</b>	ISCs		
<b>Cdk6</b>	ISCs		
<b>Aqp4</b>	ISCs		
<b>Cdca7</b>	ISCs		
<b>Rnf43</b>	ISCs		
<b>Rgmb</b>	ISCs		
<b>Rnf32</b>	ISCs		
<b>Tnfrsf19</b>	ISCs		
<b>Mbnl2</b>	ISCs		

**List of studies:**

PMID: 28285904. Mex3a Marks a Slowly Dividing Subpopulation of Lgr5+ Intestinal Stem Cells,  
 PMID: 27345837. De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome  
 PMID: 26287467. Single-cell messenger RNA sequencing reveals rare intestinal cell types,  
 PMID: 28334816. Dynamic Changes in Chromatin States During Specification and  
 PMID: 22692129. The Lgr5 intestinal stem cell signature: robust expression of proposed  
 PMID: 28059064. BMP restricts stemness of intestinal Lgr5+ stem cells by directly suppressing  
 PMID: 19269367. Transcription Factor Achaete Scute-Like 2 Controls Intestinal Stem Cell Fate.  
 PMID: 27939219. Induced Quiescence of Lgr5+ Stem Cells in Intestinal Organoids Enables

**Table S2: Cell type enrichments quantified for all drugs using mysort toolbox in R. Related to Figure 3.**

Sample	Enterocytes	EECs	Goblet	Tuft	Paneth	qISCs	ISCs	TA
Control	12.5000	12.5000	12.5000	12.5000	12.5000	12.5000	12.5000	12.5000
DC1	24.7424	11.8284	13.9311	13.6290	8.9070	11.2363	5.4694	10.2564
DC2	18.5090	19.2895	10.7123	15.5528	11.1535	11.2012	5.4455	8.1362
Epirubicin	26.8696	10.8622	13.1917	10.7509	8.4138	21.3308	4.5139	4.0671
Daunorubicin	32.3651	4.9853	9.6805	9.6528	7.9482	28.0549	3.8425	3.4705
Doxorubicin	32.8548	8.4548	10.9829	7.1564	11.9086	15.4135	4.6867	8.5424
Colchicine	20.1353	28.1459	9.4389	15.3179	8.8677	10.7426	2.7313	4.6204
Griseofulvin	10.8904	10.2568	13.7159	12.3876	8.1402	22.5472	9.5049	12.5570
Floxuridine	12.7744	15.7713	11.1067	12.4309	14.2667	17.9751	6.4320	9.2429
Pyriminium	18.8125	9.6530	12.4624	10.4699	16.6015	16.5545	7.3455	8.1006
Antimycin	19.4368	8.4941	13.7437	9.2697	9.3272	17.6077	10.9427	11.1782
Minoxidil	23.6193	10.3439	11.3008	14.2642	10.2958	14.1329	5.6293	10.4138
Mercaptopurine	12.3998	14.5268	11.6648	15.2781	20.3381	11.8429	7.3443	6.6052
Irinotecan	15.4020	11.3375	13.2379	11.6601	19.3379	15.6658	5.5713	7.7874
Gemcitabine	16.9499	17.6786	11.4869	14.7040	15.8530	10.8825	6.2595	6.1857
Nalbuphine	18.6902	13.6594	12.0865	12.4153	21.6397	10.0834	5.3519	6.0736
Erlotinib	16.5802	18.6231	7.7873	14.8757	16.2120	12.2524	9.1484	4.5208
Azaguanine	7.9272	15.3054	11.8307	13.0589	21.4605	11.7482	11.2691	7.4000
Papaverine	9.8874	5.9999	4.9647	17.4934	8.0875	34.4950	11.6576	7.4145
Carbamazepine	12.2603	13.9023	6.1388	14.0612	7.8024	14.3646	14.2437	17.2266
Calcipotriene	9.3368	9.0625	13.6042	10.5089	14.7202	27.1375	7.2826	8.3473
Toremifene	12.9030	10.6581	10.6803	24.4672	15.8123	9.3307	6.0692	10.0792
Fluphenazine	9.2928	19.1302	10.1193	12.7703	14.4328	16.2335	6.7848	11.2364
Silodosin	8.4489	10.1782	17.6725	13.5056	10.4123	14.5831	8.8494	16.3499
Fluconazole	11.7772	12.2310	15.5787	9.6634	14.4978	13.8706	9.2732	13.1081
Piperacetazine	9.0609	13.5910	12.4459	13.0440	12.8882	13.7047	14.2731	10.9923
Perphenazine	7.3965	19.6079	15.1329	10.6145	16.4714	14.4641	7.4425	8.8702
Trifluoperazine	7.0841	16.3726	11.1027	12.4591	13.0104	13.6082	13.4757	12.8872
Prochlorperazine	8.7867	13.9618	13.2937	10.9870	13.4506	14.0688	12.9188	12.5327
Tamoxifen	10.1619	13.6915	11.7950	11.8677	12.5322	16.1398	11.1096	12.7023
Clomiphene	6.2399	17.9970	13.8866	12.2982	10.6011	17.7775	10.0098	11.1900
Opipramol	7.6810	14.1555	11.5322	12.7357	22.6423	12.0678	8.7248	10.4607
Ifenprodil	9.9044	13.8659	10.4061	15.7320	14.4882	9.7877	11.6836	14.1321
Simvastatin	6.5709	24.6921	5.7758	18.0383	16.7123	13.1698	6.7410	8.2997
Fluvastatin	10.1117	39.2654	5.4349	17.7901	11.6572	7.1751	3.8777	4.6879
Lovastatin	6.2725	33.7904	10.0185	16.3474	10.2511	12.4662	5.0830	5.7710
Atorvastatin	6.6973	27.8736	12.0881	15.4824	16.9113	11.4342	4.8951	4.6180
Prednicarbate	11.2956	11.9763	13.5116	10.9711	15.2673	17.3074	9.9127	9.7580
Hydrocortisone	13.0847	12.0761	11.4320	10.6718	13.7210	19.2746	10.5680	9.1718
Budesonide	11.7773	11.6016	12.3163	11.0918	11.8339	18.2314	13.4140	9.7336
Dexamethasone	11.1076	12.1262	12.1777	11.3186	11.8788	19.7649	11.7431	9.8830
Triamcinolone	13.0879	13.5910	12.7320	12.7982	11.1985	16.2166	10.3762	9.9996
Fludrocortisone	8.8440	11.7462	13.7232	12.6150	12.6154	19.4849	10.8469	10.1244



Sample	Enterocytes	EECs	Goblet	Tuft	Paneth	qISCs	ISCs	TA
Corticosterone	9.9090	11.0674	12.0068	11.9691	11.7834	18.5988	12.4491	12.2164
Xylazine	11.7421	13.7039	9.8174	14.0256	7.6705	17.6879	11.8994	13.4532
Thalidomide	10.3653	14.3452	15.5382	11.2773	9.8289	19.1874	7.6432	11.8144
Acetohexamide	16.4991	15.7411	8.5785	14.1345	8.5820	14.2333	9.8208	12.4107
Tobramycin	12.1383	13.7366	8.6621	15.8440	9.9256	13.4666	12.6704	13.5563
Tulobuterol	12.7800	12.3265	8.5944	14.5140	6.5617	20.5283	10.5947	14.1004
Etofenamate	17.7013	11.0922	10.4839	11.5248	9.6823	13.4431	12.9945	13.0780
Tolbutamide	8.8719	11.2254	12.8682	13.0426	7.9299	20.5437	10.6627	14.8557
Itraconazole	14.6726	11.4317	14.2322	11.4521	9.8887	15.9650	10.3258	12.0318
Albendazole	16.7872	12.3320	12.1133	12.4460	7.7239	15.9635	10.4414	12.1927
Mebendazole	15.3093	9.9343	10.2286	16.9801	12.3314	14.2097	10.5253	10.4812
Butoconazole	11.8126	13.3698	12.3780	13.1692	12.5056	15.6235	10.0351	11.1064
Econazole	8.8840	15.1885	11.5889	12.6798	12.1016	13.9864	12.5368	13.0340
Hydroxyzine	10.5065	15.8577	13.5619	14.2340	10.5465	13.7553	9.2455	12.2925
Thiethylperazine	10.4975	15.6090	13.1211	13.2911	12.9836	15.1976	8.4338	10.8663
Tripelennamine	13.3963	11.6436	12.2516	12.6116	15.0113	14.9340	9.8963	10.2553
Nefazodone	12.1743	11.0682	13.2808	12.2443	10.2838	15.6188	11.8945	13.4353
Sarafloxacin	13.3997	11.9553	12.7623	11.9023	10.8127	12.7484	13.0807	13.3386

**Table S3: List of drugs inducing differentiation phenotypes in WT organoids and targeting AKP mutant organoids. Related to Figure 6.**

<b>Class</b>	<b>WT_differentiation_phenotype</b>	<b>AKP_targeting</b>
Ifenprodil	yes	yes
Opipramol	yes	yes
Prochlorperazine	yes	yes
Atorvastatin	yes	yes
Perphenazine	yes	yes
Pyrvinium	yes	yes
Lovastatin	yes	yes
Nalbuphine	yes	yes
Hydroxyzine	yes	yes
Clomiphene	yes	yes
Thiethylperazine	yes	yes
Toremifene	yes	yes
Trifluoperazine	yes	yes
Antimycin	yes	yes
Nocodazole	yes	yes
Erlotinib	yes	no
Tripelennamine	yes	no
Calcipotriene	yes	no
Thalidomide	yes	no
Silodosin	yes	no
Triamcinolone	yes	no
Xylazine	yes	no
Sarafloxacin	yes	no
Itraconazole	yes	no
Fluconazole	yes	no
Acetohexamide	yes	no
Nefazodone	yes	no

**Table S4: Cell death induction as measured by FACS in human cell lines treated by indicated drugs. Related to Figure 7.** Percentage of annexinV positive or DAPI positive cells is displayed as an average of two independent experiments (except DLD-1 cells where only one replicate was used). Combined scores were calculated based on each replicate of cell treatments: score 1 was assigned to a drug if annexin or DAPI percentage was at least 110% of the average control value (score 14 is the maximal possible score). Drugs with combined scores of more than 1 were considered to target CRC cell lines. Hits are highlighted in grey colour.

Drug_name	combined score	ANX-V LS174T	DAPI+ LS174T	ANX-V DLD1	DAPI+ DLD1	ANX-V HCT116	DAPI+ HCT116	ANX-V NCI-H508	DAPI+ NCI-H508	LS174T	DLD1	HCT116	NCI-H508
Antimycin	12	56.1	18.8	60.1	60.9	35.3	31.0	54.4	45.8	targeted	targeted	targeted	targeted
Pyruvium	12	90.7	45.9	81.3	75.6	79.3	60.7	54.8	47.1	targeted	targeted	targeted	targeted
Lovastatin	12	50.4	15.3	36.5	41.4	46.7	37.6	64.3	51.5	targeted	targeted	targeted	targeted
Itraconazole	12	88.7	25.8	63.4	63.8	74.0	62.5	88.0	79.8	targeted	targeted	targeted	targeted
Atorvastatin	11	44.3	14.0	35.1	36.8	30.1	25.1	46.3	30.1	targeted	targeted	targeted	targeted
Papaverine	10	42.9	17.3	59.2	63.3	40.3	37.4	93.3	85.8	targeted	targeted	targeted	targeted
Carbamazepine	6	39.5	9.7	32.9	47.9	18.6	17.6	29.8	24.3		targeted		targeted
Xylazine	6	46.2	15.1	43.5	57.9	16.1	16.3	19.6	18.2	targeted	targeted		targeted
Erlotinib	5	36.5	11.6	39.8	46.0	10.6	45.0	20.8	15.2		targeted	targeted	targeted
Tamoxifen	4	44.1	12.1	26.8	31.4	20.1	20.0	18.7	14.7			targeted	targeted
Opipramol	4	40.6	9.6	34.8	47.7	16.3	16.3	11.5	10.8	targeted	targeted	targeted	
Clemastine	4	35.8	7.0	33.0	32.3	19.9	16.2	22.6	13.8	targeted			targeted
Isoconazole	3	41.8	9.0	34.1	35.8	19.6	16.1	19.3	11.6		targeted		targeted
Toremifene	3	36.7	10.1	33.5	39.0	14.1	13.7	18.8	11.1		targeted		targeted
Nalbuphine	3	37.2	10.5	38.7	41.3	15.5	15.0	15.5	13.9		targeted		targeted
Clomiphene	2	37.6	8.3	29.3	28.6	18.0	15.8	23.7	15.1				targeted
Pinaverium	2	35.0	9.3	26.5	26.7	13.4	11.7	21.2	14.7				targeted
Hydroxyzine	1	41.5	12.5	23.0	22.2	16.2	13.4	9.9	6.0	targeted			
Minoxidil	1	36.5	8.8	22.2	31.8	20.9	20.5	13.7	10.5			targeted	
Benzoxiquine	1	49.4	12.0	23.4	21.7	15.5	13.6	11.0	7.5	targeted			
Econazole	1	36.0	7.1	25.3	33.3	20.3	20.0	14.8	11.8		targeted		
Calcipotriene	1	34.5	6.6	26.4	27.5	17.1	15.3	14.0	8.6	targeted			
Tulobuterol	1	35.8	10.1	20.1	19.7	21.3	18.0	7.1	4.4			targeted	
Tegaserod	1	35.8	9.0	22.0	23.4	13.0	12.4	15.9	10.7				targeted
Tobramycin	1	38.3	10.5	18.6	25.1	18.0	19.5	12.8	11.1		targeted		
Fluconazole	1	41.0	8.5	20.5	21.0	17.5	15.5	10.7	8.0	targeted			
Thalidomide	1	36.7	8.2	23.5	26.0	21.2	18.3	10.1	6.6			targeted	
Tripelennamine	1	30.8	7.9	24.9	26.2	19.0	16.3	20.8	9.0				targeted
Trifluoperazine		36.3	10.7	25.6	26.1	13.2	10.9	10.6	7.4				
Prochlorperazine		37.4	8.3	23.1	27.2	18.7	18.3	10.8	8.0				
Thiethylperazine		39.2	8.5	23.4	25.8	18.9	19.0	13.6	10.1				
Perphenazine		41.0	10.6	19.1	21.7	18.0	17.4	9.1	7.3				
Mizolastine		40.3	10.6	26.7	29.1	19.8	18.0	10.6	7.3				
Ifenprodil		38.5	11.3	22.6	18.7	13.3	11.3	11.5	8.6				
Butoconazole		37.7	10.8	19.0	19.8	14.8	13.3	15.0	11.7				
Piperacetazine		39.1	10.2	18.9	21.3	17.2	16.2	11.8	10.0				
Hydrocortisone		38.0	9.6	28.4	29.5	12.9	13.6	11.4	8.9				
Etofenamate		32.7	8.4	26.0	26.1	15.1	14.1	11.9	8.7				
Silodosin		38.4	9.7	19.6	20.2	11.2	13.0	13.4	9.8				
Sarafloxacin		41.0	9.3	22.8	22.8	18.5	16.1	11.4	7.2				
Control_rep1		31.5	8.1	30.8	31.9	16.5	17.4	13.6	11.4				
Control_rep2		37.7	12.0	18.4	21.9	17.7	20.1	12.4	10.4				
Control_rep3		38.6	10.1	16.5	17.9	15.1	20.3	15.3	11.7				
Control_rep4		40.1	10.1	16.4	16.2	22.1	17.1	13.7	10.4				
Control_rep5		37.5	12.6	17.7	22.5	15.5	14.0	12.8	10.5				
Control_rep6		36.8	8.5	13.7	14.8	18.7	16.2	10.0	7.1				
Control_average		37.0	10.2	18.9	20.9	17.6	17.5	13.0	10.3				

**Table S5: Comparison of different high-throughput sequencing methods. Related to STAR Methods.**

	Avg # of genes detected	Avg # of genes >100 counts	# of cells needed per well	format	RNA purification needed
PLATE-seq	10200	800 (8%)	10000	96	yes
DRUG-seq	11000	1000 (10%)	2500	384, 1536	no
TAC-seq	57	57 (100%)	10000	NA	yes
RASL-seq	56 ( 261 probes)	90 (35%)	1000	384	yes / no
TORNADO-seq	205	100 (50%)	1000	96	yes
	processing time	Sequencing depth	library construction cost per sample *	Sequencing cost per sample **	Total cost per sample
PLATE-seq	8-10 hrs	2 mio reads	N/A	\$10	\$15
DRUG-seq	10 hrs	2 mio reads	\$0.9	\$10	\$11
TAC-seq	NA	500k	\$6	\$2.5	~ \$8.5
RASL-seq	6 hrs	10-500k reads	\$4 ****	\$0.4	~ \$5
TORNADO-seq	6 hrs	50k reads	\$4 ***	\$0.2	~ \$5
	<b>Advantages</b>		<b>Disadvantages</b>		
PLATE-seq	Detection of more than 10 000 genes		Need of special oligo-dT plates, highest price per sample		
DRUG-seq	Detection of more than 10 000 genes		Requires custom liquid dispensing equipment, high price for sequencing, compatibility with organoid-based screens		
TAC-seq	Medium price, UMI counting improvement compare to RASL-seq		Optimization and troubleshooting of ligation-based problems, random ligation at low target level, expensive primers with 5'Phos		
RASL-seq	Lowest price for HTS		Optimization and troubleshooting of ligation-based problems, random ligation at low target level, expensive primers with 5'Phos, PCR amplification bias		
TORNADO-seq	Lowest price for HTS, UMI counts and amplicon insertion alignment allow to remove bias from unspecific amplification		Optimization of the primers		

\*not very comparable due to 96 and 384 well plate formats. Technically it appears impossible to perform organoid screens based on RNA-seq in a 384 well plate.

\*\*Price per 1 mio reads = \$5

\*\*\* \$1 cost of culture (matrigel, B27, N2 media, etc) and primers included

// Primer and probe cost is counted for 10 000 samples

**Table S7: PCR conditions. Related to STAR Methods.** For RT reaction Tth-DNA polymerase was used. For PCR GoTaq Hot Start enzyme was used.

1st stage PCR:

PCR touch-down method:

94°C - 1 min

2x:

- 94°C - 1 min

- 65°C - 5 min

- 72°C - 1 min

2x:

- 94°C - 1 min

- 64°C - 4 min

- 72°C - 1 min

2x:

- 94°C - 1 min

- 63°C - 3 min

- 72°C - 1 min

2x:

- 94°C - 1 min

- 62°C - 2 min

- 72°C - 1 min

2x:

- 94°C - 1 min

- 61°C - 1 min

- 72°C - 1 min

10x:

- 94°C - 20 sec

- 60°C - 15 sec

- 72°C - 10 sec

2nd stage PCR:

- 94°C - 20 sec

- 68°C - 15 sec

- 72°C - 10 sec

repeat 10-20 cycles

2.3. The second part of the thesis results is presented as a published paper:

## **“Genomic instability profiles at the single cell level in mouse colorectal cancers of defined genotypes?”**

Cancers, 2021, 13(6), 1267; <https://doi.org/10.3390/cancers13061267>

### **Authors:**

Vasilis S. Dionellis<sup>1</sup>, **Maxim Norkin**<sup>2</sup>, Angeliki Karamichali<sup>1</sup>, Giacomo G. Rossetti<sup>1</sup>, Joerg Huelsken<sup>2,\*</sup>, Paloma Ordonez-Moran<sup>3,\*</sup> and Thanos D. Halazonetis<sup>1,\*</sup>

### **Authors' affiliation:**

<sup>1</sup> Department of Molecular Biology, University of Geneva, 1211 Geneva, Switzerland.

<sup>2</sup> Cancer Stem Cell Laboratory, Swiss Institute of Technology Lausanne (EPFL), ISREC, 1015 Lausanne, Switzerland.

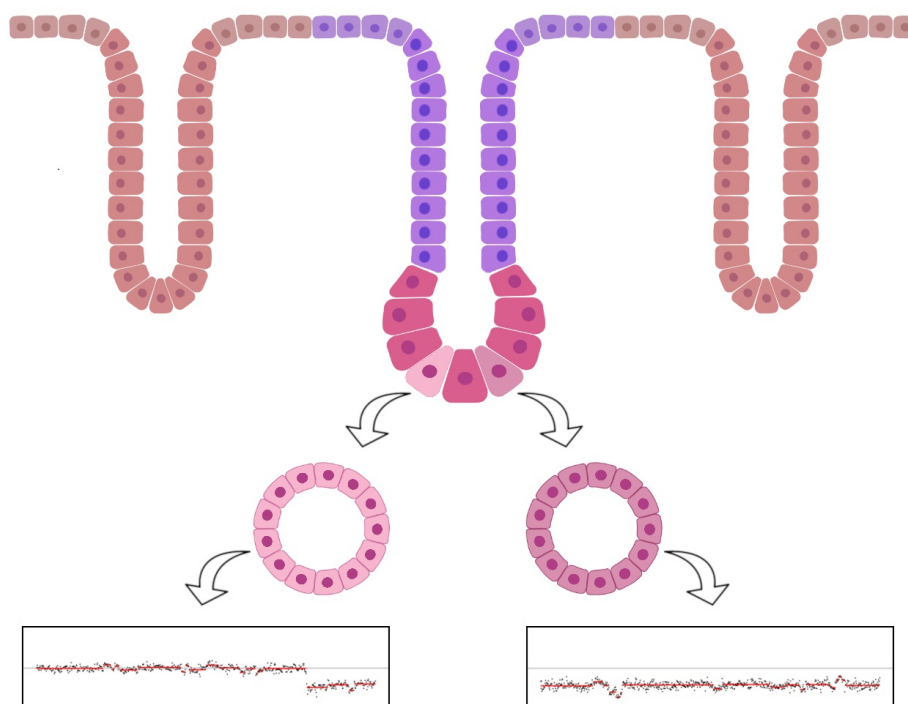
<sup>3</sup> Division of Cancer & Stem Cells, School of Medicine, Centre for Cancer Sciences, Biodiscovery Institute, University of Nottingham, NG7 2RD, Nottingham, United Kingdom.

\* Correspondence: Thanos.Halazonetis@unige.ch (T.D.H.); Palma.OrdonezMoran@nottingham.ac.uk (P.O.-M.); Joerg.Huelsken@epfl.ch (J.H.)

### **Author contributions:**

T.D.H., P.O.M. and J.H. conceived the project and designed the experiments, M.N., P.O.M., A.K. and G.R. performed experiments, V.D. established the bioinformatics pipelines and performed computational analysis with contributions from T.D.H. V.D., M.N., P.O.M., J.H. and T.D.H. wrote the manuscript.

### 2.3.1 Graphical abstract



### 2.3.2 Abstract

The genomes of many human CRCs have been sequenced, revealing a large number of genetic alterations; however, the molecular mechanisms underlying the accumulation of these alterations are still being debated. In this study, we examined colorectal tumours that developed in mice with  $Apc^{lox/lox}$ ,  $Kras^{Lsl-G12D}$  and  $Tp53^{lox/lox}$  targetable alleles. Organoids were derived from single cells and the spectrum of mutations was determined by exome sequencing. The number of single nucleotide substitutions (SNSs) correlated with the age of the tumour but was unaffected by the number of targeted cancer-driver genes. Thus, tumours that expressed mutant  $Apc$ ,  $Kras$  and  $Tp53$  alleles had as many SNSs as tumours that expressed only mutant  $Apc$ . In contrast, the presence of large-scale (>10 Mb) copy number alterations (CNAs) correlated strongly with  $Tp53$  inactivation. Comparison of the SNSs and CNAs present in organoids derived from the same tumour, revealed intratumoural heterogeneity, consistent with genomic lesions accumulating at significantly higher rates in tumour cells, than in normal cells. The rate of acquisition of SNSs increases from the early stages of cancer development, whereas large-scale CNAs accumulate after  $Tp53$  inactivation. Thus, a significant fraction of the genomic instability present in cancer cells cannot be explained by aging processes occurring in normal cells before oncogenic transformation.

### 2.3.3 Introduction

Colon is an excellent organ to study the accumulation of genomic alterations during cancer development, because precancerous and cancerous lesions can be easily harvested [1–6]. The initial event in most colorectal carcinomas (CRCs) is biallelic inactivation of the tumour-suppressor gene Adenomatous polyposis coli (APC) [2,7–9]. Further CRC development involves the accumulation of additional mutations that may be subclone-specific [6,10–15]. Activating mutations targeting KRAS are acquired in up to 40-50% of sporadic CRCs and are associated with dysplasia [2,8,16]. Up to 50–60% of human CRCs acquire inactivating mutations in the TP53 tumour-suppressor gene, an event associated with progression of dysplastic lesions to carcinoma. p53, the protein product of TP53, induces cell cycle arrest, senescence or apoptosis in response to DNA damage. Thus, its inactivation allows cancer cells to survive and proliferate, despite the presence of oncogene-induced DNA damage [17].

The first mouse model generated to study colon cancer disease was a mouse that harbours only Apc deletion in one allele (Apc<sup>MIN/+</sup>) (Moser et al., 1990). They have been extensively in research however, these mice do not recapitulate patient's CRC because first they develop tumours mainly in the small intestine and second, these are benign adenomas (Moser et al., 1990, Shibata et al., 1997). Many efforts have been done to generate a better model of CRC. Later on, Apc:Kras mutant mouse model was generated in an attempt to replicate the genetics underlying human CRC. These mice show higher tumour multiplicity than Apc deleted mice and what is more important these colonic tumours present an invasive phenotype in small intestine as in colon (Sansom et al., 2006, Janssen et al., 2006, Hung et al., 2010). This aggressive intestinal phenotype that increases intestinal tumour formation and progression was explained by synergistic activation of Wnt signalling (Janssen et al., 2006). As TP53 inactivating mutations occur at a much higher frequency in advanced human CRC (50–60%) than in adenomas (<10%), this mouse model was bred with Tp53 knock-out. This triple combination led to aggressive tumors that developed into carcinomas in the ceacum and colon (Martin et al., 2013, Tetteh et al., 2016). Indeed, intrasplenic injections or orthotopic transplantations of AKP cell lines into immunodeficient mice, these cells were able to invade, metastasize and colonize to the liver (Martin et al., 2013; Roper et al., 2017).

One of the most important hallmarks of cancer, including CRC, is genomic instability, a feature that facilitates cancer progression [18] and resistance to therapy [11,19,20].



Genomic instability can lead to the accumulation of numerous genomic alterations, including single nucleotide substitutions (SNSs), small insertions and deletions (indels), copy number alterations (CNAs) and chromosomal rearrangements. It is well established that CNAs and chromosomal rearrangements accumulate at higher rates in cancer cells, than in normal cells. However, it is less clear whether the rate of acquisition of SNSs increases after cell transformation. The early consensus in the field has been that the high number of SNSs in most human cancers simply reflect the high number of point mutations present in normal cells due to aging; since tumours are of monoclonal origin, these mutations become evident when tumour DNA is sequenced [21]. An alternative view is that SNSs accumulate at higher rates in cancer cells. Our sequencing study of human colon adenomas supported this latter view, since it revealed a higher number of SNSs in adenomas with severe dysplasia, compared to adenomas with mild dysplasia, despite similar patient age distribution [22]. One may also consider the possibility that certain types of mutations accumulate at higher rates in cancer cells, whereas other types of mutations accumulate at equal rates in normal and cancer cells due to, for example, aging. Along these lines, it is worth noting that the large-scale sequencing studies of human cancers have revealed distinct types of SNSs that are referred to as mutational signatures [23–26]. Similarly, various bulk tissue sequencing studies of genetically engineered mouse models (GEMMs), that recapitulate aspects of human cancers, have been conducted and improved our understanding of tumour initiation and progression [27–33].

The prevailing signature in human cancers is signature 1, a signature that is defined by a high number of C-to-T transitions in the context of CpG sites [25,34]. These mutations arise from failure to properly repair a methylated cytosine, after it has been deaminated by hydrolysis [35]. It has been proposed that signature 1 mutations accumulate with equal rates in normal and cancer cells [36–38]. However, the majority of mutations in colon cancer conform to signature 1 and our exome sequencing study, cited above, revealed a higher number of signature 1 SNSs in adenomas with severe dysplasia, compared to mild dysplasia. A higher mutation rate for signature 1 SNSs can be rationalized on the basis that cancer cells have DNA replication stress, which leads to the formation of single-stranded DNA [17]. The rate of cytosine deamination is hundred times higher in single-stranded, than double-stranded DNA [35], and deamination of a methylated cytosine to thymine in single-stranded DNA cannot be detected by the repair

machinery, because it does not lead to base pair mismatches and because thymine is a naturally occurring base in DNA.

Understandably, interpretation of cancer sequencing data is complicated by the presence of intratumoural heterogeneity [21,39]. Mutations that are present in a subset of cancer cells may have very low overall allele frequencies and not be counted. To address this problem, one can prepare single cell-derived tumour organoids for sequencing. In a previous study, we sequenced the exomes of organoids derived from normal or precancerous single cells isolated from the intestines of *Apc*<sup>Min/+</sup> mice and observed eleven times more mutations in the organoids derived from the precancerous cells [40]. Another study also examined single cell-derived organoids; in this latter study organoids were prepared from cancer and normal cells obtained from three CRC human patients. Comparison of the number of mutations in the tumour-derived organoids, as compared to the normal tissue-derived organoids revealed a modest increase in the number of SNSs conforming to signature 1 and a more significant increase in the number of SNSs conforming to signature 17 [41].

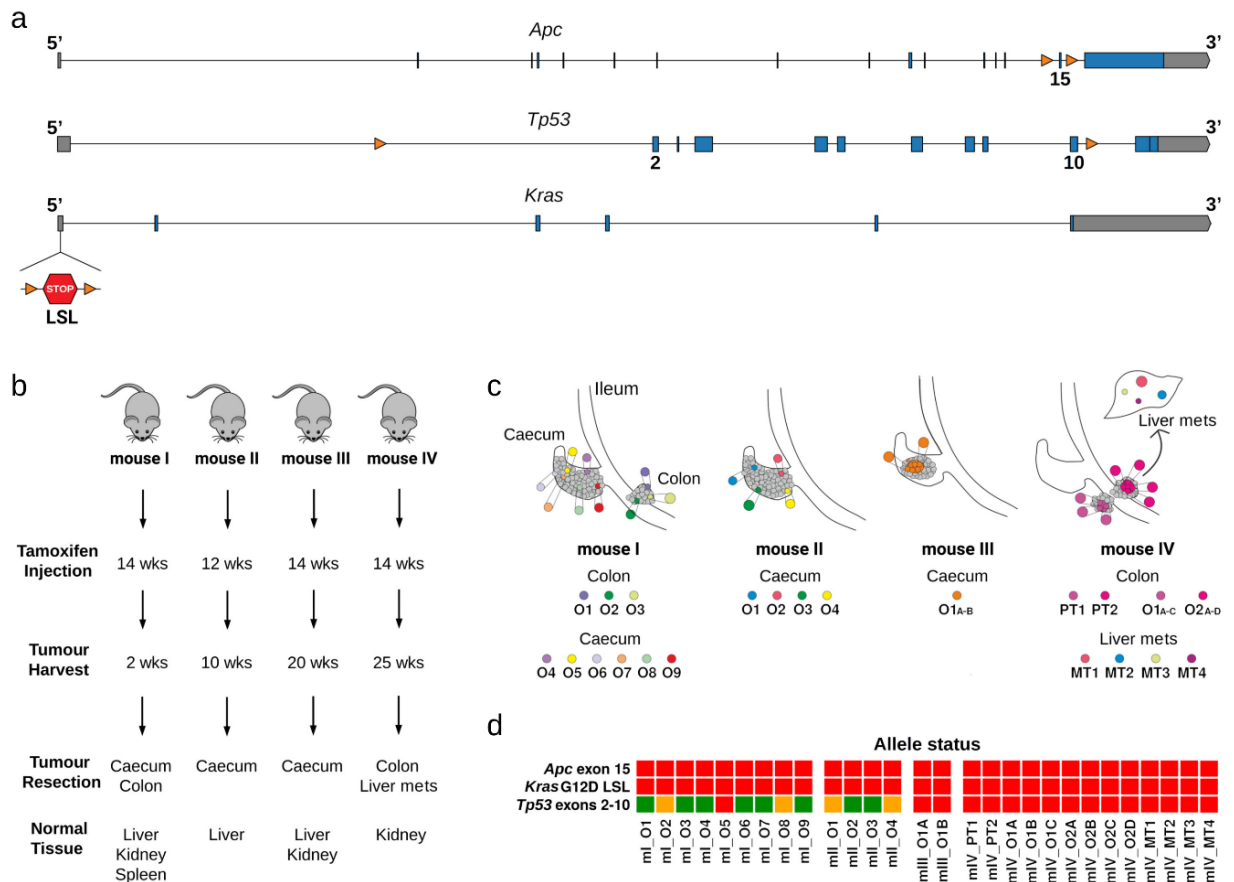
To gain a better understanding of mutation rates in cancer cells, we turned to a mouse model of CRC that is driven by several cancer driver genes. Specifically, we examined mice that had three targetable alleles: *Apc*<sup>lox/lox</sup>, *Kras*<sup>Lsl-G12D</sup> and *Tp53*<sup>lox/lox</sup>, corresponding to the most frequently mutated genes in human colon cancer [42,43]. The mice also harboured a transgene that was expressed specifically in the colon and which encoded a tamoxifen-inducible recombinase (*Cdx2*<sup>CreERT2</sup>), allowing the three cancer-driver genes to be targeted in an inducible manner. As before, we prepared single cell-derived organoids from CRCs that developed in these mice and sequenced their exomes. Our results provide a better understanding of the role of oncogenes and tumour suppressor genes on the accumulation of SNSs and CNAs in cancer cells.

## 2.3.4 Results

### 2.3.4.1 Clonal organoid cultures derived from single tumoral cells

We used a mouse model of colonic tumorigenesis (*AKP-Cdx2*<sup>CreERT2</sup> mice), in which three endogenous cancer-driver genes were modified, so they could be targeted by a tamoxifen-inducible Cre gene. The targeted cancer-driver genes were *Apc*, whose exon 15 was flanked by loxP sites (*Apc*<sup>lox/lox</sup>), *Kras*, which contained a G12D mutation and a transcription termination site flanked by loxP sites upstream of the first coding exon

(*Kras*<sup>Lsl-G12D/+</sup>), and *Tp53*, whose exons 2-10 were flanked by loxP sites (*Tp53*<sup>lox/lox</sup>) (Fig. 1a; Fig. S1a). The *Cre* gene was under control of the *Cdx2* promoter to confer specific expression in the large intestine.



**Figure 1.** CRC mouse model, resected samples and genotype of organoids. (a) Structure of conditional alleles of *Apc*, *Trp53*, and *LSL-KrasG12D*. Blue boxes and orange triangles indicate exons and loxP sites, respectively. (b) Four 3-4 months old mice were injected with tamoxifen to initiate tumor development. Resections from caecum and colon tumours were obtained for subsequent isolation of intestinal single cells. Normal tissue was also resected and used as control sample during sequencing analysis. Mouse IV developed metastatic events in liver tissue and metastatic tumor samples were also resected and sequenced. (c) Multiregion sampling of each mouse is illustrated by coloured labels. PT: primary tumours, MT: metastasis, O: tumour organoids. (d) Recombination events in organoids validated by PCR-based genotyping and by read depth analysis of sequencing data. Red, orange and green colour indicate homozygous, heterozygous and no recombination for *Apc* and *Trp53*, respectively. In the case of *LSL-KrasG12D*, red colour represents recombination of the conditional allele.

We administered tamoxifen to four mice (Fig. 1b). The first mouse (ml) received a dose of 30 mg/kg; this mouse developed multiple tumours and was sacrificed 2 weeks after tamoxifen administration. The other three mice received a tamoxifen dose of 3 mg/kg, which led to low levels of recombination and few tumours developing. These mice were sacrificed 10 to 25 weeks after tamoxifen injection. Tissue biopsies with macroscopically visible tumours were used to prepare suspensions of single cells, which

were then aliquoted into wells of 96-well plates for expansion as 3D-organoid cultures (Fig. 1c; Fig. S1b). We only propagated cultures from wells, in which initially only a single organoid grew, and we considered that these cultures were derived from a single cell. The organoids were spheroid-shaped and lacked the crypt-like projections that are typical of intestinal organoids derived from non-transformed cells (Fig. S1b).

A diagnostic PCR, supported by analysis of the number of exome sequencing reads, was used to determine whether the *Apc*, *Kras*<sup>G12D</sup> and *Tp53* genes had undergone recombination by Cre (Fig. 1d). As expected, all tumour-derived organoids retained the Cre gene (Fig. S2a) and had suffered biallelic deletions of exon 15 of *Apc* (Figs S2b, S3). The G12D *Kras* allele had also recombined in all organoids (Fig. 1d; Figs S2c, S4, S5). Finally, in most organoids derived from mice I and II, the *Tp53* gene had not recombined, whereas in all organoids from mice III and IV both *Tp53* alleles had recombined (Fig. 1d; Figs S2d, S6).

#### **2.3.4.2 Single nucleotide substitutions – Prevalence**

Single nucleotide substitutions (SNSs) were identified by comparing the exome sequencing data of the single cell-derived tumour organoids to the corresponding data of bulk normal tissue from the same mice (Fig. 1b). We used the liver as the reference normal tissue, except for mouse IV, for which we used the kidney, because its liver had metastatic lesions (Fig. 1c). For mouse IV, we also obtained exome sequencing data of primary and metastatic tumour tissue biopsies (Fig. 1c).

One concern when analysing cancer sequencing data, in which few SNSs are expected per sample, is that a significant fraction of the identified variants may be germline polymorphisms and not somatic SNSs. This type of error will occur, when, due to low sequencing coverage, germline variants are identified in the sequences of the tumour organoid, but not in the sequences of the reference normal tissue. Given the very large number of non-annotated germline variants in the mouse, such errors may be frequent. To minimize them, we performed exome sequencing, which allowed us to have high sequencing coverage. In addition, we restricted the analysis to the protein coding regions, because these are well-annotated.

The exome sequencing data of 22 tumour organoids prepared from the four mice revealed, in total, 206 somatic SNSs mapping to gene coding regions or splice sites (Data S1). Of the 206 SNSs, 149 were missense, 3 nonsense, 44 synonymous and 10 were targeting splice-sites (Fig. 2a). We classified all the SNSs as passenger mutations, since

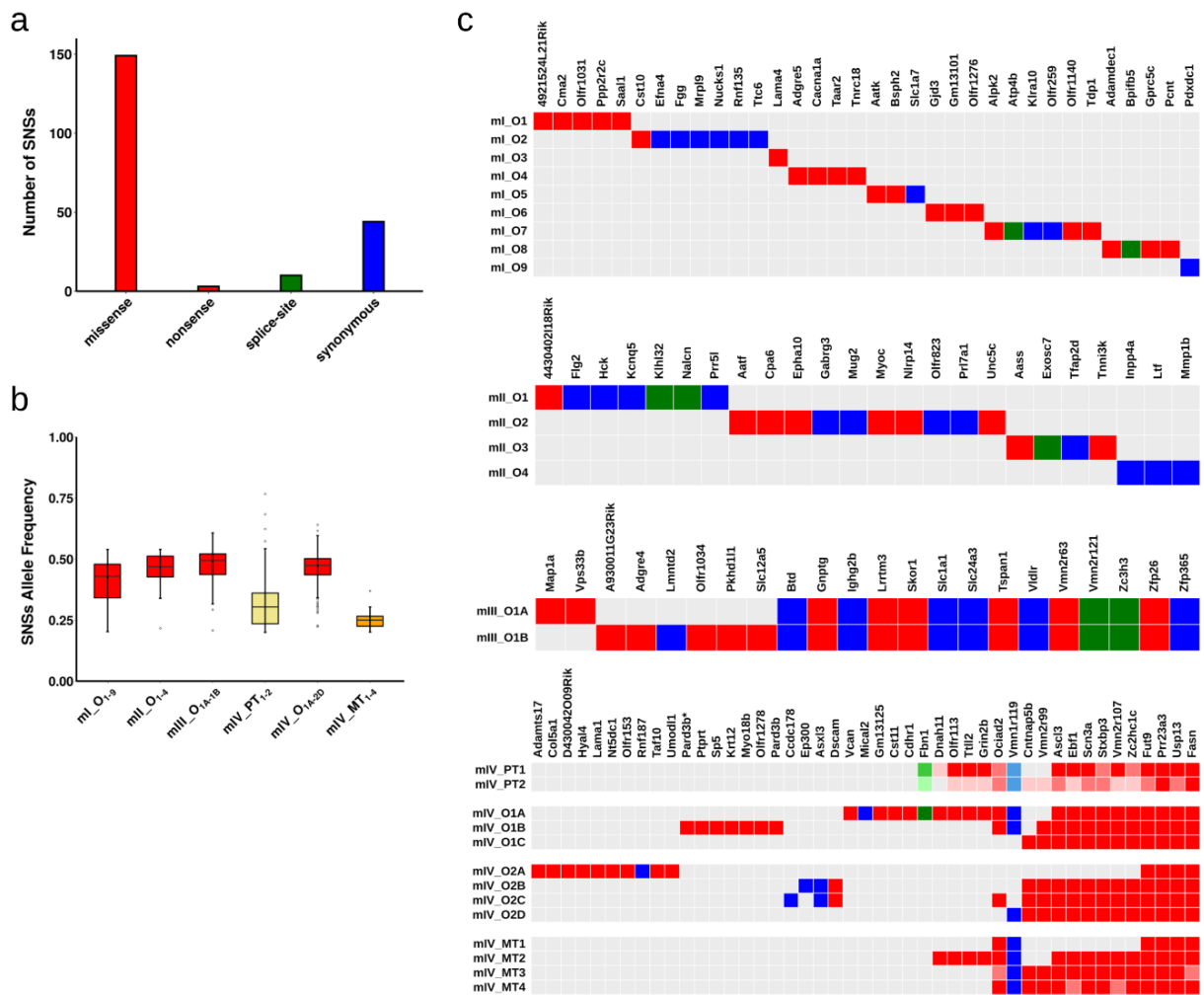
none of them targeted known oncogenes or tumour suppressor genes (as defined by ICGC/TCGA).

The average allele frequency of the SNSs in the organoids was approximately 50% (Fig. 2b). Considering that the organoids were derived from a single cell, these allele frequencies are consistent with heterozygous mutations acquired *in vivo*. If the SNSs had been acquired during tissue culture, they would not be present in all cells and would have lower allele frequencies. By comparison, the allele frequencies of the SNSs in bulk tumour tissue were lower than 50% (Fig. 2b), reflecting intratumoural heterogeneity (see below) and the presence of normal cells in bulk tumour tissue.

We next studied the spectrum of the identified SNSs. For mice I and II, tumour biopsies that were a few mm apart from each other were used to prepare organoid cultures and one organoid culture was sequenced per tissue biopsy (Fig. 1c). In these mice, no SNSs were shared between the different organoids indicating that independent transformation events took place in the different biopsies (Fig. 2c). Indeed, mice I and II showed signs of tumour development quite early after tamoxifen injection, consistent with the development of many tumours in parallel.

For mouse III, a single tumour tissue biopsy was used to obtain organoid cultures, two of which were subjected to exome sequencing (Fig. 1c). Fourteen SNSs were shared between the two organoids and eight SNSs were private (two and six, respectively, in the two organoids) (Fig. 2c). At the very minimum, the private SNSs must have been acquired after tumour development was initiated.

For mouse IV, two tumour tissue biopsies were harvested; three and four single cell-derived organoid cultures were then sequenced from the two biopsies, respectively (Fig. 1c). Interestingly, all organoids shared four SNSs, whereas six of the seven organoids shared an additional six SNSs (Fig. 2c). Thus, the organoids derived from the two tumour biopsies were related to each other. Given that the average frequency for all mutant alleles was about 50% (Fig. 2b), we infer that the SNSs were acquired *in vivo*. It is possible that the four shared SNSs might have been acquired prior to neoplastic transformation. In contrast, the private SNSs (29 in total; defined as SNSs found in only one of the seven organoids) and the semi-private SNSs (12 in total; defined as SNSs found in more than one, but not in all organoids) must have been acquired after neoplastic transformation (Fig. 2c).



**Figure 2.** Types and allele frequencies of identified point mutations. (a) Number of SNSs within the protein coding regions that were identified in the tumour organoids of all four mice. (b) Allele frequencies of coding and splicing somatic SNSs present in tumour organoids of mice I-IV and in primary and metastatic tumours of mouse IV. As expected the majority of somatic variants in organoids have a frequency around 50% (red boxes), whereas primary and metastatic tumors contain somatic SNSs with lower frequency (yellow and orange boxes, respectively) due to contamination by normal cells. (c) Coding and splice site SNSs present in sequenced tissues and organoids. Red, blue and green colours indicate nonsynonymous, synonymous and splice-site point mutations, respectively. For the primary and metastatic biopsies of mouse IV, SNSs with allele frequencies less than 20% or less than 10% are marked by boxes with decreased colour saturation.

Mouse IV was the only mouse that developed metastases to the liver. We attempted to establish organoids from the metastatic lesions, but were unable to do so (Fig. 1c, Fig. S1b). However, we sequenced the exomes of four distinct metastatic lesions (MT1-MT4) and also the exomes of the two primary tumour biopsies (PT1, PT2), from which we had successfully obtained organoid cultures (O1A-C-O2A-D). The four SNSs that were shared by all organoids were also present in the primary and metastatic tumour biopsies. The majority of the semi-private SNSs were present in at least one of the biopsies,

whereas the majority of the private SNSs were absent (Fig. 2c). The two primary tumour biopsies had similar, although not identical SNSs profiles, whereas three of the four metastatic lesions had distinct SNS profiles, indicating at least three independent metastasis seeding events (Fig. 2c).

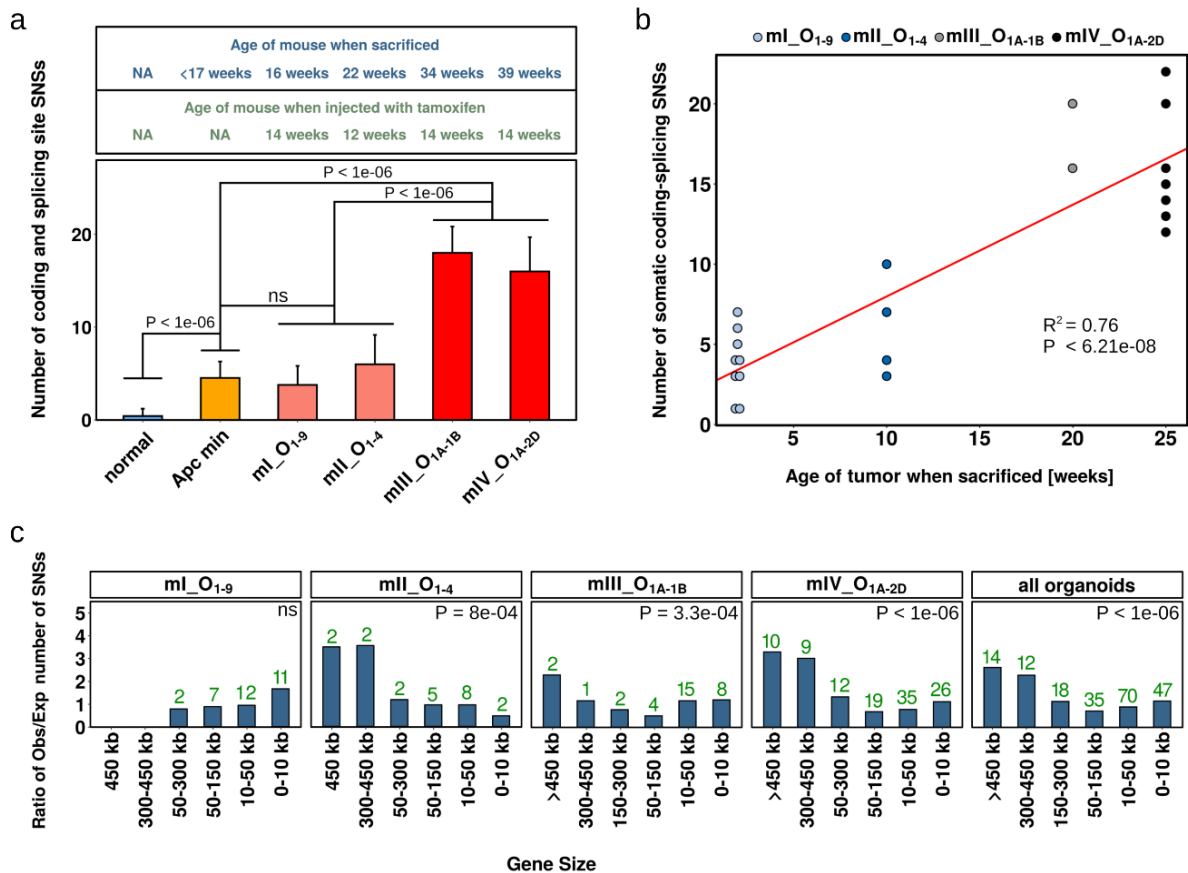
### **2.3.4.3 Single nucleotide substitutions – Link to genotype and Distribution**

We next examined whether there was a correlation between the genotype of the organoids and the number of SNSs acquired. The organoids from mice I and II, in which Apc was inactivated and mutant Kras was expressed, had accumulated fewer SNSs than the organoids derived from mice III and IV, in which Tp53 was also inactivated (Fig. 3a). Nevertheless, this difference might not be related to the genotype, because mice III and IV were sacrificed 20 and 25 weeks after tamoxifen administration, whereas mice I and II were sacrificed after two and ten weeks, respectively (Fig. 1b). Thus, the number of SNSs in the four mice correlated well with the time over which the tumours developed (Fig. 3b), implying a similar rate of SNS acquisition over time in all mice.

In a previous study, we had performed exome sequencing of single cell-derived organoids from Apc<sup>Min/+</sup> mice; these organoids originated either from adenomatous polyp or normal intestine tissue. The average number of SNSs in the transformed Apc<sup>Min/+</sup> organoids was very similar to the number of SNSs present in the organoids from mice I and II, which expressed mutant Kras, in addition to having inactivated Apc (Fig. 3a). In contrast, the organoids derived from non-transformed cells from the same mice had significantly fewer SNSs than the transformed organoids (Fig. 3a). These results suggest that expression of mutant Kras did not have a significant effect on the rate by which SNSs accumulate; on the other hand, transformed cells had a higher mutation rate than normal cells. We note that colon cells that have as cancer-drivers only mutant Apc or only mutant Apc and mutant Kras are precancerous in humans.

An interesting feature of SNSs in human precancerous and cancerous lesions is that they target more frequently large genes than small genes [22]. Mechanistically, one explanation is that a slower progression of replication forks in cancer cells could result in the central segments of large genes being replicated in mitosis by the break-induced replication mechanism, which is error-prone [17,44–50]. We examined the distribution of somatic SNSs according to gene size in the mouse tumour organoids and observed more

SNSs per Mb, than expected, in the large genes (Fig. 3c). This effect was statistically significant for mice II, III and IV, but not for mouse I, which developed tumours rapidly.



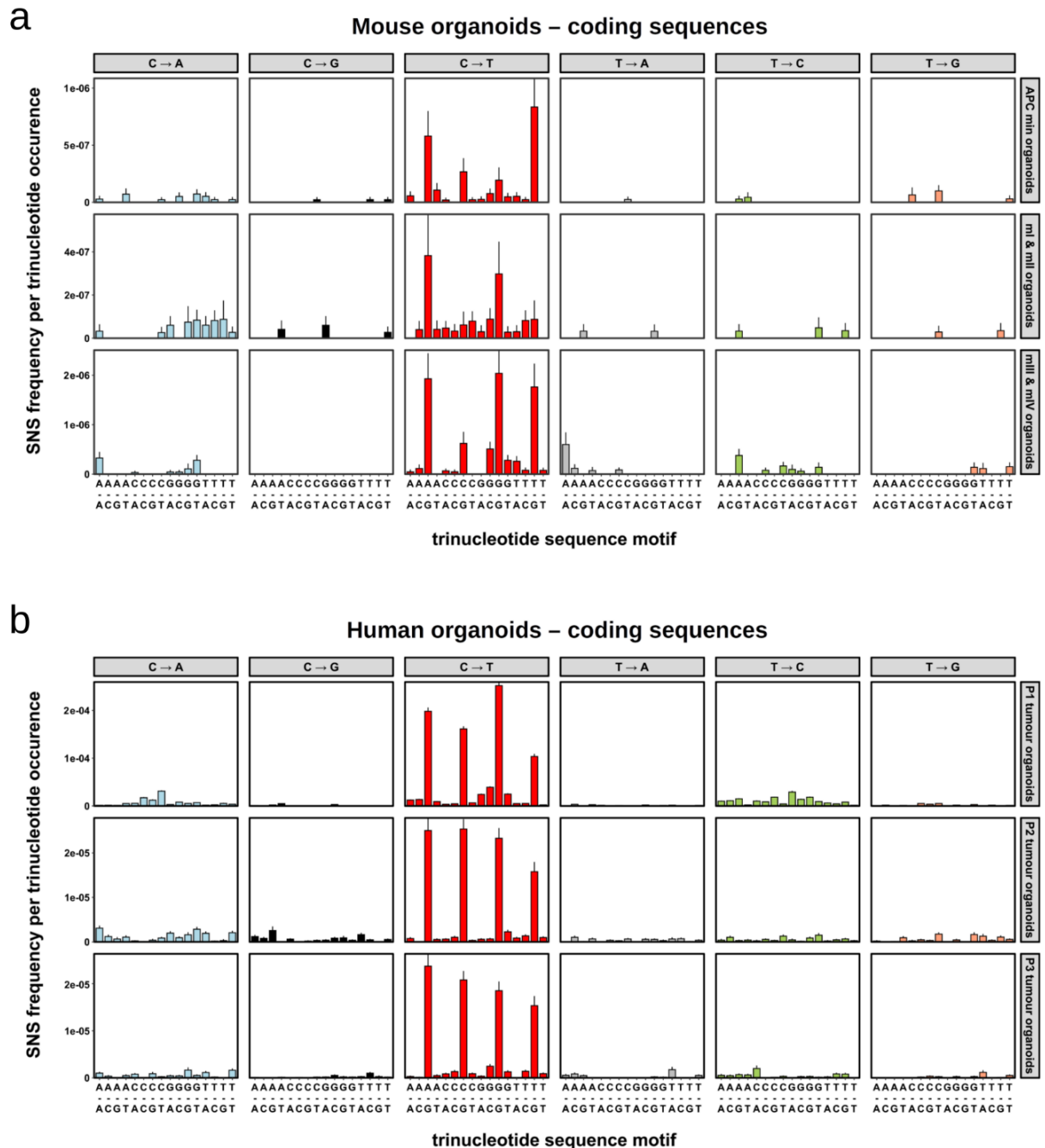
**Figure 3.** Relation of number of SNSs to age and distribution of SNSs according to gene size. a) Average number of somatic SNSs in normal and tumour organoids of  $Apc^{Min/+}$  mice and in tumour organoids of AKP mice. Organoids originated from mice I and II displayed a similar mutational burden to these of  $Apc^{Min/+}$  mice but lower than mice III and IV organoids. b) Linear correlation of the number of somatic coding and splice site SNSs with the time over which the tumours developed, calculated from the dates at which the mice were injected with tamoxifen and sacrificed. c) Distribution of SNSs according to gene size. The graphs show the ratios of observed (Obs) versus expected (Exp) number of SNSs for each gene category. The observed number of SNSs is indicated by the green letters. With the exception of mouse I, the somatic SNSs were significantly more prevalent in the large genes.

### 2.3.4.4 Single nucleotide substitutions – Mutational signature

SNSs in human cancers often target nucleotides in specific sequence contexts, which are referred to as mutational signatures. Signature 1, which is the most prevalent mutational signature in human cancers, describes the substitution of cytosines by thymines in the context of NpCpG motifs [25,34]. This signature, although present in most cancers, exhibits some tissue specificity and is particularly prevalent in precancerous lesions (adenomas) and cancers of the colon [22]. We had previously observed this



signature in organoids derived from intestinal adenomas of *Apc<sup>Min/+</sup>* mice [40]. In the current study, signature 1 was again the most prevalent signature (Fig. S7a). The high prevalence of signature 1 became even more evident when the number of substitutions was normalized by the frequency of the respective triplets in the genome (Fig. 4a), since the NpCpG triplet is quite underrepresented in the mouse genome (Table S1).



**Figure 4.** Mutational signatures of somatic SNSs in tumour organoids. a) Normalized signatures of protein-coding SNSs identified in the tumour organoids from mice I-IV of this study and from *Apc<sup>Min</sup>* mice. b) Normalized signatures of protein-coding SNSs identified in the tumour organoids from three CRC patients. The SNS frequencies were normalized according to the prevalence of the respective nucleotide triplets in the protein-coding sequences of the mouse (a) and human (b) reference genomes.

Interestingly, we also observed evidence for the presence of SNSs conforming to signature 17 in the mouse organoids of our current study and of our previous study of *Apc*<sup>Min/+</sup> mice (Fig. 4a and Fig S7a). Signature 17 is characterized by an elevated number of T to G and T to C substitutions in the context of CpTpT trinucleotides. Its origin is unknown, but it is particularly present in esophageal, stomach and colon human cancers [25,51–53].

To determine whether the mutational profiles observed in the mouse organoids were similar to those present in human cancers, we reanalyzed the published sequencing data of 36 human organoids derived from the tumours and normal tissues of three colorectal cancer patients [41]. The SNSs within the protein coding sequences revealed a strong signature 1 profile and a weak signature 17 profile, similar to what we observed in the mouse organoids (compare Figs 4b and S7b to Figs 4a and S7a, respectively).

Since the human organoids were subjected to whole-genome sequencing, we were able to examine more thoroughly their mutational signature profile. At the genome-wide level, signature 1 was by far the most prevalent signature, followed by signature 17 in patients 1 and 2 (Fig. S8a). Further analysis of the SNSs conforming to these signatures revealed a strong dependence on replication timing with more SNSs being present in late S than in early S replication regions (Fig. S8b). Within each replication timing domain, the frequency of SNSs was similar in the protein coding, intronic and intergenic regions; this was true for both signatures 1 and 17 (Fig. S8c). Finally, we note that the SNSs conforming to signature 1 and the C to T transitions in non-CpG contexts were the only SNS types present in organoids derived from non-transformed cells (Fig. S8a).

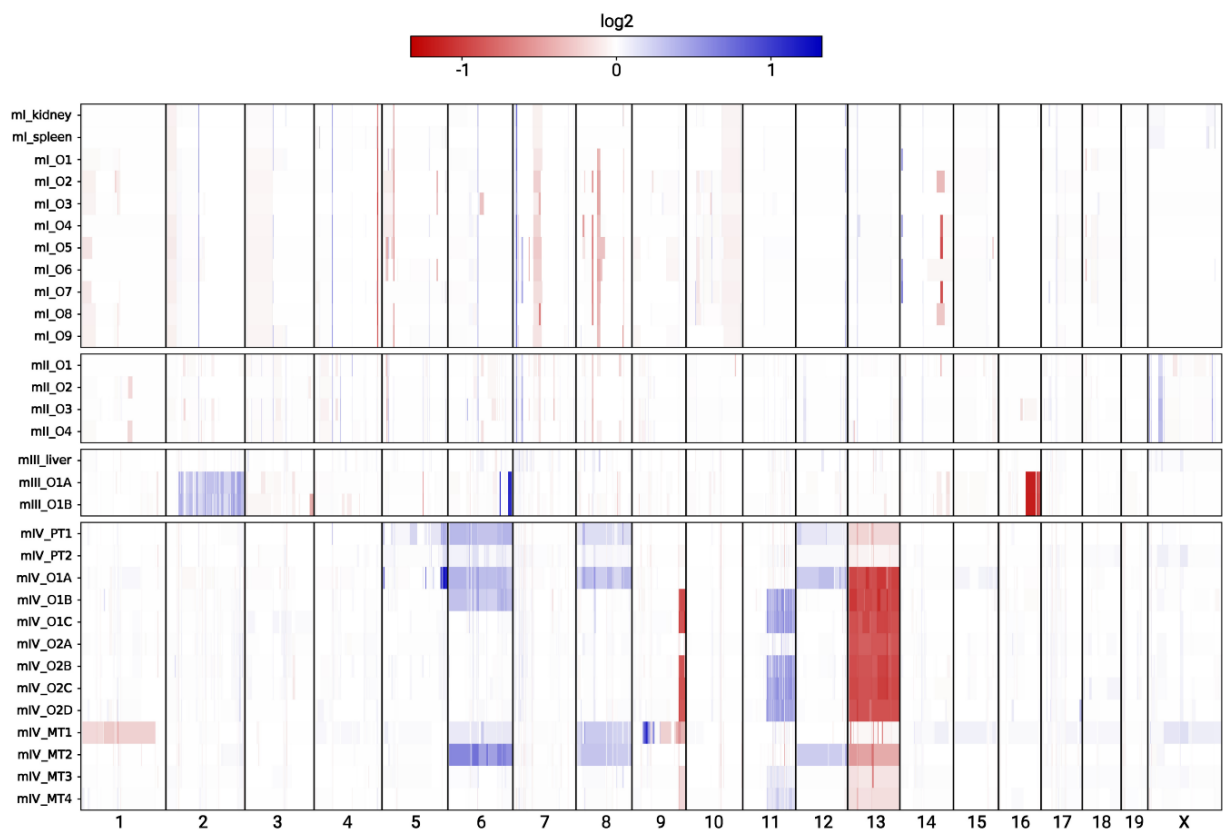
#### **2.3.4.5 Copy number alterations**

To probe for copy number alterations (CNAs) in the mouse tumour organoids, we compared the number of sequencing reads across the genome to the number of normal tissue reads from the same mouse. For this type of analysis, whole genome sequencing data are superior to exome sequencing data; yet the high read coverage of our data allowed us to identify copy number changes with a high degree of certainty (Fig. S9).

For the organoids from mice I and II, the analysis did not reveal any convincing CNAs (Fig. 5). Small genomic regions with different ratios of the number of sequencing reads in the organoid and reference bulk normal tissue were observed, but similar differences were observed when comparing the number of reads between different normal tissues.

Moreover, some of these differences were even shared between the organoids of mice I and II, which strongly indicates that they were noise (Fig. 5).

In contrast to mice I and II, several CNAs were evident in the organoids from mice III and IV. The two organoids from mouse III shared CNAs on chromosomes 2, 6 and 16, indicating that they were related (Fig. 5). These two organoids also shared 14 SNSs (Fig. 2c). The seven organoids from mouse IV all shared loss of one copy of chr 13; five of the seven organoids shared CNAs on chromosomes 9 and 11; and two of the seven organoids shared a duplication of chr 6. Finally, organoid O1A had private CNAs on chromosomes 5, 8 and 12 (Fig. 5). Many of the CNAs observed in the organoids from mouse IV were also evident in the primary and metastatic tumour biopsies. Interestingly, metastatic lesion 1 had several CNAs that were not present in any of the organoids or primary tumour samples; these CNAs included amplifications in chromosomes 9 and X and deletions in chromosomes 1 and 9 (Fig. 5).



**Figure 5.** Profiles of copy number alterations (CNAs) in the tumour organoids and tissue biopsies from mice I-IV of this study. Red and blue colours correspond to deletions and duplications, respectively, and colour saturation indicates the ratio of the number of reads in the organoid or biopsy compared to the normal reference tissue.

### 2.3.4.6 Evolution of tumour clones in mouse IV

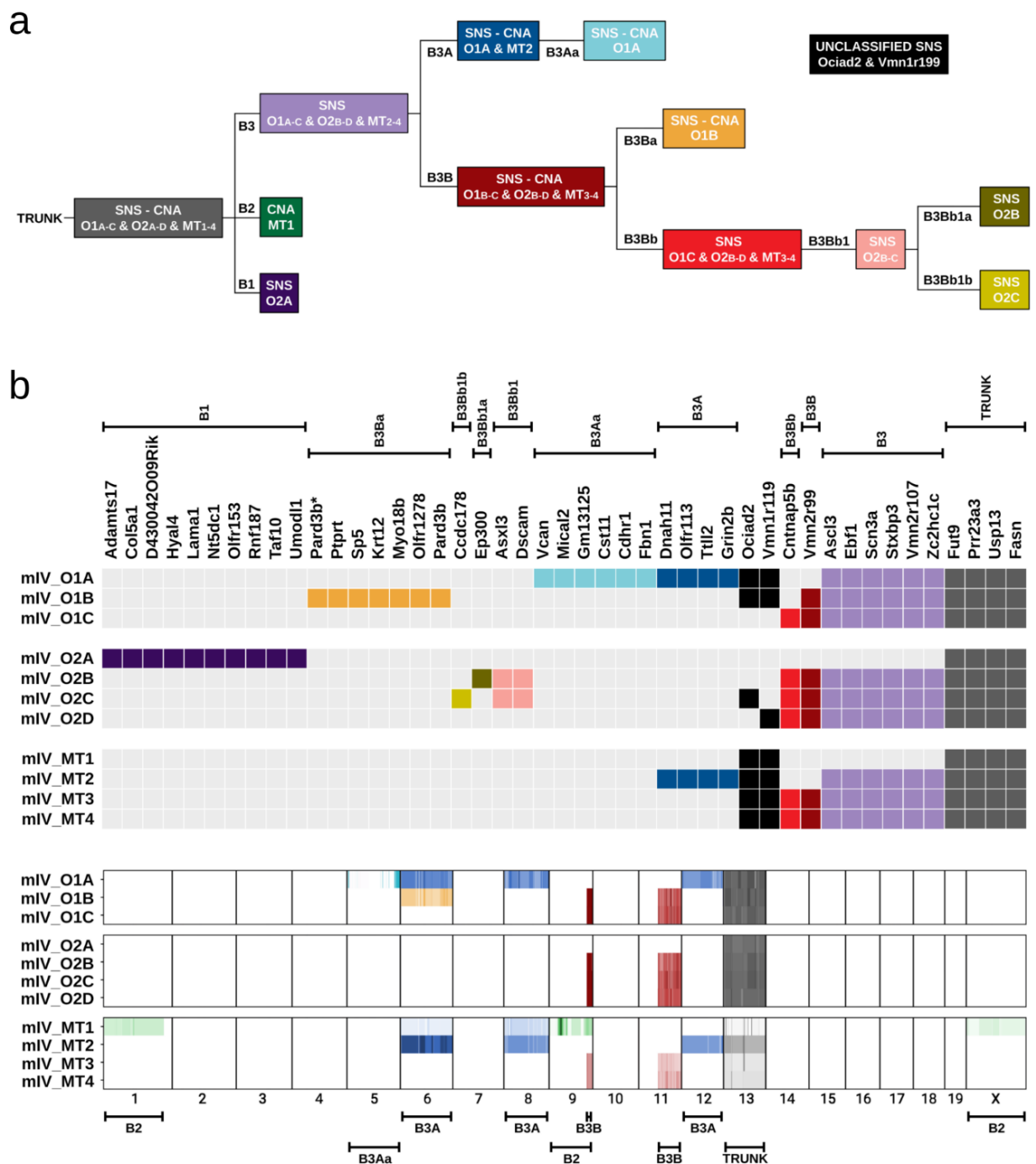
The availability of SNS and CNA data from several organoids and from primary and metastatic tumour samples of mouse IV provided an opportunity to establish an order in which these mutations were acquired during tumour evolution.

First, we plotted the allele frequencies of all the SNSs identified in the organoids and tumours of mouse IV (Fig. S10). In the organoids, most allele frequencies were close to 50%, consistent with one allele being mutated in a diploid region of the genome. For a few SNSs, the allele frequencies deviated significantly from 50%, but all these SNSs were located within genomic regions affected by CNAs. In the primary tumours, the SNS allele frequencies ranged between less than 5% to about 30%: the four SNSs that were identified in all organoids had allele frequencies of about 30%; whereas, the allele frequencies of the remaining SNSs were lower with the interesting exception of the SNSs that targeted the *Ociad2* and *Vmn1r119* genes, which had allele frequencies close to 30%, even though they were identified in only 3 out of the 7 organoids (Fig. S10). In the metastatic lesions all SNSs had similar allele frequencies, consistent with each metastasis having been seeded by a single cell or by a microcolony of genetically identical cells (Fig. S10). Accordingly, we included the metastatic lesions in the phylogenetic tree. We note that metastatic lesions 3 and 4 had exactly the same SNSs and CNAs (Fig. 2c and 5).

To plot the phylogenetic tree, we started with the four SNSs (SNS-TRUNK) and the deletion of one copy of chr 13 (CNA-TRUNK) that were present in all the organoids and metastatic lesions (Fig. 6). Three branches could be projected from the trunk of the phylogenetic tree. The first branch (B1) was formed by O2A and contained several private mutations (SNS-B1); the second branch (B2) was formed by MT1 and contained private CNAs affecting chromosomes 1, 9 and X (CNA-B2), as well as a duplications of chr 6 and chr 8; the third branch (B3) was formed by all other samples and contained a group of six SNSs (SNS-B3). In regard to branch B2, we note that the duplications of chr 6 and chr 8 were also present in the O1A and MT2 samples, raising the possibility that the metastatic lesion MT1 might not be monoclonal. Therefore, we attributed only the CNAs affecting chromosomes 1, 9 and X to branch B2 (Fig. 6b).

From branch B3, two branches originated; the first branch (B3A) encompassed O1A and MT2 and was characterized by the presence of four SNSs (SNS-B3A), and duplications of chromosomes 6, 8 and 12 (CNA-B3A). From this first branch, a sub-branch emerged containing O1A and characterized by a group of six SNSs (SNS-B3Aa) and

amplification events in chr 5 (CNA-B3Aa). The second branch (B3B) emerging from branch B3, was characterized by a SNS targeting Vnn2r99 (SNS-B3B), an amplification of part of chr 11 and a deletion of part of chr 9 (CNA-B3B).



**Figure 6.** Phylogenetic analysis of tumour progression in mouse IV. a) Phylogenetic tree illustrating the temporal order of acquisition of the SNSs and CNAs. The branches (B) of the tree are numbered. b) Colour-coding of the SNSs and CNAs according to the phylogenetic analysis. The SNSs and CNAs belonging to the same branch of the phylogenetic tree have the same colour.

In turn, two branches arose from branch B3B: branch B3Ba, which was formed by O1B and was characterized by 7 private SNSs (SNS-B3Ba) and duplication of chr 6 (CNA-B3Ba; we consider this to be an independent event from the duplication of chr 6 observed in O1A and MT2) and branch B3Bb, which was formed by O1C, O2B, O2C, O2D and MT3/MT4 and was characterized by a SNS targeting the *Cntnap5b* gene (SNS-B3Bb). From branch B3Bb, a branch (B3Bb1) containing O2B and O2C emerged; in turn, this branch gave rise to two branches characterized by SNS-B3Bb1a and SNS-B3Bb1b, respectively (Fig. 6).

The phylogenetic tree encompassed all the SNSs and CNAs with the notable exception of the SNSs targeting the *Ociad2* and *Vmn1r119* genes (Fig. 6), as these SNSs could not be incorporated in a way that made sense to us. What is evident is that tumour development was associated with the parallel emergence of SNSs and CNAs.

### 2.3.5 Discussion

Genomic instability is considered a major culprit of tumour development and emergence of resistance to therapy. While the presence of genomic instability in cancer was recognized at the beginning of the previous century, the advent of massive parallel sequencing has significantly advanced our understanding of the mechanisms leading to this cancer hallmark [19].

Our study focused on two major types of genomic instability: chromosomal instability and instability at the level of SNSs. The key question that we wanted to address was the extent to which the *Apc*, *Kras* and *Tp53* genes, which are frequently implicated in colon cancer development, contribute to the types of genomic instability mentioned above. As a model system we used tumour-prone mice that express mutant *Kras* and that inactivate *Apc* and, stochastically, *Tp53*. Single cell-derived organoids were examined to allow us to obtain a better understanding of cancer development in the face of tumour heterogeneity. We compared the current results to the results generated from our previous study, in which we performed exome sequencing of organoids derived from precancerous lesions and matching normal intestinal epithelium of *Apc*<sup>Min/+</sup> mice [40].

The analysis of the data revealed a rather simple picture. The rate of accumulation of SNSs was higher in the tumour-derived organoids than in the organoids derived from the normal epithelium, but unaffected by the number of targeted cancer-driver genes. Thus, the organoids with mutant *Apc* accumulated SNSs with the same rate as the organoids with mutant *Apc* and mutant *Kras* and even the organoids with mutant *Apc*, mutant *Kras* and mutant *Tp53*. In contrast, CNAs were present almost exclusively in the

organoids harbouring mutant Tp53. The observation that the rate of accumulation of SNSs was independent of the number of mutant cancer-driver genes was unexpected. In contrast, other studies had previously linked Tp53 mutations to the induction of CNAs [42,54,55].

As mentioned above, SNSs can exhibit specific patterns that are referred to as mutational signatures [23,25,56]. Signature 1 is the most dominant signature and is characterized by the presence of C to T substitutions in a CpG context; this signature arises from spontaneous deamination of methylated cytosines to thymines [22,35,56]. It has been proposed that signature 1 SNSs accumulate at a constant rate in both normal and transformed cells [36–38]. Thus, their number would reflect the age of the organism. However, in our previous study of organoids derived from *Apc*<sup>Min/+</sup> mice, the tumour-derived organoids had a significantly higher number of CpG to TpG transitions than the organoids derived from non-transformed cells [40]. Similarly, an analysis of the sequencing data of organoids derived from three human colon cancer patients revealed more signature 1 SNSs in the tumour-derived organoids than the organoids derived from normal tissue [41 and Fig. S8]. Nevertheless, signature 1 SNSs were present to a significant degree in the normal tissue-derived organoids. In contrast, SNSs that do not conform to signature 1, were present almost exclusively in the tumour-derived organoids (Fig. S8). We conclude that the mutagenic processes leading to signature 1 operate also in normal cells, albeit at a lower level than in tumour cells, whereas the mutagenic processes that lead to the other signatures are highly tumour-specific. Interestingly, all SNSs showed a dependence on replication timing with late S replicating regions being significantly more prone to mutagenesis than the early S replicating regions. This dependency may explain why there is a higher density of protein coding sequences in the early S replicating regions of the genome.

The sequencing of single cell-derived organoids from the same tumour makes it possible to construct a phylogenetic tree marking tumour development. In our study, this was possible for mouse IV. Our analysis revealed a phylogenetic tree characterized by the parallel emergence of SNSs and CNAs. In addition, analysis of metastatic lesions demonstrated that apart from MT1, the rest were derived from a single cell or genetically identical cells. In contrast, sequencing of the primary tumour tissue revealed a spectrum of mutations typical of a heterogeneous population of cancer cells.

### **2.3.6 Conclusions**

An analysis of organoids derived from a mouse model of colorectal cancer has allowed us to study the accumulation of SNSs and CNAs at the single cell level during tumour evolution. The key conclusions are that the rate of accumulation of SNSs is higher in transformed than non-transformed cells and that it is unaffected by the number of cancer-driver genes that are active in the tumour. Thus, tumours with mutant Apc accumulate as many SNSs as tumours with mutant Apc, mutant Kras and mutant Tp53. Signature 1 SNSs are the most prevalent in our model but are also present to a lower degree in normal cells. Moreover, late S replicating genomic regions are more prone to accumulate SNSs. In contrast to SNSs, CNAs were observed only in cells with mutant Tp53. So far, very few studies have been published sequencing single cell-derived tumour organoids. Nevertheless, this approach has great potential to elucidate the mutagenic processes present in cancer and, therefore, to contribute to our understanding of genomic instability.



## 2.3.7 Materials and Methods

### Mice

All mice were kept on a 12-hour light/dark cycle in individually ventilated cages. The  $Apc^{lox/lox}$  mice [57,58],  $Kras^{Lsl-G12D}$  mice [59] and  $Tp53^{lox/lox}$  mice [60] were crossed to  $Cdx2^{CreERT2/+}$  mice (The Jackson Laboratory, Charles River, L'Arbresle, France) [61] to obtain  $Apc^{lox/lox};Kras^{Lsl-G12D};Tp53^{lox/lox};Cdx2^{CreERT2/+}$  (AKP- $Cdx2^{CreERT2}$ ) animals (Fig. 1a, 1b). All experiments were authorized by the Canton of Vaud (license VD3396) and were performed according to accepted guidelines for animal handling.

### Genotyping

A small amount of tissue from each mouse was used for genotyping. Confirmation of recombination events upon tamoxifen injection was assessed by genotyping the corresponding organoids selected for sequencing. For genotyping, the organoids from one well of a 48-well plate were lysed in 200  $\mu$ L lysis buffer supplemented with 150  $\mu$ g of Proteinase K and then incubated overnight at 55 °C. The lysates were diluted 10 times with water and subjected to PCR amplification using GoTaq Hot Start Polymerase (M7423, Promega).

The following primers were used for genotyping: for the cre allele: Cre\_Fw (5'-CACCAGCCAGCTATCAACTCG-3') and Cre\_Rev (5'-TTACATTGGTCCAGCCA CCAG-3'); for the Apc allele: Apc\_Fw (5'-GTTCTGTATCATGGAAAGATAGGTG GTC-3') and Apc\_Rev1 (5'-CACTCAAACGCTTTTGGAGGGTTGATTC-3') or Apc\_Rev2 (5'-GAGTACGGGGTCTCTGTCTCAGTGAA-3'); for the Tp53 allele: Tp53\_Fw1 (5'-CACAAAAACAGGTTAAACCCA-3') or Tp53\_Fw2 (5'-AAGGGG TATGAGGGACAAGG-3') and Tp53\_Rev (5'-GAAGACAGAAAAGGGGAGGG-3'); for the KrasG12D allele: Kras\_WT\_Fw (5'-TGTCTTTCCCCAGCACAGT-3') or Kras\_MUT\_Fw (5'-CCATGGCTTGAGTAAGTCTGC-3') and Kras\_common\_rev (5'-CTGCATAGTACGCTATAACCCTGT-3'). The PCR conditions and DNA fragment sizes obtained are described in the Supplementary Information section.

### Induction of tumour formation

Tamoxifen (Sigma), 3 mg/kg (or 30 mg/kg for mouse I), was administered either as a single i.p. injection with sunflower oil (mouse II) or by gavage with peanut oil (mouse I, mouse III and mouse IV), when the mice were 10-14 weeks old. The mice were sacrificed 2 weeks (mouse I), 10 weeks (mouse II), 20 weeks (mouse III) or 25 weeks (mouse IV) after tamoxifen administration.

## **Tissue isolation, organoid culture and expansion**

Intestinal tissue (colon or caecum) was isolated from AKP-Cdx2<sup>CreERT2</sup> mice. Colonic tumours were distinct and each was dissected and treated separately. For caecum tumours, which effectively occupied nearly all the caecum space, the caecum was divided in several parts and each part was considered as a separate tumour. For each tumour, the resected tissue was cut into 2-3 mm wide cubes, that were separated by tissue also 2-3 mm wide. The tumours of mouse III, isolated 20 weeks after tamoxifen administration, were small in size and were not cut into separate pieces before processing (see Fig. 1c and Figure S1b).

The tumour fragments were washed thoroughly in PBS-EDTA at 4 °C and then homogenized with a teflon pestle in 1.5 ml Eppendorf tubes. Tissue homogenates were treated with Trypsin-EDTA for 3-4 min and quickly pipetted up and down, approximately 100-200 times, using 200 µl tips to disrupt any cell aggregates. After centrifugation, the pellets were resuspended in ENR media, filtered through 70 µm cell strainers (BD Bioscience) and single cell suspensions were mixed with cold Matrigel® (Corning) and plated in 96-well plates. The tissue culture media (ENR) for these organoids was based on DMEM/F12 (Life Technologies) supplemented with B27 and N2 (Life Technologies), supplemented with 10 mM HEPES, 100 units/mL of penicillin and 100 µg/mL of streptomycin (Life Technologies), 2mM L-Glutamine (Life Technologies) and 1.25 µM N-Acetylcysteine (Sigma-Aldrich). The following growth factors were also added: 50 ng/mL murine recombinant EGF (Life Technologies), R-Spondin1-Fc and Noggin-6xHis [62]. After the Matrigel solidified, ENR tissue culture media was added on top. The suspensions of single cells were seeded at different cell concentrations to obtain wells containing a single organoid. Selected organoids from each tumour piece were then expanded to obtain enough material for DNA sequencing. For these organoids, the medium was changed every 2 days and organoids were split every 3-4 days by mechanical dissociation. Organoids were kept in culture as short as possible to obtain the necessary amount of DNA for exome sequencing; on average, each organoid culture was split three times. Around four to six 48-well plates full of organoids were harvested and the organoid pellets were washed and frozen at -80 °C.

## **DNA Extraction and Exome Sequencing**

Genomic DNA from the organoids was extracted and fragmented by sonication. The resultant fragments (~200 bp) were subjected to exome capture using the SureSelect

Mouse All Exon Kit (Agilent Technologies) and paired-end libraries were prepared and sequenced on an Illumina HiSeq 4000 platform.

### **Sequence Analysis**

Sequencing reads were aligned on the mouse reference genome NCBI Build GRCm38/mm10 using the Burrows-Wheeler Alignment tool v.0.7.17. Bam conversion, sorting, removal of PCR duplicates and indexing of the sequence alignment files was conducted by SAMtools v.1.9. Somatic variant calling was performed by GATK v.4.11.0 using healthy tissue from the liver or kidney of the same mouse as matching normal sample. False positive calls were filtered out using a panel of normal samples constructed from all normal tissues of 4 mice. Variants present in common mouse dbSnp142 were also discarded. The mutational spectra of detected somatic SNSs were examined using the SomaticSignature v.2.20.0 R package for the analysis of all the 96 possible trinucleotide changes. CNA events in bam files were analysed by VarScan2 v.2.4.3 using the recommended workflow. To filter out somatic CNA events, we excluded CNAs that were present in the liver, kidney and spleen tissues of the mice from which the organoids were prepared. Segmentation was applied by DNACopy R package v.1.58.

### **SNS signature normalization**

The SNS signatures in mouse and human samples were normalized using the genomic mouse and human sequences, respectively, downloaded from the NCBI RefSeq curated dataset at the UCSC server (<http://hgdownload.soe.ucsc.edu/goldenPath/>). The coordinates of early-S, mid-S and late-S replicating regions of the human genome were obtained from our previous analysis of U2OS cells [63].

### **Supplementary Materials:**

The following are available online at [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1), Figure S1-S10, Table S1 and Data S1. The FASTQ files of mouse exome sequencing have been deposited at the NCBI SRA database under the BioProject ID PRJNA675998.

### **Author Contributions:**

T.D.H., P.O.M. and J.H. conceived the project and designed the experiments, M.N., P.O.M., A.K. and G.R. performed experiments, V.D. established the bioinformatics pipelines and performed computational analysis with contributions from T.D.H. V.D., M.N., P.O.M., J.H. and T.D.H. wrote the manuscript.

### **Data Availability Statement:**

The FASTQ files of mouse exome sequencing have been deposited at the NCBI SRA database under the BioProject ID PRJNA675998.

### **Institutional Review Board Statement:**

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Canton de Vaud (license VD3396 and date of approval 08.01.2019).”

**Funding:**

This research was funded by the Carigest Foundation.

**Acknowledgments:**

The Apcl<sup>ox</sup>/lox mice were a kind gift of Tatiana Petrova, University of Lausanne, Switzerland. We thank Michael R. Stratton for granting us access to the human colorectal cancer data presented in the publication doi: 10.1038/s41586-018-0024-3.

**Conflicts of Interest:**

The authors declare no conflict of interest.

### 2.3.8 References

1. Kinzler, K.W.; Vogelstein, B. Lessons from hereditary colorectal cancer. *Cell* 1996, 87, 159-170.
2. Hamilton, S.R.; Powell, S.M.; Papadopoulos, N.; Smith, K.J.; Vogelstein, B.; Kinzler, K.W. Molecular Determinants of Dysplasia in Colorectal Lesions. *Cancer Res* 1994, 54, 5523-5526.
3. Vogelstein, B.; Kinzler, K.W. The multistep nature of cancer. *Trends Genet.* 1993, 9, 138-141.
4. Powell, S.M.; Zilz, N.; Beazer-Barclay, Y.; Bryan, T.M.; Hamilton, S.R.; Thibodeau, S.N.; Vogelstein, B.; Kinzler, K.W. APC mutations occur early during colorectal tumorigenesis. *Nature* 1992, 359, 235-237.
5. Fearon, E.R.; Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* 1990, 61, 759-767.
6. Vogelstein, B.; Fearon, E.R.; Hamilton, S.R.; Kern, S.E.; Preisinger, A.C.; Leppert, M.; Smits, A.M.M.; Bos, J.L. Genetic Alterations during Colorectal-Tumor Development. *N. Engl. J. Med.* 1988, 319, 525-532.
7. Roberts, D.M.; Pronobis, M.I.; Poulton, J.S.; Kane, E.G.; Peifer, M. Regulation of Wnt signaling by the tumor suppressor adenomatous polyposis coli does not require the ability to enter the nucleus or a particular cytoplasmic localization. *Mol. Biol. Cell* 2012, 23, 2041-2056;
8. Ghazvini, M.; Sonneveld, P.; Kremer, A.; Franken, P.; Sacchetti, A.; Atlasi, Y.; Roth, S.; Joosten, R.; Smits, R.; Fodde, R. Cancer Stemness in Apc- vs. Apc/KRAS-Driven Intestinal Tumorigenesis. *PLoS One* 2013, 8, e73872.
9. Schneikert, J.; Behrens, J. The canonical Wnt signalling pathway and its APC partner in colon cancer development. *Gut* 2007, 56, 417-425.
10. Stastna, M.; Janeckova, L.; Hrckulak, D.; Kriz, V.; Korinek, V. Human colorectal cancer from the perspective of mouse models. *Genes (Basel)*. 2019, 10, 788.
11. Tauriello, D.V.F.; Calon, A.; Lonardo, E.; Batlle, E. Determinants of metastatic competency in colorectal cancer. *Mol. Oncol.* 2017, 11, 97-119.
12. Jackstadt, R.; Sansom, O.J. Mouse models of intestinal cancer. *J. Pathol.* 2016, 238, 141-151.
13. Sottoriva, A.; Kang, H.; Ma, Z.; Graham, T.A.; Salomon, M.P.; Zhao, J.; Marjoram, P.; Siegmund, K.; Press, M.F.; Shibata, D.; et al. A big bang model of human colorectal tumor growth. *Nat. Genet.* 2015, 47, 209-216.

14. Rodrigues, N.R.; Rowan, A.; Smith, M.E.F.; Kerr, I.B.; Bodmer, W.F.; Gannon, J. V.; Lane, D.P. p53 mutations in colorectal cancer. *Proc. Natl. Acad. Sci. U. S. A.* 1990, 87, 7555-7559.
15. Weinberg, R.A. Oncogenes, Antioncogenes, and the Molecular Bases of Multistep Carcinogenesis. *Cancer Res.* 1989, 49, 3713-3721.
16. Janssen, K.P.; Alberici, P.; Fsihi, H.; Gaspar, C.; Breukel, C.; Franken, P.; Rosty, C.; Abal, M.; El Marjou, F.; Smits, R.; et al. APC and Oncogenic KRAS Are Synergistic in Enhancing Wnt Signaling in Intestinal Tumor Formation and Progression. *Gastroenterology* 2006, 131, 1096-1109.
17. Halazonetis, T.D.; Gorgoulis, V.G.; Bartek, J. An oncogene-induced DNA damage model for cancer development. *Science* 2008, 319, 1352-1355.
18. Hadjihannas, M. V.; Behrens, J. CIN by Wnt: Growth pathways, mitotic control and chromosomal instability in cancer. *Cell Cycle* 2006, 5, 2077-2081.
19. Negrini, S.; Gorgoulis, V.G.; Halazonetis, T.D. Genomic instability an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.* 2010, 11, 220-228.
20. Swanton, C. Intratumor heterogeneity: Evolution through space and time. *Cancer Res.* 2012, 72, 4875-4882.
21. Gerlinger, M.; Rowan, A.J.; Horswell, S.; Larkin, J.; Endesfelder, D.; Gronroos, E.; Martinez, P.; Matthews, N.; Stewart, A.; Tarpey, P.; et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *N. Engl. J. Med.* 2012, 366, 883-892.
22. Nikolaev, S.I.; Sotiriou, S.K.; Pateras, I.S.; Santoni, F.; Sougioultzis, S.; Edgren, H.; Almusa, H.; Robyr, D.; Guipponi, M.; Saarela, J.; et al. A single-nucleotide substitution mutator phenotype revealed by exome sequencing of human colon adenomas. *Cancer Res.* 2012, 72, 6279-6289.
23. Alexandrov, L.B.; Kim, J.; Haradhvala, N.J.; Huang, M.N.; Tian Ng, A.W.; Wu, Y.; Boot, A.; Covington, K.R.; Gordenin, D.A.; Bergstrom, E.N.; et al. The repertoire of mutational signatures in human cancer. *Nature* 2020, 578, 94-101.
24. Van Hoeck, A.; Tjoonk, N.H.; Van Boxtel, R.; Cuppen, E. Portrait of a cancer: Mutational signature analyses for cancer diagnostics. *BMC Cancer* 2019, 19, 457.
25. Alexandrov, L.B.; Nik-Zainal, S.; Wedge, D.C.; Aparicio, S.A.J.R.; Behjati, S.; Biankin, A. V.; Bignell, G.R.; Bolli, N.; Borg, A.; Børresen-Dale, A.L.; et al. Signatures of mutational processes in human cancer. *Nature* 2013, 500, 415-421.

26. Nik-Zainal, S.; Alexandrov, L.B.; Wedge, D.C.; Van Loo, P.; Greenman, C.D.; Raine, K.; Jones, D.; Hinton, J.; Marshall, J.; Stebbings, L.A.; et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* 2012, 149, 979-993.
27. Kane, A.M.; Fennell, L.J.; Liu, C.; Borowsky, J.; McKeone, D.M.; Bond, C.E.; Kazakoff, S.; Patch, A.M.; Koufariotis, L.T.; Pearson, J.; et al. Alterations in signaling pathways that accompany spontaneous transition to malignancy in a mouse model of BRAF mutant microsatellite stable colorectal cancer. *Neoplasia (United States)* 2020, 22, 120-128.
28. Niknafs, N.; Zhong, Y.; Moral, J.A.; Zhang, L.; Shao, M.X.; Lo, A.; Makohon-Moore, A.; Iacobuzio-Donahue, C.A.; Karchin, R. Characterization of genetic subclonal evolution in pancreatic cancer mouse models. *Nat. Commun.* 2019, 10, 5435.
29. Connor, F.; Rayner, T.F.; Aitken, S.J.; Feig, C.; Lukk, M.; Santoyo-Lopez, J.; Odom, D.T. Mutational landscape of a chemically-induced mouse model of liver cancer. *J. Hepatol.* 2018, 69, 840-850.
30. Fantini, D.; Glaser, A.P.; Rimar, K.J.; Wang, Y.; Schipma, M.; Varghese, N.; Rademaker, A.; Behdad, A.; Yellapa, A.; Yu, Y.; et al. A Carcinogen-induced mouse model recapitulates the molecular alterations of human muscle invasive bladder cancer. *Oncogene* 2018, 37, 1911-1925.
31. Liu, H.; Murphy, C.J.; Karreth, F.A.; Emdal, K.B.; Yang, K.; White, F.M.; Elemento, O.; Toker, A.; Wulf, G.M.; Cantley, L.C. Identifying and targeting sporadic oncogenic genetic aberrations in mouse models of triple-negative breast cancer. *Cancer Discov.* 2018, 8, 354-369.
32. Tauriello, D.V.F.; Palomo-Ponce, S.; Stork, D.; Berenguer-Llargo, A.; Badia-Ramentol, J.; Iglesias, M.; Sevillano, M.; Ibiza, S.; Cañellas, A.; Hernando-Momblona, X.; et al. TGF $\beta$  drives immune evasion in genetically reconstituted colon cancer metastasis. *Nature* 2018, 554, 538-543.
33. Nassar, D.; Latil, M.; Boeckx, B.; Lambrechts, D.; Blanpain, C. Genomic landscape of carcinogen-induced and genetically induced mouse skin squamous cell carcinoma. *Nat. Med.* 2015, 21, 946-954.
34. Sjöblom, T.; Jones, S.; Wood, L.D.; Parsons, D.W.; Lin, J.; Barber, T.D.; Mandelker, D.; Leary, R.J.; Ptak, J.; Silliman, N.; et al. The consensus coding sequences of human breast and colorectal cancers. *Science* 2006, 314, 268-274.
35. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* 1993, 362, 709-715.

36. Tomasetti, C.; Vogelstein, B. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science*. 2015, 347, 78-81.
37. Vogelstein, B.; Papadopoulos, N.; Velculescu, V.E.; Zhou, S.; Diaz, L.A.; Kinzler, K.W. Cancer genome landscapes. *Science*. 2013, 339, 1546-1558.
38. Tomasetti, C.; Vogelstein, B.; Parmigiani, G. Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. *Proc. Natl. Acad. Sci. U.S.A.* 2013, 110, 1999-2004.
39. McGranahan, N.; Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* 2017, 168, 613-628.
40. Lugli, N.; Dionellis, V.S.; Ordóñez-Morán, P.; Kamileri, I.; Sotiriou, S.K.; Huelsken, J.; Halazonetis, T.D. Enhanced Rate of Acquisition of Point Mutations in Mouse Intestinal Adenomas Compared to Normal Tissue. *Cell Rep.* 2017, 19, 2185-2192.
41. Roerink, S.F.; Sasaki, N.; Lee-Six, H.; Young, M.D.; Alexandrov, L.B.; Behjati, S.; Mitchell, T.J.; Grossmann, S.; Lightfoot, H.; Egan, D.A.; et al. Intra-tumour diversification in colorectal cancer at the single-cell level. *Nature* 2018, 556, 457-462.
42. Drost, J.; Van Jaarsveld, R.H.; Ponsioen, B.; Zimmerlin, C.; Van Boxtel, R.; Buijs, A.; Sachs, N.; Overmeer, R.M.; Offerhaus, G.J.; Begthel, H.; et al. Sequential cancer mutations in cultured human intestinal stem cells. *Nature* 2015, 521, 43-47.
43. Matano, M.; Date, S.; Shimokawa, M.; Takano, A.; Fujii, M.; Ohta, Y.; Watanabe, T.; Kanai, T.; Sato, T. Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat. Med.* 2015, 21, 256-262.
44. Elango, R.; Osia, B.; Harcy, V.; Malc, E.; Mieczkowski, P.A.; Roberts, S.A.; Malkova, A. Repair of base damage within break-induced replication intermediates promotes kataegis associated with chromosome rearrangements. *Nucleic Acids Res.* 2019, 47, 9666-9684.
45. Costantino, L.; Sotiriou, S.K.; Rantala, J.K.; Magin, S.; Mladenov, E.; Helleday, T.; Haber, J.E.; Iliakis, G.; Kallioniemi, O.P.; Halazonetis, T.D. Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science* 2014, 343, 88-91.
46. Deem, A.; Keszthelyi, A.; Blackgrove, T.; Vayl, A.; Coffey, B.; Mathur, R.; Chabes, A.; Malkova, A. Break-induced replication is highly inaccurate. *PLoS Biol.* 2011, 9, e1000594.

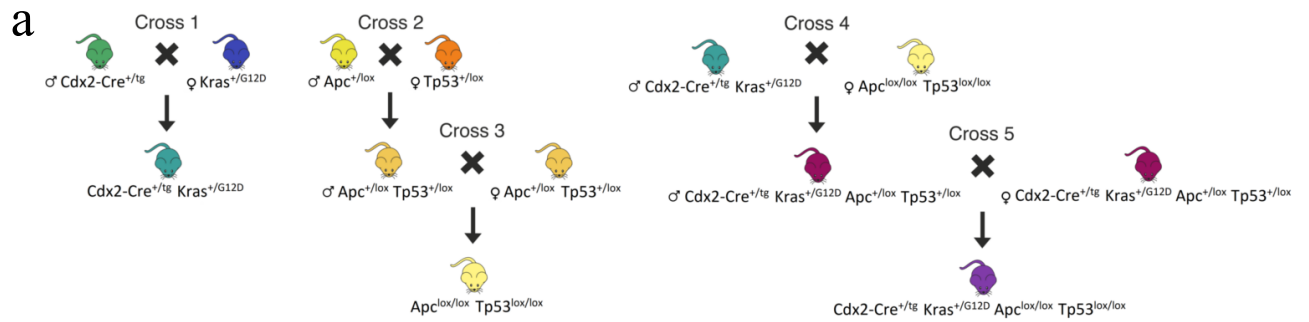


47. McAvoy, S.; Ganapathiraju, S.C.; Ducharme-Smith, A.L.; Pritchett, J.R.; Kosari, F.; Perez, D.S.; Zhu, Y.; James, C.D.; Smith, D.I. Non-random inactivation of large common fragile site genes in different cancers. *Cytogenet. Genome Res.* 2007, 118, 260-269.
48. Helmrich, A.; Stout-Weider, K.; Hermann, K.; Schrock, E.; Heiden, T. Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes. *Genome Res.* 2006, 16, 1222-1230.
49. Smith, D.I.; Zhu, Y.; McAvoy, S.; Kuhn, R. Common fragile sites, extremely large genes, neural development and cancer. *Cancer Lett.* 2006, 232, 48-57.
50. Gorgoulis, V.G.; Vassiliou, L.V.F.; Karakaidos, P.; Zacharatos, P.; Kotsinas, A.; Liloglou, T.; Venere, M.; DiTullio, R.A.; Kastriakis, N.G.; Levy, B.; et al. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* 2005, 434, 907-913.
51. Secrier, M.; Li, X.; De Silva, N.; Eldridge, M.D.; Contino, G.; Bornschein, J.; Macrae, S.; Grehan, N.; O'Donovan, M.; Miremadi, A.; et al. Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance. *Nat. Genet.* 2016, 48, 1131-1141.
52. Dulak, A.M.; Stojanov, P.; Peng, S.; Lawrence, M.S.; Fox, C.; Stewart, C.; Bandla, S.; Imamura, Y.; Schumacher, S.E.; Shefler, E.; et al. Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. *Nat. Genet.* 2013, 45, 478-486.
53. Christensen, S.; Van der Roest, B.; Besselink, N.; Janssen, R.; Boymans, S.; Martens, J.W.M.; Yaspo, M.L.; Priestley, P.; Kuijk, E.; Cuppen, E.; et al. 5-Fluorouracil treatment induces characteristic T>G mutations in human cancer. *Nat. Commun.* 2019, 10(1), 4571.
54. Donehower, L.A.; Soussi, T.; Korkut, A.; Liu, Y.; Schultz, A.; Cardenas, M.; Li, X.; Babur, O.; Hsu, T.K.; Lichtarge, O.; et al. Integrated Analysis of TP53 Gene and Pathway Alterations in The Cancer Genome Atlas. *Cell Rep.* 2019, 28, 1370-1384.e5.
55. Eischen, C.M. Genome stability requires p53. *Cold Spring Harb. Perspect. Med.* 2016, 6, a026096.
56. Helleday, T.; Eshtad, S.; Nik-Zainal, S. Mechanisms underlying mutational signatures in human cancers. *Nat. Rev. Genet.* 2014, 15, 585-598.
57. Sansom, O.J.; Reed, K.R.; Hayes, A.J.; Ireland, H.; Brinkmann, H.; Newton, I.P.; Battle, E.; Simon-Assmann, P.; Clevers, H.; Nathke, I.S.; et al. Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. *Genes Dev.* 2004, 18, 1385-1390.

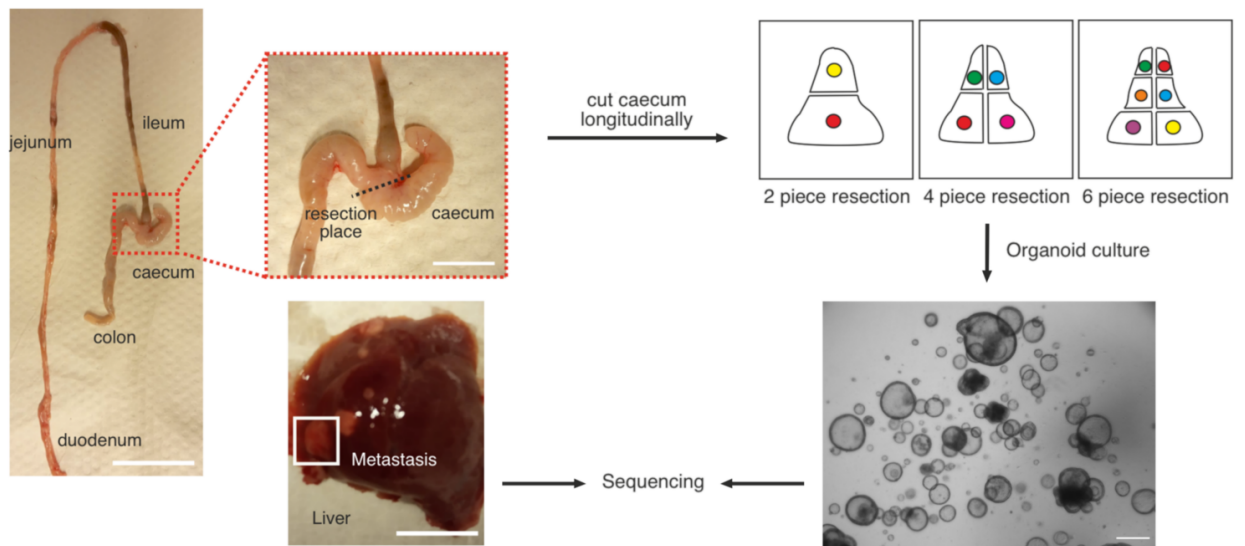
58. Shibata, H.; Toyama, K.; Shioya, H.; Ito, M.; Hirota, M.; Hasegawa, S.; Matsumoto, H.; Takano, H.; Akiyama, T.; Toyoshima, K.; et al. Rapid colorectal adenoma formation initiated by conditional targeting of the APC gene. *Science* 1997, 278, 120-123.
59. Jackson, E.L.; Willis, N.; Mercer, K.; Bronson, R.T.; Crowley, D.; Montoya, R.; Jacks, T.; Tuveson, D.A. Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* 2001, 15, 3243-3248.
60. Marino, S.; Vooijs, M.; Van Der Gulden, H.; Jonkers, J.; Berns, A. Induction of medulloblastomas in p53-null mutant mice by somatic inactivation of Rb in the external granular layer cells of the cerebellum. *Genes Dev.* 2000, 14, 994-1004.
61. Feng, Y.; Sentani, K.; Wiese, A.; Sands, E.; Green, M.; Bommer, G.T.; Cho, K.R.; Fearon, E.R. Sox9 induction, ectopic paneth cells, and mitotic spindle axis defects in mouse colon adenomatous epithelium arising from conditional biallelic Apc inactivation. *Am. J. Pathol.* 2013, 183, 493-503.
62. Ordóñez-Morán, P.; Dafflon, C.; Imajo, M.; Nishida, E.; Huelsken, J. HOXA5 Counteracts Stem Cell Traits by Inhibiting Wnt Signaling in Colorectal Cancer. *Cancer Cell* 2015, 28, 815-829.
63. Macheret, M., Halazonetis, T.D. Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. *Nature* 2018, 555(7694), 112-116.

## 2.3.9 Supplementary figures and tables

**Fig.S1**

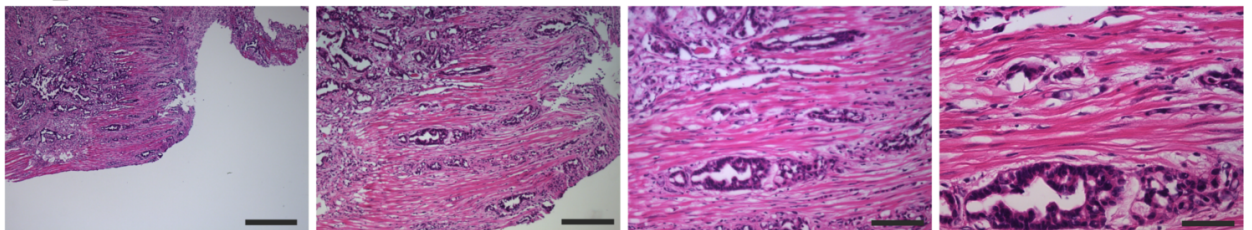


**b**

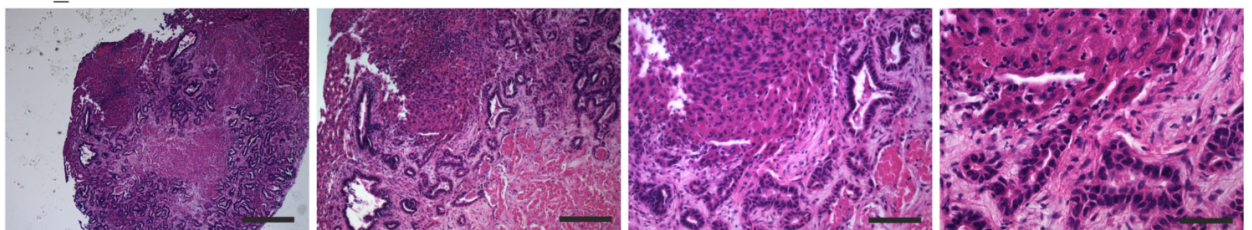


**c**

**mIV\_PT2**



**mIV\_MT1**

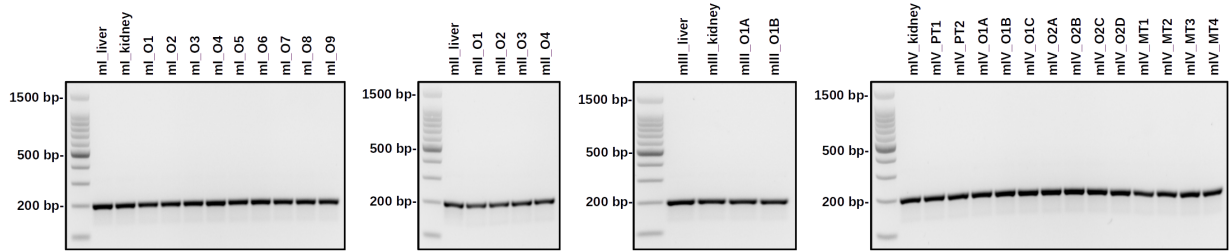


**Fig S1. Mouse breeding and preparation of organoids.** **a)** Breedings to generate AKP-Cdx2<sup>CreERT2</sup> mice. **b)** Resection of primary tumours and metastatic lesions. Macroscopically visible intestinal tumors were resected, followed by isolation of single cells from intestinal crypts and expansion as 3D organoid cultures. Tumour organoids derived from the colonic tumoral cells grew as spheroids and lacked crypt-like projections typical of healthy intestinal cells. Phase contrast images of AKP-Cdx2<sup>CreERT2</sup> tumour organoids expanded from a single cancer cell, after 1 week in culture. Scale bars, 100  $\mu$ m. **c)** H&E staining was performed on primary tumour 2 (PT2) and liver metastatic tissue (MT1) of mouse IV. Scale bars correspond to 500  $\mu$ m, 200 $\mu$ m, 100 $\mu$ m, 50 $\mu$ m from left to right.

Fig.S2

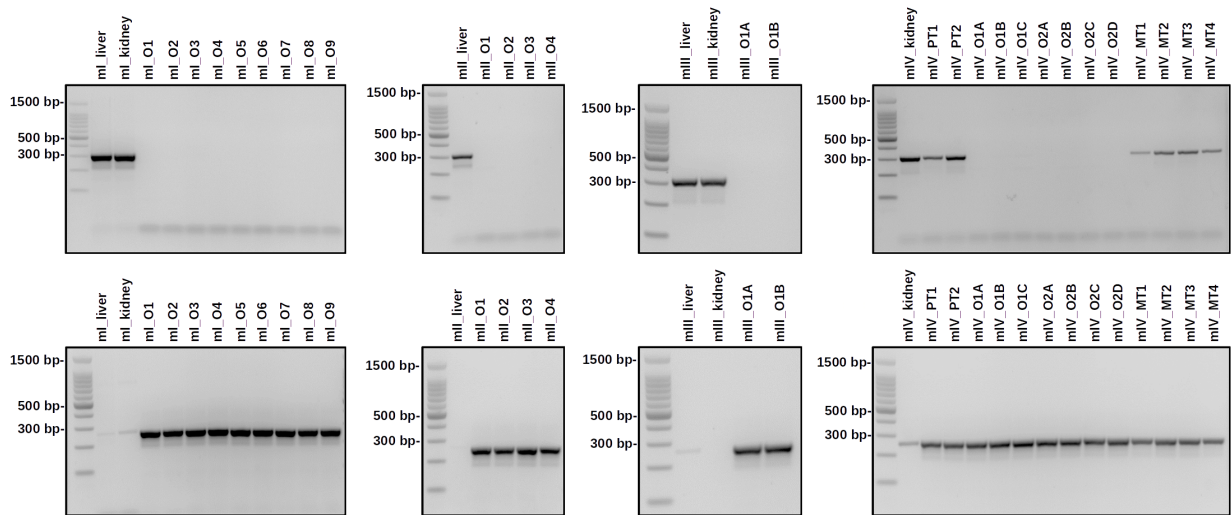
a

*Cre*



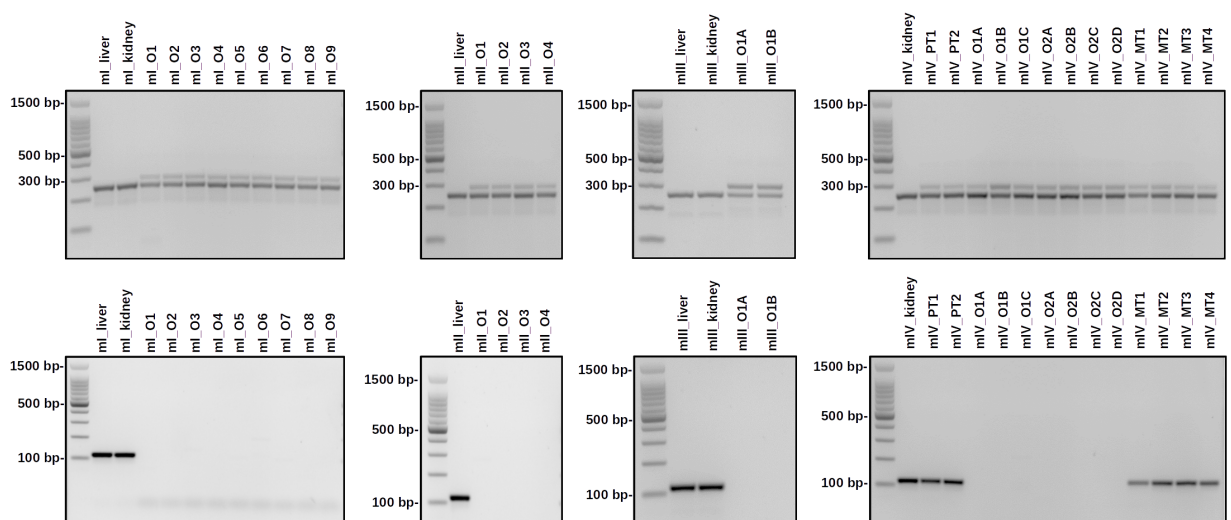
b

*Apc*

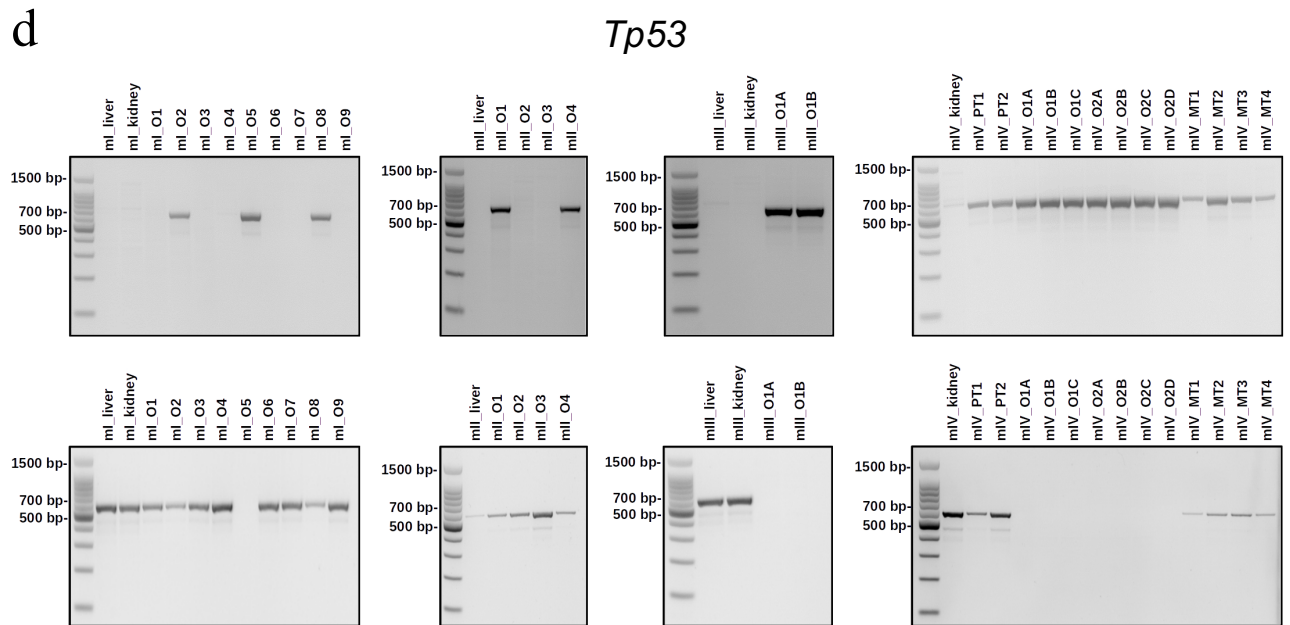


c

*Kras*



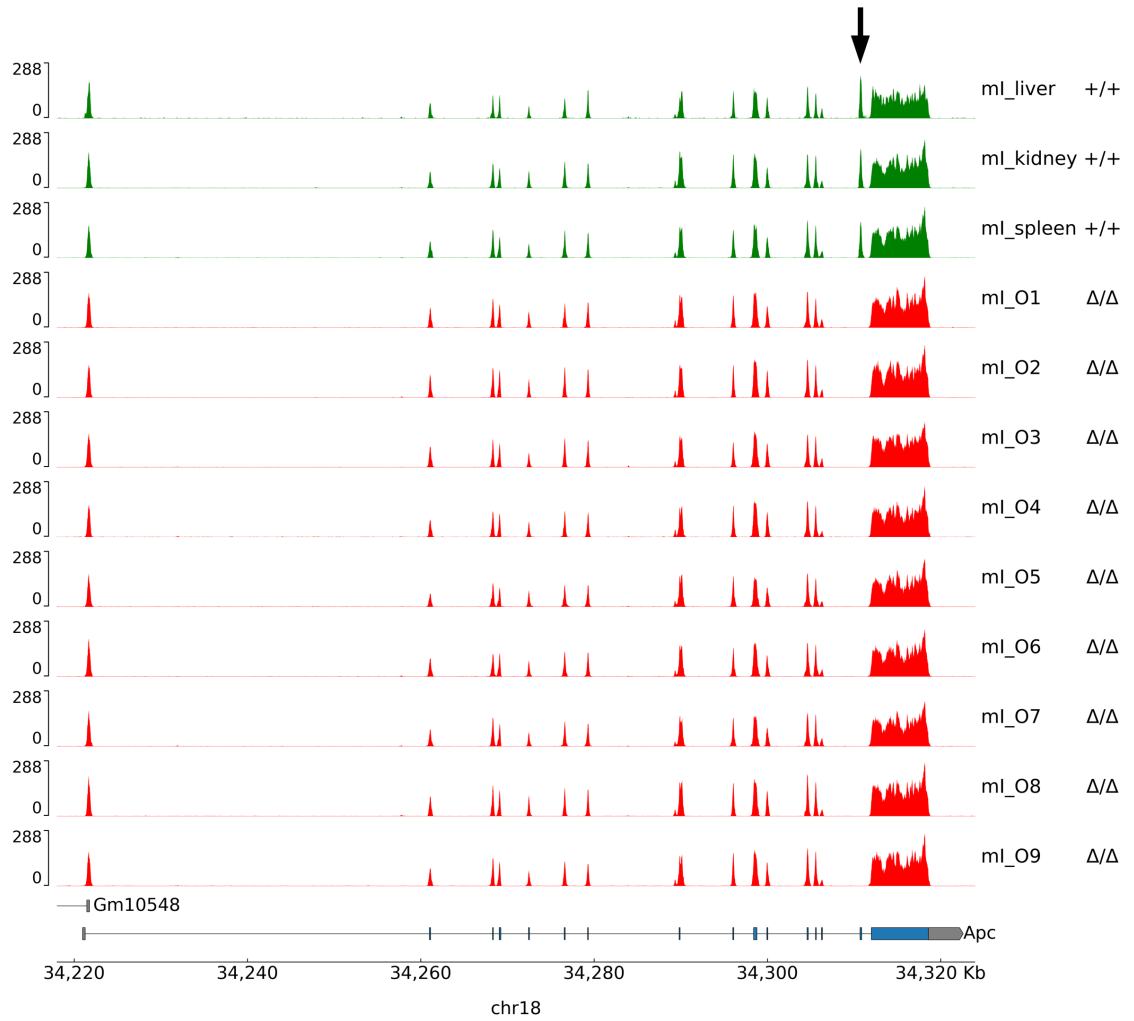
## Fig.S2 (continued)



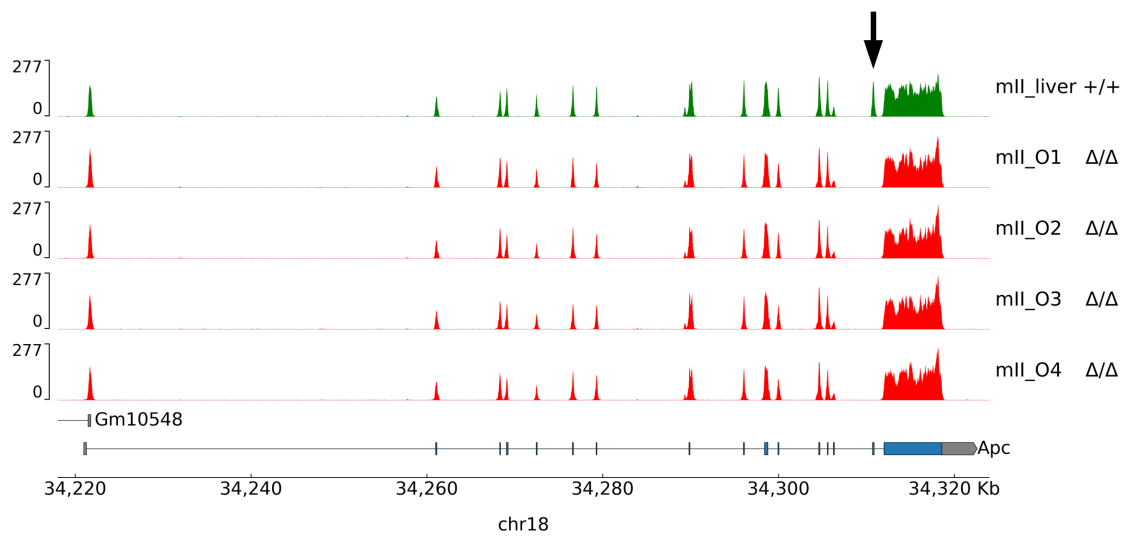
**Fig S2. Mouse genotyping by PCR. a)** Genotyping using the primer pair Cre\_Fw-Cre\_Rev. The observed size of a 200-bp PCR product indicates *Cre*-positive samples. **b)** Genotyping for *Apc* using the pair of primers Apc\_Fw-Apc\_Rev1 (top panels) and Apc\_Fw-Apc\_Rev2 (bottom panels). PCR product sizes of 314-bp and 225-bp correspond to mutant and wt alleles, respectively (top panels). A PCR product size of 258-bp designates the mutant recombinant allele (bottom panels). **c)** Mouse genotyping for *Kras* using the pairs of primers Kras\_WT\_Fw-Kras\_common\_rev (top panels) and Kras\_WT\_Fw-Kras\_MUT\_rev (bottom panels). PCR product sizes of 250-bp and 290-bp correspond to wt and recombined mutant alleles, respectively (top panels). A PCR product size of 119-bp designates the presence of the mutant non-recombined allele (bottom panels). **d)** Mouse genotyping for *Tp53* using the pairs of primers Tp53\_Fw-Tp53\_Rev (top panels) and Tp53\_Fw2-Tp53\_Rev (bottom panels). A PCR product size of 612-bp indicates the presence of the mutant recombined allele (top panels). PCR product sizes of 584-bp and 431-bp indicate the mutant non-recombined and wt alleles, respectively (bottom panels).

**Fig.S3**

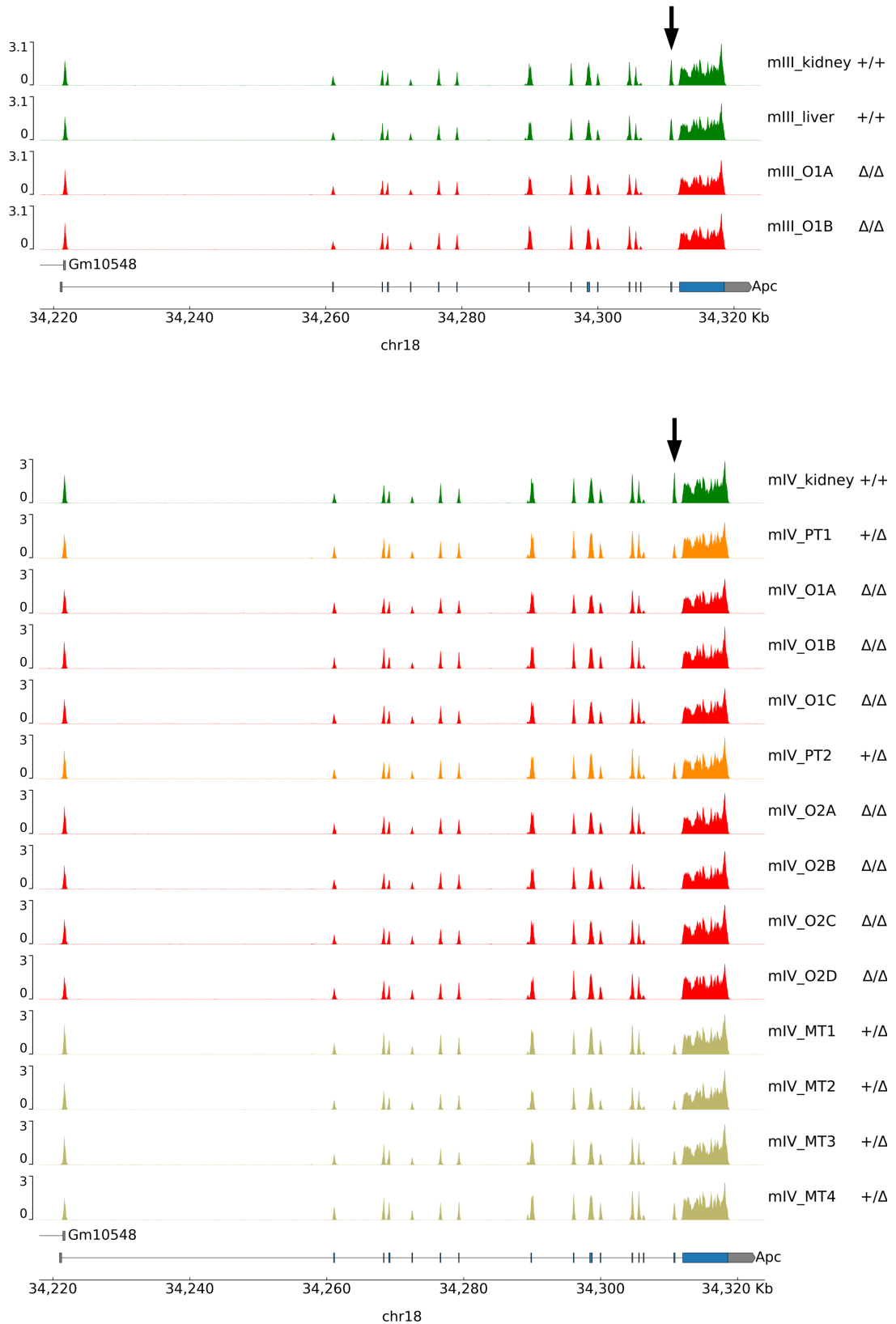
**a**



**b**



**Fig.S3 (continued)**

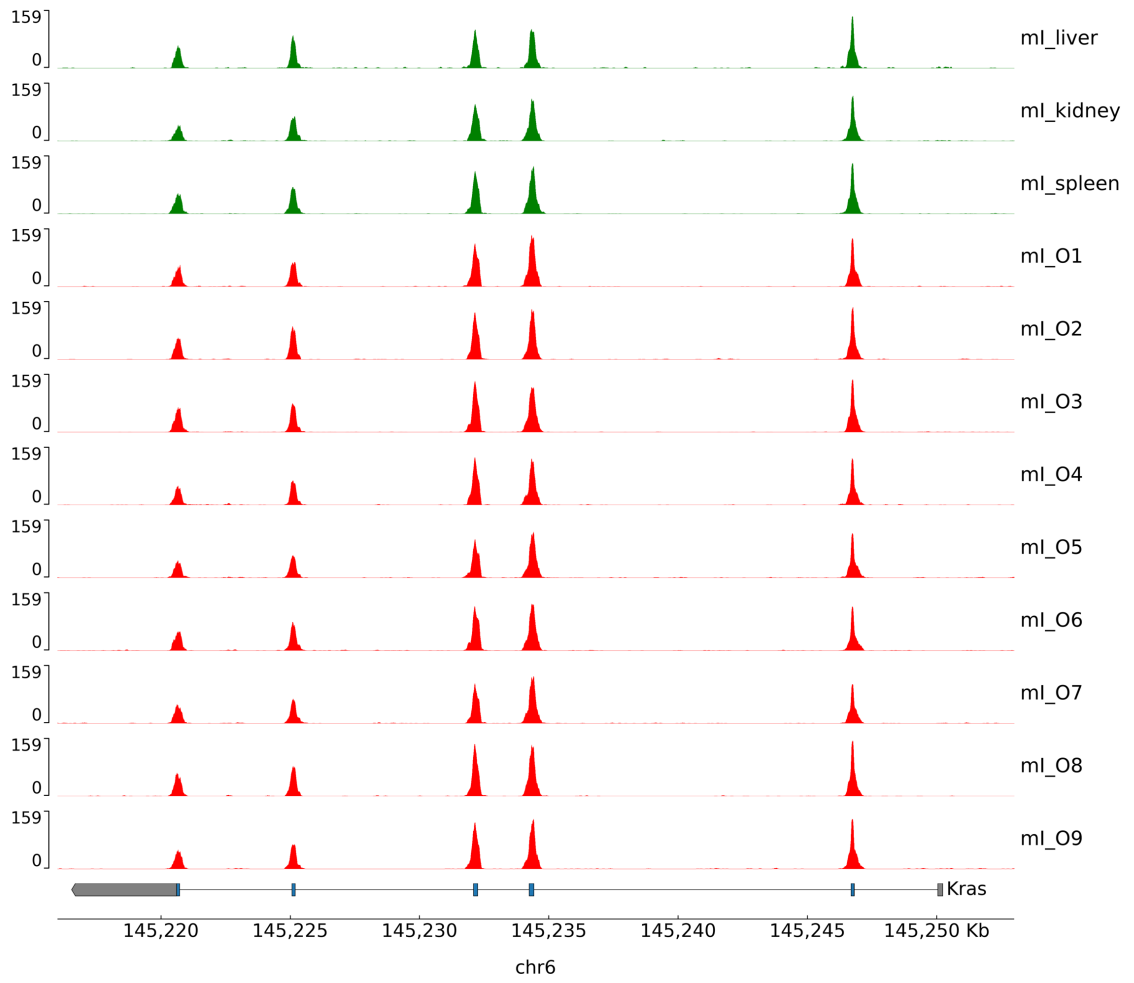


**Fig S3. Whole exome sequencing read depth of *Apc* genomic locus for mouse I (a), mouse II (b), mouse III (c) and mouse IV(d).** Normal tissues and tumour organoids are represented by green and red colours, respectively. Primary tumours and metastatic tissues are illustrated by orange and yellow colours, respectively. Black arrows point to the deletion of exon 15. The genotype is indicated next to the sample name.

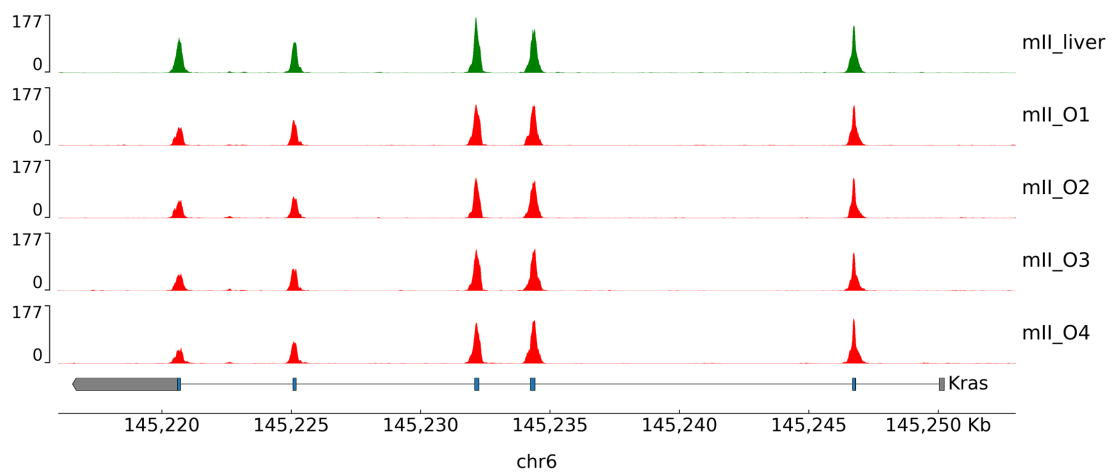


**Fig.S4**

**a**



**b**

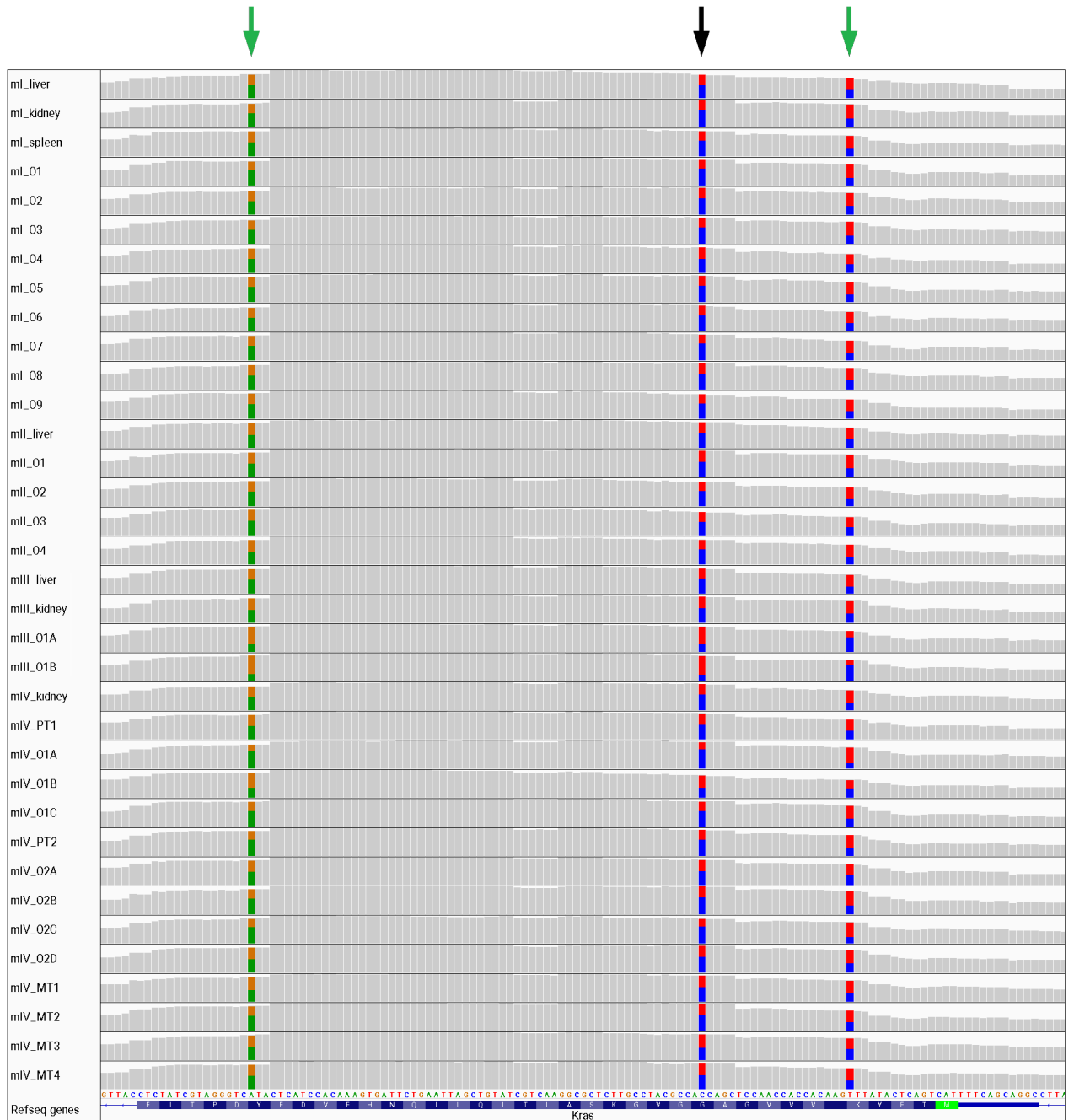


**Fig.S4 (continued)**



**Fig S4. Whole exome sequencing read depth of *Kras* genomic locus for mouse I (a), mouse II (b), mouse III (c) and mouse IV(d).** Normal tissues and tumour organoids are represented by green and red colours, respectively. Primary tumours and metastatic tissues are illustrated by orange and yellow colours, respectively.

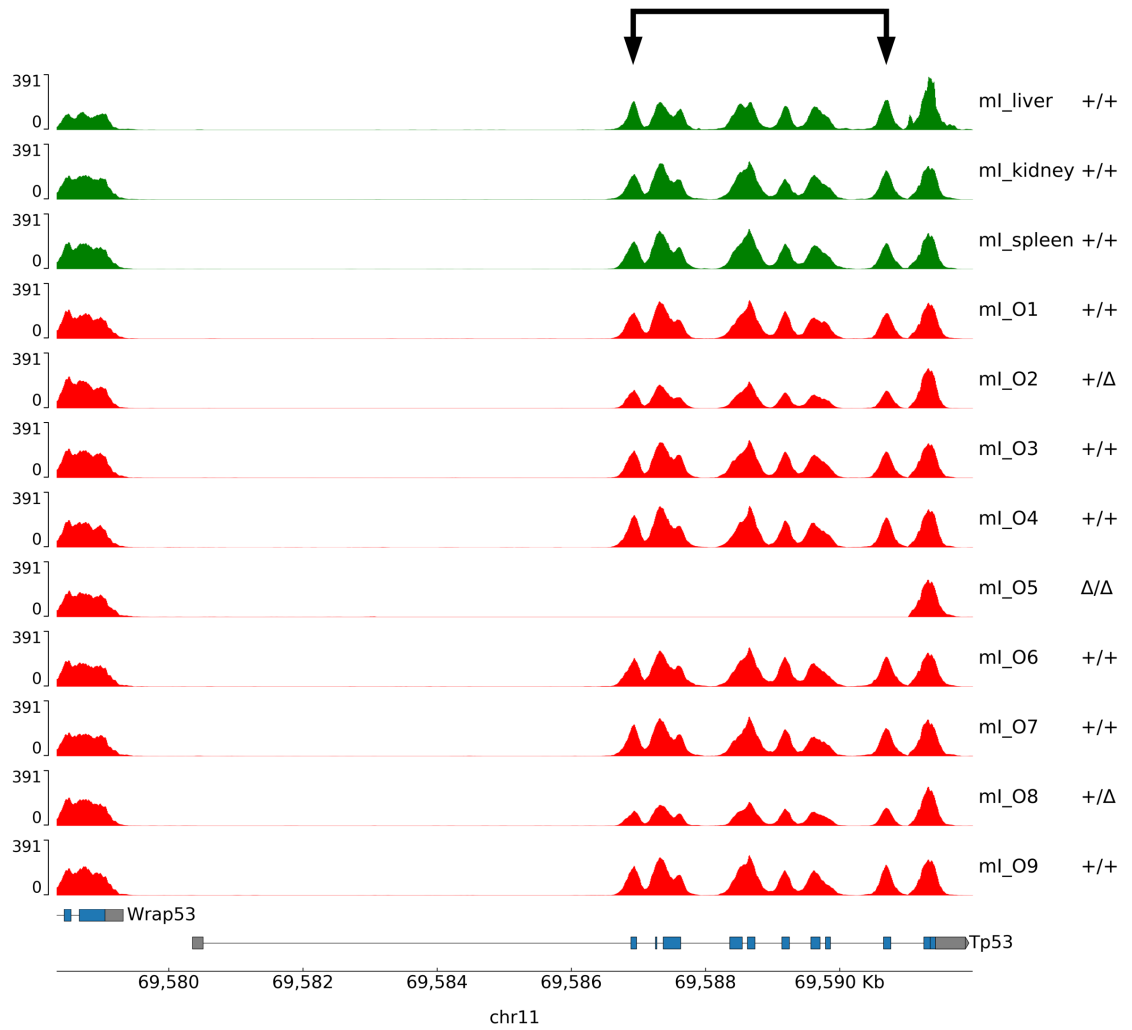
**Fig.S5**



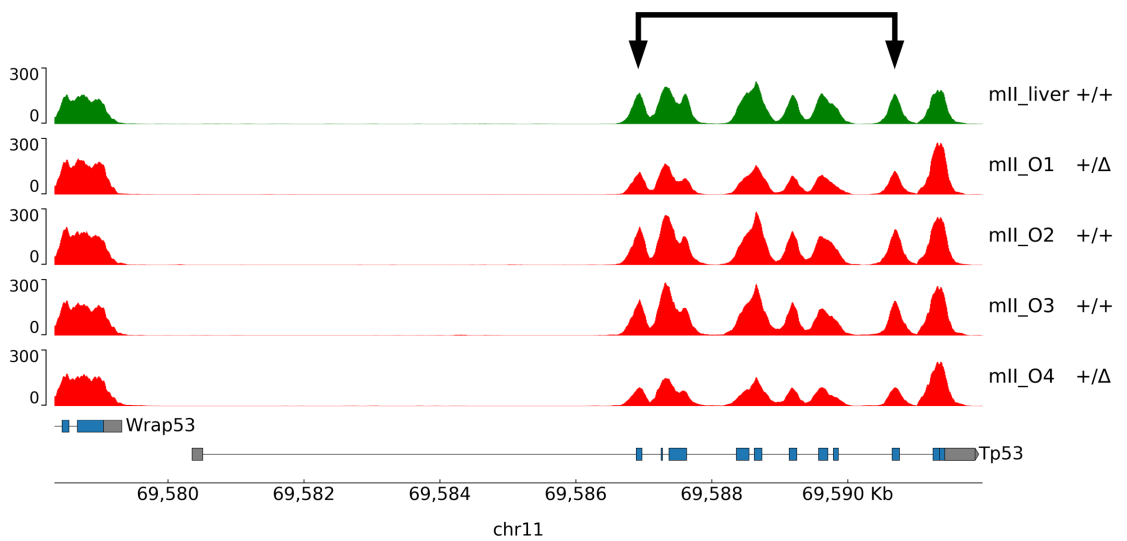
**Fig S5. *Kras* exon2 genotype detected by exome sequencing of mouse samples.** All sequenced tissues and organoids of mice I, II, III and IV carried the *G12D Kras* mutation (black arrow). This mutation is accompanied by two additional strain specific silent mutations (green arrows). Green, blue, red and orange colours correspond to A,C,T and G, respectively.

Fig.S6

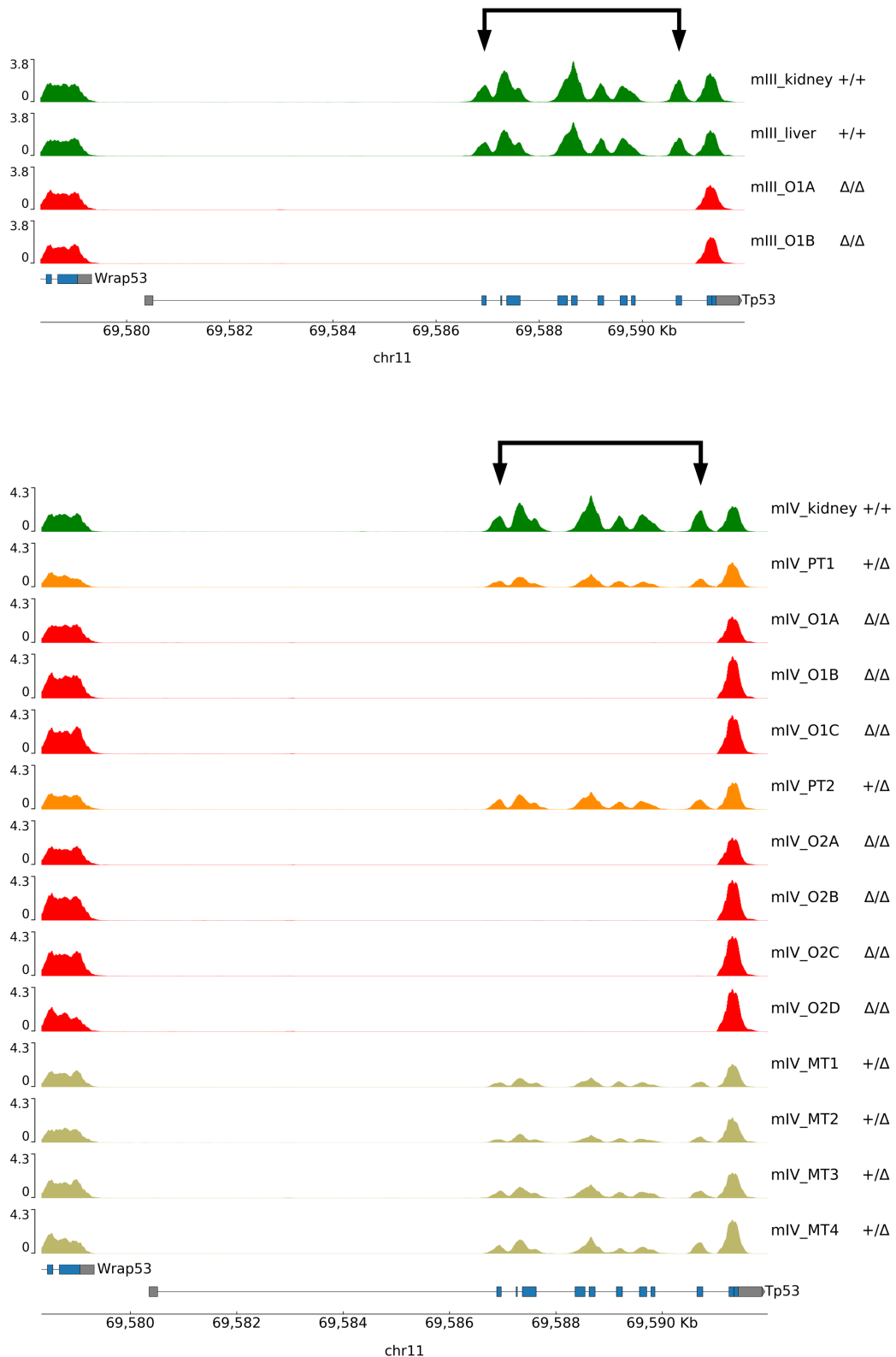
a



b

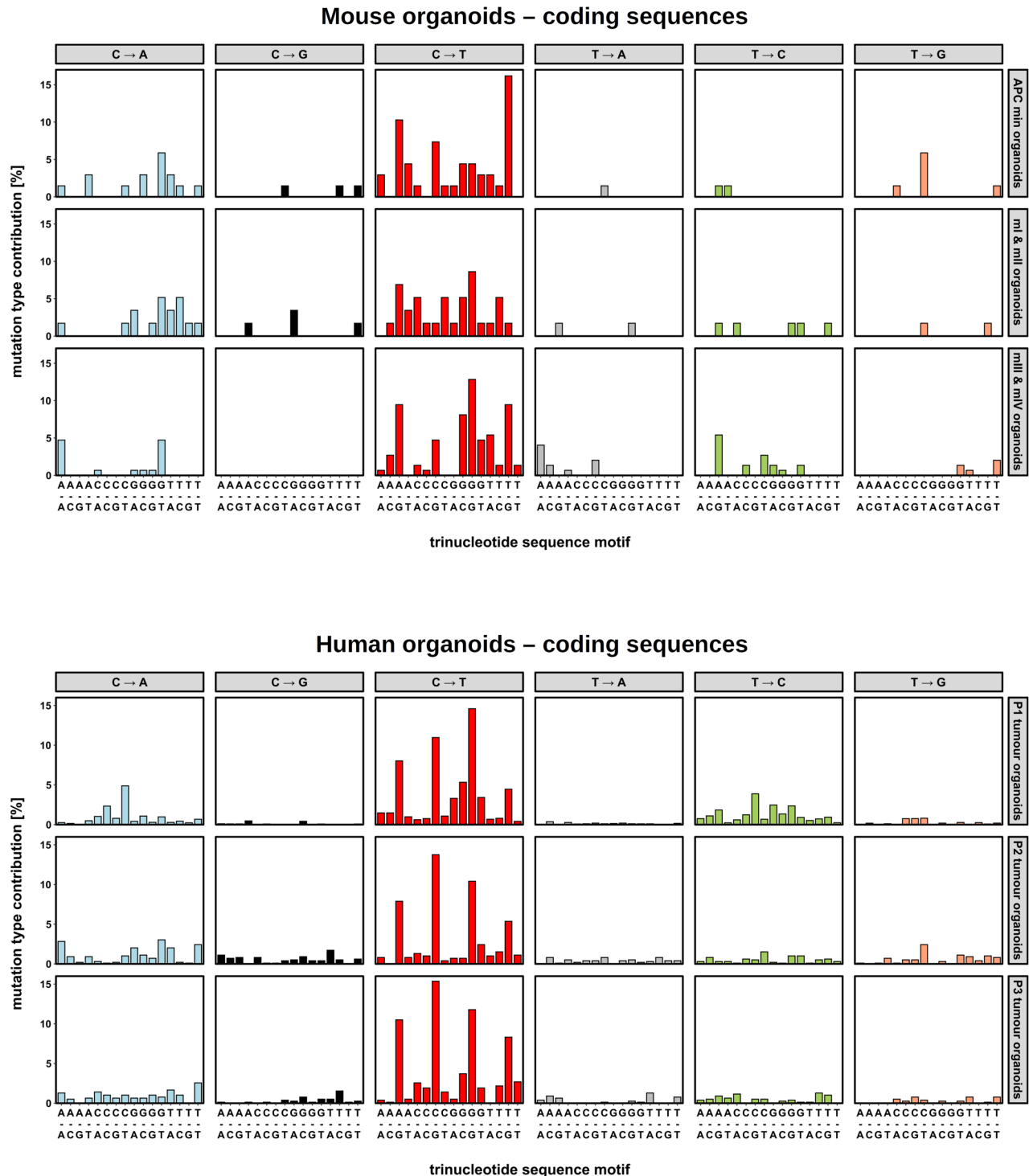


**Fig.S6 (Continued)**



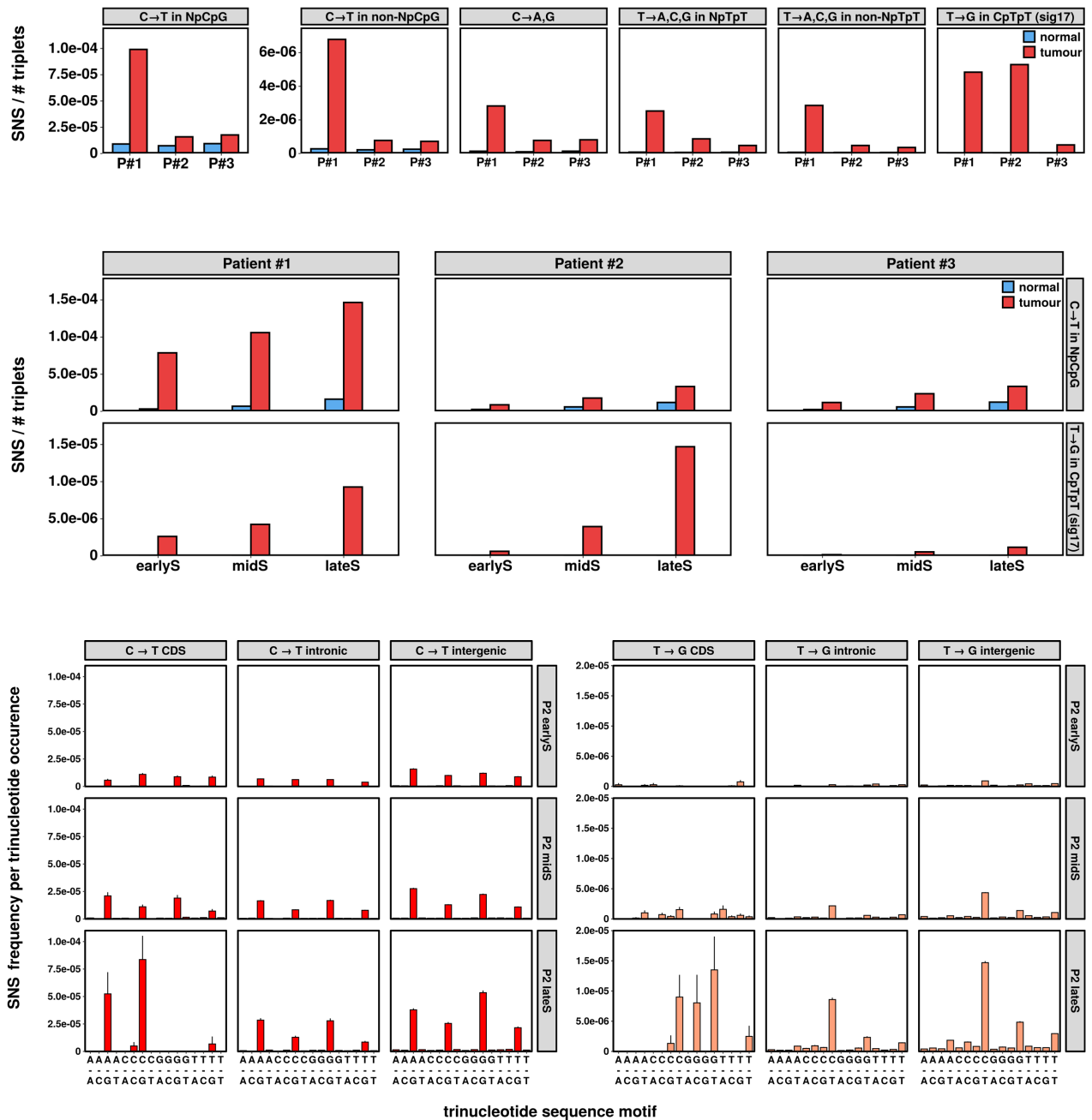
**Fig S6. Whole exome sequencing read depth of *Tp53* genomic locus for mouse I (a), mouse II (b), mouse III (c) and mouse IV(d).** Normal tissues and tumour organoids are represented by green and red colours, respectively. Primary tumours and metastatic tissues are illustrated by orange and yellow colours, respectively. Black arrows point to the deletion of exons 2-10. The genotype is indicated next to the sample name.

**Fig.S7**



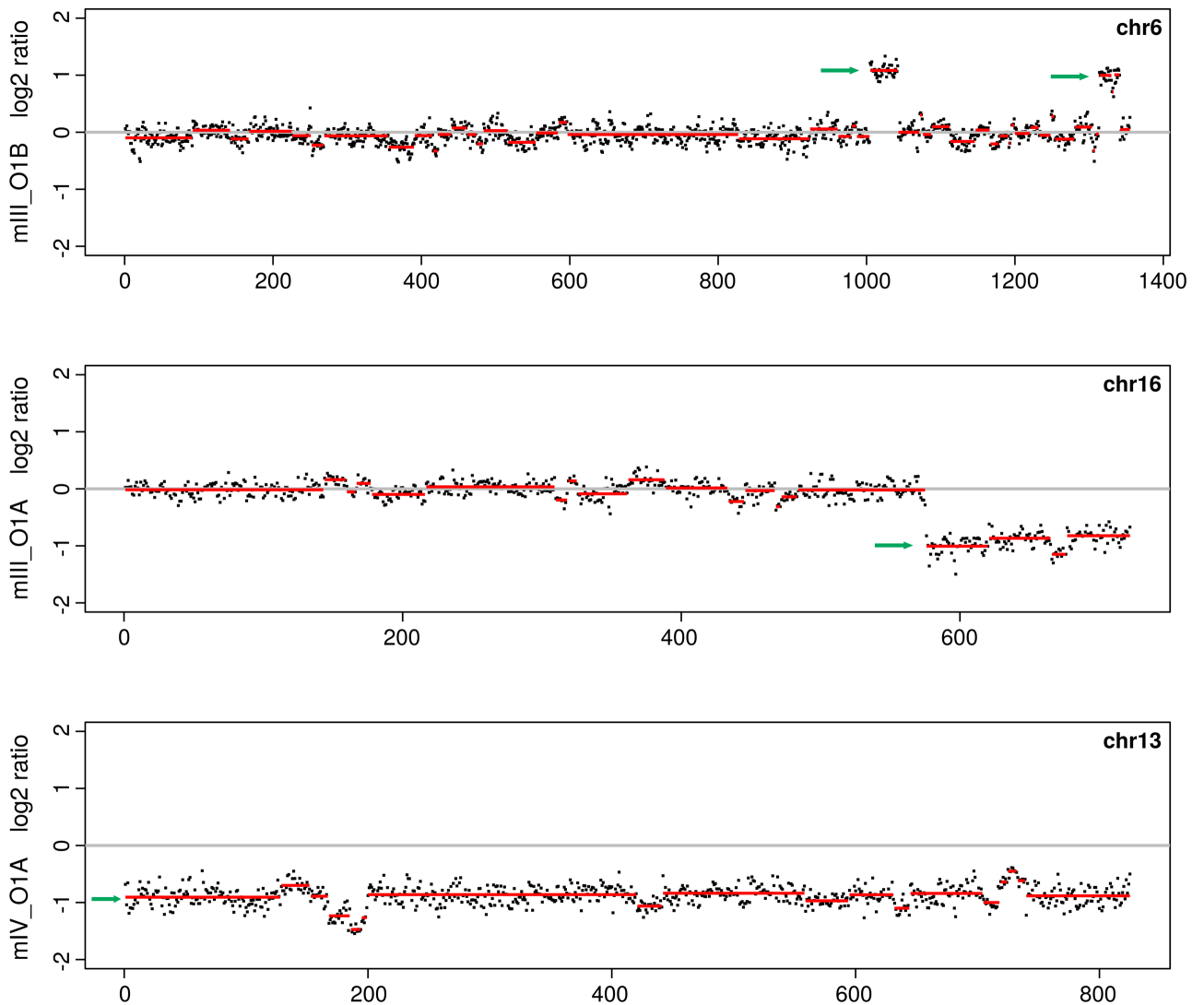
**Fig S7. Mutational signatures of somatic SNSs in tumour organoids.** a) Non-normalized signatures of protein-coding SNSs identified in the tumour organoids from mice I-IV of this study and from *Apc<sup>min</sup>* mice. b) Non-normalized signatures of protein-coding SNSs identified in the tumour organoids from three CRC patients.

**Fig.S8**



**Fig S8. Mutational signatures of somatic SNSs in human normal and tumour cell-derived organoids as a function of replication timing and gene annotation.** **a)** Normalized frequencies of all SNSs identified in the organoids from normal and tumour cells from three CRC patients. **b)** Normalized frequencies of SNSs conforming to signature 1 (C to T in NpCpG triplet) and to signature 17 (T to G in CpTpT triplet), according to replication timing. **c)** Normalized frequencies of C to T and T to G SNSs according to gene annotation and replication timing. The SNS frequencies were normalized according to the prevalence of the respective nucleotide triplets in the corresponding regions of the human genome.

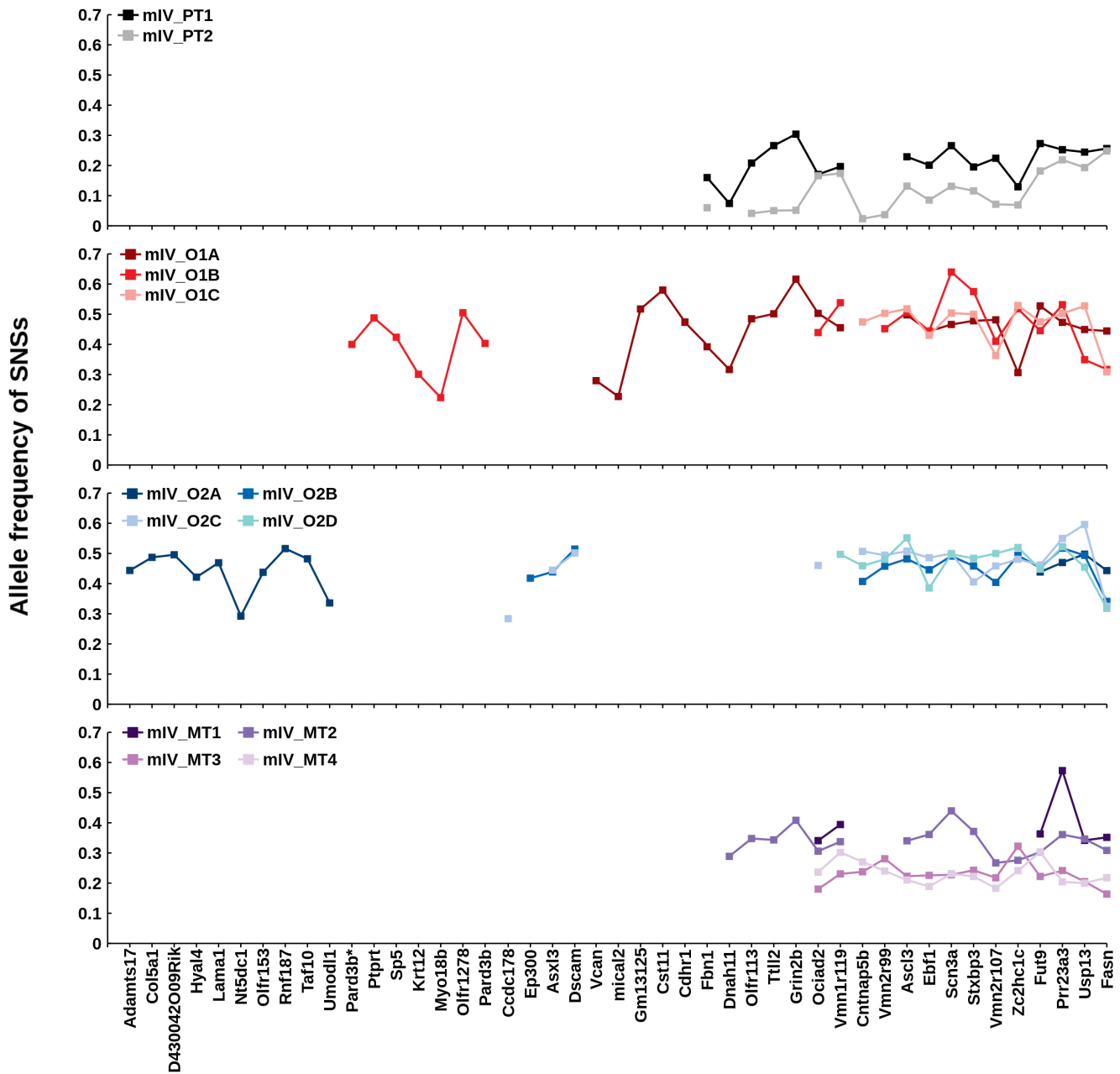
**Fig.S9**



**Fig S9. Examples of detected copy number alterations in mouse tumour organoids.** Green arrows indicate somatic copy number alteration events. In the example shown in the bottom panel, one copy of chromosome 13 has been lost.



**Fig.S10**



**Fig S10. Allele frequencies of identified somatic SNVs.** The allele frequencies of SNVs in primary tumour biopsies (PT1, PT2) varied, consistent with tumour heterogeneity, and did not exceed 30% (top panel). The SNVs identified in tumour organoids had allele frequencies close to 50%, apart from the SNVs that localized within genomic regions affected by CNA events (middle panels). Finally, the allele frequencies of the SNVs in the metastatic biopsies showed little variation, consistent with the metastases originating from one or more genetically identical cells (bottom panel).

## Chapter 3: Conclusions and outlook

### First part of the thesis.

Despite the generous efforts of many research labs and clinicians to find efficient treatment for CRC, it's still one of the deadliest cancers and leads to over 1 million deaths worldwide. Metastatic patients have the worst survival rates and new approaches to treat this disease are in urgent need. A high level of intra- and inter-tumor heterogeneity is one of the reasons for many drugs failure: the drug might be active in one cohort of patients, but others will not respond to it. Moreover, tumor cell plasticity and acquired resistance allow tumor cells to escape the drug action. The latter is driven by cancer stem cells, a small population of the most aggressive and resistant cells in the tumor. Additionally, the research shows that other cancer cell population - differentiated cells are less proliferative and are the first to vanish from the tumor site [Shimokawa M. et al., 2017].

Here we considered and implemented two conceptual ideas in an effort to find a treatment for CRC. **First**, we address the differentiation therapy approach to treat CRC by differentiating cancer stem cells to less aggressive differentiated-like cancer cells. **Second**, we established an RNA-seq-based platform that monitors organoid responses to drug treatments and allows: 1) discovery of cell type perturbations in wt organoids, 2) discovery of novel drug treatments inducing a particular differentiation phenotype both in wt and cancer organoids, 3) identification of drug resistance mechanism in cancer organoids. Further, this approach could be used in predicting drug combinations for targeting several signaling pathways involved in tumor escape from the treatment. Moreover, large collection of gene expression data of drug treated cancer organoid allows discovering novel connections between the pathways which might be implicated in resistance mechanisms.

Here we developed RNA-seq based high-throughput screening platform (TORNADO-seq) in organoids which allows monitoring the expression of large gene signatures and allows to quantify and evaluate complex cellular phenotypes and cellular alterations upon drug treatment. While the majority of drug screens are based on few parameters such as the activity of a single reporter or the measurement of cell viability, they lack rich data required for the analysis of complex phenotypes and the MOA of drugs. These limitations can be easily overcome with the use of TORNADO-seq (**T**argeted **ORgANOiD** sequencing). TORNADO-seq compares favourably with other high-content methods with respect to availability, cost, and applicability to organoid-based screens.

TORNADO-seq shows high-reproducibility, requires only 6h to perform, is cost-efficient (5 USD per sample including culture, library preparation, and sequencing cost), and doesn't require any specialized equipment. Due to the implemented UMI-counting and sequencing of amplicons, TORNADO-seq lacks PCR- and ligation-based bias and misdetections. Primer design is straightforward, needs minor optimization, and only requires moderate effort. To our knowledge, this is the first NGS-based HTS platform developed for organoid cultures.

Taken that, we applied this approach to identify differentiation-inducing drugs in wt and cancer organoids (APC<sup>lof</sup>;KRAS<sup>G12D</sup>;TP53<sup>lof</sup>). In wt organoids, we were able to identify drugs specifically enriching organoids for a certain cell type. In particular, lovastatin and trifluoperazine specifically induced EEC enrichment more precisely than previously proposed drug cocktails (Basak et al., 2017). More importantly, we identified many novel drug candidates targeting colon cancer organoids that were not described previously for CRC. Among them are antipsychotic phenothiazines, cholesterol-lowering statins, anti-mycotic conazoles, SERMs, glucocorticoids, and anti-histamines. Based on the obtained gene expression profiles we were able to propose and confirm MOAs for some of these drugs which we found to act by targeting the cholesterol pathway. Several of these drugs (ifenprodil, opi Pramol, perphenazine, toremifene) showed a beneficial therapeutic window targeting preferentially cancer compared to wt organoids. Finally, statins, opi Pramol and toremifene also targeted human CRC cell lines warranting future *in vivo* validation. Additional experiments with wider range of drug concentration, of course, could bring an additional layer of information to our understanding on the mechanism of action of identified drug hits.

Interestingly, many of the identified drugs targeting cancer organoids were acting via cholesterol-biosynthesis pathway. This is in line with high consumption of energy by cancer cells for fuelling their growth. Future research will have to find a way to target this pathway in patients, maybe using some locally delivered particles/or targeted radiation of certain nature which would allow to disrupt cholesterol accumulation in cancer cells. Interestingly, several observational studies showed improved survival in CRC patients taking statins as a medication with a primarily goal to lower cholesterol levels [Melloni C. et al., 2019; Poynter J.N. et al., 2005].

Colorectal cancer research is a fast-developing field and new approaches for CRC treatment are constantly emerging. So, recent application of patient-derived organoids in drug screenings and in predicting patient-specific responses to the drug treatments may be a promising tool for fighting CRC. We believe that the implementation of RNA-seq

based screenings such as TORNADO-seq in combination with PDO technology will allow the identification of resistance mechanisms in tumors. Altogether, the combined application of those technologies will definitely improve our understanding of this fatal disease.

## **Second part of the thesis.**

Accumulation of mutations in tissues occurs constantly. Genomic instability (single nucleotide substitutions (SNSs), copy number alterations (CNAs)) is a hallmark of cancer. Acquiring of mutations eventually will lead to the development of malignant tumors, including colorectal cancer. It has been reported that mutation acquisition rate in CRC tumors is much higher than in normal cells [Lugli N. et al., 2017]. Even though the process of mutation acquisition is studied well, the contribution of driver genes to this process in CRC is not fully elucidated. According to the tumor development theory by Fearon and Vogelstein [Fearon, E. R. et al., 1990] the acquiring of mutations (SNSs) in tumors appear in the following sequence – APC, KRAS, P53, before mutations altering TGF- $\beta$  and PI3K/AKT signaling, such as SMAD4 and PIK3CA, appear leading to enhanced chemoresistance and metastatic properties of tumor cells. Further, copy number alterations (CNAs) develop in more aggressive tumors with the time and can in the same way as SNSs affect the expression of important genes and driver the tumor progression.

Here we combined single-cell derived tumor organoids and exome sequencing to elucidate the role of p53 and KRAS driver genes. We used either APC<sup>lof</sup> (A) APC<sup>lof</sup>;KRAS<sup>G12D</sup> (AK) or APC<sup>lof</sup>;KRAS<sup>G12D</sup>;TP53<sup>lof</sup> (AKP) mutation genotypes that reflect early stages of CRC development. Using low-dose tamoxifen injections and Cre<sup>ERT2</sup> inducible system we generated colonic tumors developing over long period of time (more than 5-6 months). Some of those tumors (AKP) developed as carcinomas and gave metastasis to the liver. We saw a high increase in the number of acquired mutations in tumors which developed over long period of time compared to short lived tumors (2 weeks). But the rate of mutation acquisition was not affected by the number of cancer-driver genes (A, AK or AKP). At the same time, CNAs were present almost exclusively in tumors harboring TP53 mutations, which was also shown by other studies. Additional *ex vivo* experiments on proliferation rate of organoids showed that the difference in SNSs present in non-transformed and transformed cells cannot be explained by differences in proliferation rates. Limited sample size in this study of course is a weakness and can't fully exclude the possibility that timing is also important for CNAs alongside TP53 mutation.

We showed that tumors harboring AKP phenotype are able, over long period of time (6 months) to progress to carcinoma stage and give metastasis to distant organs (liver). Surprisingly, those tumors didn't have much more mutations than tumors which were only 10 days old. At the same time "old" tumors had other changes in genomic instability such as CNAs, sometime with the disappearance of a bigger part of the whole chromosome, telling us that the evolving of the tumor on later stages of its development is most likely to be regulated via CNAs rather than SNSs. Further, we were able to track the evolution of the tumor clones in the mouse which had metastasis and found that those separated metastatic lesions had both common and unique mutations. Interestingly, we also able to detect further chromosomal alterations in some of the distant metastasis indicating further tumor adaptation and evolution in new microenvironment. Taken together, in the model of early colorectal cancer development the influence of KRAS mutation for genomic instability is minimal, at the same time TP53 affect genomic instability via CNAs.

## Chapter 4: Bibliography

Artis, David, et al. "RELMbeta/FIZZ2 Is a Goblet Cell-Specific Immune-Effector Molecule in the Gastrointestinal Tract." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 37, Sept. 2004, pp. 13596–600. PubMed, <https://doi.org/10.1073/pnas.0404034101>.

Baker, Ann-Marie, et al. "Characterization of LGR5 Stem Cells in Colorectal Adenomas and Carcinomas." *Scientific Reports*, vol. 5, no. 1, Aug. 2015, p. 8654. DOI.org (Crossref), <https://doi.org/10.1038/srep08654>.

Barcellos-Hoff, M. H., et al. "Functional Differentiation and Alveolar Morphogenesis of Primary Mammary Cultures on Reconstituted Basement Membrane." *Development (Cambridge, England)*, vol. 105, no. 2, Feb. 1989, pp. 223–35.

Barker, Nick, et al. "Identification of Stem Cells in Small Intestine and Colon by Marker Gene Lgr5." *Nature*, vol. 449, no. 7165, Oct. 2007, pp. 1003–07. DOI.org (Crossref), <https://doi.org/10.1038/nature06196>.

Barker Nick. "Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration." *Nature Reviews Molecular Cell Biology*, volume 15, pages 19–33, 2014. <https://doi.org/10.1038/nrm3721>

Barriga, Francisco M., et al. "Mex3a Marks a Slowly Dividing Subpopulation of Lgr5+ Intestinal Stem Cells." *Cell Stem Cell*, vol. 20, no. 6, June 2017, pp. 801-816.e7. PubMed, <https://doi.org/10.1016/j.stem.2017.02.007>.

Basak, Onur, et al. "Induced Quiescence of Lgr5+ Stem Cells in Intestinal Organoids Enables Differentiation of Hormone-Producing Enteroendocrine Cells." *Cell Stem Cell*, vol. 20, no. 2, Feb. 2017, pp. 177-190.e4. DOI.org (Crossref), <https://doi.org/10.1016/j.stem.2016.11.001>.

Biller, Leah H., and Deborah Schrag. "Diagnosis and Treatment of Metastatic Colorectal Cancer: A Review." *JAMA*, vol. 325, no. 7, Feb. 2021, p. 669. DOI.org (Crossref), <https://doi.org/10.1001/jama.2021.0106>.

Bissell, M. J., et al. "Microenvironmental Regulators of Tissue Structure and Function Also Regulate Tumor Induction and Progression: The Role of Extracellular Matrix and Its Degrading Enzymes." *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 70, no. 0, Jan. 2005, pp. 343–56. DOI.org (Crossref), <https://doi.org/10.1101/sqb.2005.70.013>.

Biton, Moshe, et al. "Epithelial MicroRNAs Regulate Gut Mucosal Immunity via Epithelium-T Cell Crosstalk." *Nature Immunology*, vol. 12, no. 3, Mar. 2011, pp. 239–46. PubMed, <https://doi.org/10.1038/ni.1994>.

Boehnke, Karsten, et al. "Assay Establishment and Validation of a High-Throughput Screening Platform for Three-Dimensional Patient-Derived Colon Cancer Organoid Cultures." *Journal of Biomolecular Screening*, vol. 21, no. 9, Oct. 2016, pp. 931–41. PubMed, <https://doi.org/10.1177/1087057116650965>.

Brandenberg, Nathalie, et al. "High-Throughput Automated Organoid Culture via Stem-Cell Aggregation in Microcavity Arrays." *Nature Biomedical Engineering*, vol. 4, no. 9, Sept. 2020, pp. 863–74. PubMed, <https://doi.org/10.1038/s41551-020-0565-2>.

Brattain, M. G., et al. "Heterogeneity of Human Colon Carcinoma." *Cancer Metastasis Reviews*, vol. 3, no. 3, 1984, pp. 177–91. PubMed, <https://doi.org/10.1007/BF00048384>.

Bruun, Jarle, et al. "Patient-Derived Organoids from Multiple Colorectal Cancer Liver Metastases Reveal Moderate Intra-Patient Pharmacotranscriptomic Heterogeneity." *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, vol. 26, no. 15, Aug. 2020, pp. 4107–19. PubMed, <https://doi.org/10.1158/1078-0432.CCR-19-3637>.

Buczacki, Simon J. A., et al. "Intestinal Label-Retaining Cells Are Secretory Precursors Expressing Lgr5." *Nature*, vol. 495, no. 7439, Mar. 2013, pp. 65–69. PubMed, <https://doi.org/10.1038/nature11965>.

Buikhuisen, Joyce Y., et al. "Exploring and Modelling Colon Cancer Inter-Tumour Heterogeneity: Opportunities and Challenges." *Oncogenesis*, vol. 9, no. 7, July 2020, p. 66. PubMed, <https://doi.org/10.1038/s41389-020-00250-6>.

Bürtin, Florian, et al. "Mouse Models of Colorectal Cancer: Past, Present and Future Perspectives." *World Journal of Gastroenterology*, vol. 26, no. 13, Apr. 2020, pp. 1394–426. PubMed, <https://doi.org/10.3748/wjg.v26.i13.1394>.

Bush, Erin C., et al. "PLATE-Seq for Genome-Wide Regulatory Network Analysis of High-Throughput Screens." *Nature Communications*, vol. 8, no. 1, Dec. 2017, p. 105. DOI.org (Crossref), <https://doi.org/10.1038/s41467-017-00136-z>.

Carragher, Linda A. S., et al. "V600EBraf Induces Gastrointestinal Crypt Senescence and Promotes Tumour Progression through Enhanced CpG Methylation of P16INK4a." *EMBO Molecular Medicine*, vol. 2, no. 11, Nov. 2010, pp. 458–71. PubMed, <https://doi.org/10.1002/emmm.201000099>.

Cho, Yong Beom, et al. "Colorectal Cancer Patient-Derived Xenografted Tumors Maintain Characteristic Features of the Original Tumors." *The Journal of Surgical Research*, vol. 187, no. 2, Apr. 2014, pp. 502–09. PubMed, <https://doi.org/10.1016/j.jss.2013.11.010>.

Clevers, Hans. "The Intestinal Crypt, a Prototype Stem Cell Compartment." *Cell*, vol. 154, no. 2, July 2013, pp. 274–84. PubMed, <https://doi.org/10.1016/j.cell.2013.07.004>.

Coffee, Erin M., et al. "Concomitant BRAF and PI3K/MTOR Blockade Is Required for Effective Treatment of BRAF(V600E) Colorectal Cancer." *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, vol. 19, no. 10, May 2013, pp. 2688–98. PubMed, <https://doi.org/10.1158/1078-0432.CCR-12-2556>.

Crosnier, Cécile, et al. "Organizing Cell Renewal in the Intestine: Stem Cells, Signals and Combinatorial Control." *Nature Reviews Genetics*, vol. 7, no. 5, May 2006, pp. 349–59. DOI.org (Crossref), <https://doi.org/10.1038/nrg1840>.

Cunningham, David, et al. "Cetuximab Monotherapy and Cetuximab plus Irinotecan in Irinotecan-Refractory Metastatic Colorectal Cancer." *The New England Journal of*

Medicine, vol. 351, no. 4, July 2004, pp. 337–45. PubMed, <https://doi.org/10.1056/NEJMoa033025>.

De Roock, W., et al. “KRAS Wild-Type State Predicts Survival and Is Associated to Early Radiological Response in Metastatic Colorectal Cancer Treated with Cetuximab.” *Annals of Oncology: Official Journal of the European Society for Medical Oncology*, vol. 19, no. 3, Mar. 2008, pp. 508–15. PubMed, <https://doi.org/10.1093/annonc/mdm496>.

de Santa Barbara, P., et al. “Development and Differentiation of the Intestinal Epithelium.” *Cellular and Molecular Life Sciences: CMLS*, vol. 60, no. 7, July 2003, pp. 1322–32. PubMed, <https://doi.org/10.1007/s00018-003-2289-3>.

de Sousa e Melo, Felipe, et al. “A Distinct Role for Lgr5+ Stem Cells in Primary and Metastatic Colon Cancer.” *Nature*, vol. 543, no. 7647, Mar. 2017, pp. 676–80. PubMed, <https://doi.org/10.1038/nature21713>.

De Sousa E Melo, Felipe, et al. “Poor-Prognosis Colon Cancer Is Defined by a Molecularly Distinct Subtype and Develops from Serrated Precursor Lesions.” *Nature Medicine*, vol. 19, no. 5, May 2013, pp. 614–18. PubMed, <https://doi.org/10.1038/nm.3174>.

Del Rio, M., et al. “Molecular Subtypes of Metastatic Colorectal Cancer Are Associated with Patient Response to Irinotecan-Based Therapies.” *European Journal of Cancer (Oxford, England: 1990)*, vol. 76, May 2017, pp. 68–75. PubMed, <https://doi.org/10.1016/j.ejca.2017.02.003>.

Di Fiore, F., et al. “Clinical Relevance of KRAS Mutation Detection in Metastatic Colorectal Cancer Treated by Cetuximab plus Chemotherapy.” *British Journal of Cancer*, vol. 96, no. 8, Apr. 2007, pp. 1166–69. PubMed, <https://doi.org/10.1038/sj.bjc.6603685>.

Dionellis, Vasilis S., et al. “Genomic Instability Profiles at the Single Cell Level in Mouse Colorectal Cancers of Defined Genotypes.” *Cancers*, vol. 13, no. 6, Mar. 2021, p. 1267. PubMed, <https://doi.org/10.3390/cancers13061267>.

Douillard, J. Y., et al. “Final Results from PRIME: Randomized Phase III Study of Panitumumab with FOLFOX4 for First-Line Treatment of Metastatic Colorectal Cancer.” *Annals of Oncology: Official Journal of the European Society for Medical Oncology*, vol. 25, no. 7, July 2014, pp. 1346–55. PubMed, <https://doi.org/10.1093/annonc/mdu141>.

Drost, Jarno, et al. “Sequential Cancer Mutations in Cultured Human Intestinal Stem Cells.” *Nature*, vol. 521, no. 7550, May 2015, pp. 43–47. PubMed, <https://doi.org/10.1038/nature14415>.

Du, Yuhong, et al. “Development of a Miniaturized 3D Organoid Culture Platform for Ultra-High-Throughput Screening.” *Journal of Molecular Cell Biology*, vol. 12, no. 8, Aug. 2020, pp. 630–43. PubMed, <https://doi.org/10.1093/jmcb/mjaa036>.

Egerod, Kristoffer L., et al. “A Major Lineage of Enteroendocrine Cells Coexpress CCK, Secretin, GIP, GLP-1, PYY, and Neurotensin but Not Somatostatin.” *Endocrinology*, vol. 153, no. 12, Dec. 2012, pp. 5782–95. PubMed, <https://doi.org/10.1210/en.2012-1595>.



el Marjou, Fatima, et al. "Tissue-Specific and Inducible Cre-Mediated Recombination in the Gut Epithelium." *Genesis (New York, N.Y.:* 2000), vol. 39, no. 3, July 2004, pp. 186–93. PubMed, <https://doi.org/10.1002/gene.20042>.

Enquist, Ida B., et al. "Lymph Node-Independent Liver Metastasis in a Model of Metastatic Colorectal Cancer." *Nature Communications*, vol. 5, no. 1, May 2014, p. 3530. DOI.org (Crossref), <https://doi.org/10.1038/ncomms4530>.

Fearon, E. R., and B. Vogelstein. "A Genetic Model for Colorectal Tumorigenesis." *Cell*, vol. 61, no. 5, June 1990, pp. 759–67. PubMed, [https://doi.org/10.1016/0092-8674\(90\)90186-i](https://doi.org/10.1016/0092-8674(90)90186-i).

Feng, Ying, et al. "Sox9 Induction, Ectopic Paneth Cells, and Mitotic Spindle Axis Defects in Mouse Colon Adenomatous Epithelium Arising from Conditional Biallelic Apc Inactivation." *The American Journal of Pathology*, vol. 183, no. 2, Aug. 2013, pp. 493–503. PubMed, <https://doi.org/10.1016/j.ajpath.2013.04.013>.

Folkesson, Evelina, et al. "High-Throughput Screening Reveals Higher Synergistic Effect of MEK Inhibitor Combinations in Colon Cancer Spheroids." *Scientific Reports*, vol. 10, no. 1, July 2020, p. 11574. PubMed, <https://doi.org/10.1038/s41598-020-68441-0>.

Fujii, Masayuki, et al. "A Colorectal Tumor Organoid Library Demonstrates Progressive Loss of Niche Factor Requirements during Tumorigenesis." *Cell Stem Cell*, vol. 18, no. 6, June 2016, pp. 827–38. PubMed, <https://doi.org/10.1016/j.stem.2016.04.003>.

Fumagalli, Arianna, et al. "Plasticity of Lgr5-Negative Cancer Cells Drives Metastasis in Colorectal Cancer." *Cell Stem Cell*, vol. 26, no. 4, Apr. 2020, pp. 569-578.e7. PubMed, <https://doi.org/10.1016/j.stem.2020.02.008>.

Gassler, N., et al. "Molecular Characterisation of Non-Absorptive and Absorptive Enterocytes in Human Small Intestine." *Gut*, vol. 55, no. 8, Aug. 2006, pp. 1084–89. PubMed, <https://doi.org/10.1136/gut.2005.073262>.

Gehart, Helmuth, et al. "Identification of Enteroendocrine Regulators by Real-Time Single-Cell Differentiation Mapping." *Cell*, vol. 176, no. 5, Feb. 2019, pp. 1158-1173.e16. PubMed, <https://doi.org/10.1016/j.cell.2018.12.029>.

Gehart, Helmuth, and Hans Clevers. "Tales from the Crypt: New Insights into Intestinal Stem Cells." *Nature Reviews Gastroenterology & Hepatology*, vol. 16, no. 1, Jan. 2019, pp. 19–34. DOI.org (Crossref), <https://doi.org/10.1038/s41575-018-0081-y>.

Ghoos, Y., and G. Vantrappen. "The Cytochemical Localization of Lysozyme in Paneth Cell Granules." *The Histochemical Journal*, vol. 3, no. 3, May 1971, pp. 175–78. PubMed, <https://doi.org/10.1007/BF01002560>.

Gocek, Elzbieta, and Ewa Marcinkowska. "Differentiation Therapy of Acute Myeloid Leukemia." *Cancers*, vol. 3, no. 2, May 2011, pp. 2402–20. PubMed, <https://doi.org/10.3390/cancers3022402>.

Grothey, Axel, et al. "Regorafenib Monotherapy for Previously Treated Metastatic Colorectal Cancer (CORRECT): An International, Multicentre, Randomised, Placebo-

Controlled, Phase 3 Trial.” *The Lancet*, vol. 381, no. 9863, Jan. 2013, pp. 303–12. DOI.org (Crossref), [https://doi.org/10.1016/S0140-6736\(12\)61900-X](https://doi.org/10.1016/S0140-6736(12)61900-X).

Grün, Dominic, Mauro J. Muraro, et al. “De Novo Prediction of Stem Cell Identity Using Single-Cell Transcriptome Data.” *Cell Stem Cell*, vol. 19, no. 2, Aug. 2016, pp. 266–77. PubMed, <https://doi.org/10.1016/j.stem.2016.05.010>.

Grün, Dominic, Anna Lyubimova, et al. “Single-Cell Messenger RNA Sequencing Reveals Rare Intestinal Cell Types.” *Nature*, vol. 525, no. 7568, Sept. 2015, pp. 251–55. DOI.org (Crossref), <https://doi.org/10.1038/nature14966>.

Guinney, Justin, et al. “The Consensus Molecular Subtypes of Colorectal Cancer.” *Nature Medicine*, vol. 21, no. 11, Nov. 2015, pp. 1350–56. PubMed, <https://doi.org/10.1038/nm.3967>.

Gustavsson, Bengt, et al. “A Review of the Evolution of Systemic Chemotherapy in the Management of Colorectal Cancer.” *Clinical Colorectal Cancer*, vol. 14, no. 1, Mar. 2015, pp. 1–10. PubMed, <https://doi.org/10.1016/j.clcc.2014.11.002>.

Haber, Adam L., et al. “A Single-Cell Survey of the Small Intestinal Epithelium.” *Nature*, vol. 551, no. 7680, Nov. 2017, pp. 333–39. DOI.org (Crossref), <https://doi.org/10.1038/nature24489>.

Habib, Abdella M., et al. “Overlap of Endocrine Hormone Expression in the Mouse Intestine Revealed by Transcriptional Profiling and Flow Cytometry.” *Endocrinology*, vol. 153, no. 7, July 2012, pp. 3054–65. PubMed, <https://doi.org/10.1210/en.2011-2170>.

Hansen, L. A., et al. “Retinoids in Chemoprevention and Differentiation Therapy.” *Carcinogenesis*, vol. 21, no. 7, July 2000, pp. 1271–79.

Henrikson, Nora B., et al. “Family History and the Natural History of Colorectal Cancer: Systematic Review.” *Genetics in Medicine*, vol. 17, no. 9, Sept. 2015, pp. 702–12. DOI.org (Crossref), <https://doi.org/10.1038/gim.2014.188>.

Hinoi, Takao, et al. “Mouse Model of Colonic Adenoma-Carcinoma Progression Based on Somatic Apc Inactivation.” *Cancer Research*, vol. 67, no. 20, Oct. 2007, pp. 9721–30. PubMed, <https://doi.org/10.1158/0008-5472.CAN-07-2735>.

Howitt, Michael R., et al. “Tuft Cells, Taste-Chemosensory Cells, Orchestrate Parasite Type 2 Immunity in the Gut.” *Science (New York, N.Y.)*, vol. 351, no. 6279, Mar. 2016, pp. 1329–33. PubMed, <https://doi.org/10.1126/science.aaf1648>.

Hurwitz, Herbert, et al. “Bevacizumab plus Irinotecan, Fluorouracil, and Leucovorin for Metastatic Colorectal Cancer.” *The New England Journal of Medicine*, vol. 350, no. 23, June 2004, pp. 2335–42. PubMed, <https://doi.org/10.1056/NEJMoa032691>.

Indra, A. K., et al. “Temporally-Controlled Site-Specific Mutagenesis in the Basal Layer of the Epidermis: Comparison of the Recombinase Activity of the Tamoxifen-Inducible Cre-ER(T) and Cre-ER(T2) Recombinases.” *Nucleic Acids Research*, vol. 27, no. 22, Nov. 1999, pp. 4324–27. PubMed, <https://doi.org/10.1093/nar/27.22.4324>.

Itzkovitz, Shalev, et al. "Single-Molecule Transcript Counting of Stem-Cell Markers in the Mouse Intestine." *Nature Cell Biology*, vol. 14, no. 1, Nov. 2011, pp. 106–14. PubMed, <https://doi.org/10.1038/ncb2384>.

Jadhav, Unmesh, et al. "Dynamic Reorganization of Chromatin Accessibility Signatures during Dedifferentiation of Secretory Precursors into Lgr5+ Intestinal Stem Cells." *Cell Stem Cell*, vol. 21, no. 1, July 2017, pp. 65–77.e5. PubMed, <https://doi.org/10.1016/j.stem.2017.05.001>.

Jasperson, Kory W., et al. "Hereditary and Familial Colon Cancer." *Gastroenterology*, vol. 138, no. 6, June 2010, pp. 2044–58. PubMed, <https://doi.org/10.1053/j.gastro.2010.01.054>.

Kazakevych, Juri, et al. "Dynamic Changes in Chromatin States during Specification and Differentiation of Adult Intestinal Stem Cells." *Nucleic Acids Research*, vol. 45, no. 10, June 2017, pp. 5770–84. PubMed, <https://doi.org/10.1093/nar/gkx167>.

Kim, Eun Ran, and Dong Kyung Chang. "Colorectal Cancer in Inflammatory Bowel Disease: The Risk, Pathogenesis, Prevention and Diagnosis." *World Journal of Gastroenterology*, vol. 20, no. 29, Aug. 2014, pp. 9872–81. PubMed, <https://doi.org/10.3748/wjg.v20.i29.9872>.

Kopetz, Scott, et al. "Encorafenib, Binimetinib, and Cetuximab in BRAF V600E-Mutated Colorectal Cancer." *The New England Journal of Medicine*, vol. 381, no. 17, Oct. 2019, pp. 1632–43. PubMed, <https://doi.org/10.1056/NEJMoa1908075>.

Kuipers, Ernst J., et al. "Colorectal Cancer." *Nature Reviews Disease Primers*, vol. 1, no. 1, Dec. 2015, p. 15065. DOI.org (Crossref), <https://doi.org/10.1038/nrdp.2015.65>.

Lakso, M., et al. "Targeted Oncogene Activation by Site-Specific Recombination in Transgenic Mice." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 14, July 1992, pp. 6232–36. PubMed, <https://doi.org/10.1073/pnas.89.14.6232>.

Lasorella, Anna, et al. "The ID Proteins: Master Regulators of Cancer Stem Cells and Tumour Aggressiveness." *Nature Reviews Cancer*, vol. 14, no. 2, Feb. 2014, pp. 77–91. DOI.org (Crossref), <https://doi.org/10.1038/nrc3638>.

Li, Huipeng, et al. "Reference Component Analysis of Single-Cell Transcriptomes Elucidates Cellular Heterogeneity in Human Colorectal Tumors." *Nature Genetics*, vol. 49, no. 5, May 2017, pp. 708–18. PubMed, <https://doi.org/10.1038/ng.3818>.

Li, Ning, et al. "Single-Cell Analysis of Proxy Reporter Allele-Marked Epithelial Cells Establishes Intestinal Stem Cell Hierarchy." *Stem Cell Reports*, vol. 3, no. 5, Nov. 2014, pp. 876–91. PubMed, <https://doi.org/10.1016/j.stemcr.2014.09.011>.

Lièvre, Astrid, et al. "KRAS Mutation Status Is Predictive of Response to Cetuximab Therapy in Colorectal Cancer." *Cancer Research*, vol. 66, no. 8, Apr. 2006, pp. 3992–95. PubMed, <https://doi.org/10.1158/0008-5472.CAN-06-0191>.

Lukonin, Ilya, et al. "Phenotypic Landscape of Intestinal Organoid Regeneration." *Nature*, vol. 586, no. 7828, Oct. 2020, pp. 275–80. PubMed, <https://doi.org/10.1038/s41586-020-2776-9>.

Madison, Blair B., et al. "Cis Elements of the Villin Gene Control Expression in Restricted Domains of the Vertical (Crypt) and Horizontal (Duodenum, Cecum) Axes of the Intestine." *The Journal of Biological Chemistry*, vol. 277, no. 36, Sept. 2002, pp. 33275–83. PubMed, <https://doi.org/10.1074/jbc.M204935200>.

Marisa, Laetitia, et al. "Gene Expression Classification of Colon Cancer into Molecular Subtypes: Characterization, Validation, and Prognostic Value." *PLoS Medicine*, vol. 10, no. 5, 2013, p. e1001453. PubMed, <https://doi.org/10.1371/journal.pmed.1001453>.

Martin, Eric S., et al. "Development of a Colon Cancer GEMM-Derived Orthotopic Transplant Model for Drug Discovery and Validation." *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, vol. 19, no. 11, June 2013, pp. 2929–40. PubMed, <https://doi.org/10.1158/1078-0432.CCR-12-2307>.

Melloni C. et al., "Abstract 15660: Statin Use is Associated With Increased Overall Survival in Patients With Colorectal Cancer: Findings From a Cohort of 29,498 United States Veterans". Nov 2019, *Circulation*, 2019, 140:A15660.

Mendelaar, Pauline A. J., et al. "Whole Genome Sequencing of Metastatic Colorectal Cancer Reveals Prior Treatment Effects and Specific Metastasis Features." *Nature Communications*, vol. 12, no. 1, Jan. 2021, p. 574. PubMed, <https://doi.org/10.1038/s41467-020-20887-6>.

Michels, Birgitta E., et al. "Pooled In Vitro and In Vivo CRISPR-Cas9 Screening Identifies Tumor Suppressors in Human Colon Organoids." *Cell Stem Cell*, vol. 26, no. 5, May 2020, pp. 782-792.e7. PubMed, <https://doi.org/10.1016/j.stem.2020.04.003>.

Miyamoto, Shingo, et al. "Novel Screening System Revealed That Intracellular Cholesterol Trafficking Can Be a Good Target for Colon Cancer Prevention." *Scientific Reports*, vol. 9, no. 1, Apr. 2019, p. 6192. PubMed, <https://doi.org/10.1038/s41598-019-42363-y>.

Montgomery, Robert K., et al. "Mouse Telomerase Reverse Transcriptase (MTert) Expression Marks Slowly Cycling Intestinal Stem Cells." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 1, Jan. 2011, pp. 179–84. PubMed, <https://doi.org/10.1073/pnas.1013004108>.

Moor, Andreas E., et al. "Spatial Reconstruction of Single Enterocytes Uncovers Broad Zonation along the Intestinal Villus Axis." *Cell*, vol. 175, no. 4, Nov. 2018, pp. 1156-1167.e15. PubMed, <https://doi.org/10.1016/j.cell.2018.08.063>.

Moser, A. R., et al. "A Dominant Mutation That Predisposes to Multiple Intestinal Neoplasia in the Mouse." *Science (New York, N.Y.)*, vol. 247, no. 4940, Jan. 1990, pp. 322–24. PubMed, <https://doi.org/10.1126/science.2296722>.

Mullins, Christina S., et al. "Integrated Biobanking and Tumor Model Establishment of Human Colorectal Carcinoma Provides Excellent Tools for Preclinical Research."

Cancers, vol. 11, no. 10, Oct. 2019, p. E1520. PubMed, <https://doi.org/10.3390/cancers11101520>.

Muñoz, Javier, et al. "The Lgr5 Intestinal Stem Cell Signature: Robust Expression of Proposed Quiescent '+4' Cell Markers." *The EMBO Journal*, vol. 31, no. 14, June 2012, pp. 3079–91. PubMed, <https://doi.org/10.1038/emboj.2012.166>.

Narasimhan, Vignesh, et al. "Medium-Throughput Drug Screening of Patient-Derived Organoids from Colorectal Peritoneal Metastases to Direct Personalized Therapy." *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, vol. 26, no. 14, July 2020, pp. 3662–70. PubMed, <https://doi.org/10.1158/1078-0432.CCR-20-0073>.

Nasr, Rihab, et al. "Eradication of Acute Promyelocytic Leukemia-Initiating Cells through PML-RARA Degradation." *Nature Medicine*, vol. 14, no. 12, Dec. 2008, pp. 1333–42. PubMed, <https://doi.org/10.1038/nm.1891>.

Nigmatullina, Lira, et al. "Id2 Controls Specification of Lgr5 + Intestinal Stem Cell Progenitors during Gut Development." *The EMBO Journal*, vol. 36, no. 7, Apr. 2017, pp. 869–85. DOI.org (Crossref), <https://doi.org/10.15252/embj.201694959>.

Noah, Taeko K., et al. "SAM Pointed Domain ETS Factor (SPDEF) Regulates Terminal Differentiation and Maturation of Intestinal Goblet Cells." *Experimental Cell Research*, vol. 316, no. 3, Feb. 2010, pp. 452–65. PubMed, <https://doi.org/10.1016/j.yexcr.2009.09.020>.

Okita, Akira, et al. "Consensus Molecular Subtypes Classification of Colorectal Cancer as a Predictive Factor for Chemotherapeutic Efficacy against Metastatic Colorectal Cancer." *Oncotarget*, vol. 9, no. 27, Apr. 2018, pp. 18698–711. PubMed, <https://doi.org/10.18632/oncotarget.24617>.

Ooft, Salo N., et al. "Patient-Derived Organoids Can Predict Response to Chemotherapy in Metastatic Colorectal Cancer Patients." *Science Translational Medicine*, vol. 11, no. 513, Oct. 2019, p. eaay2574. PubMed, <https://doi.org/10.1126/scitranslmed.aay2574>.

Ordóñez-Morán, Paloma, et al. "HOXA5 Counteracts Stem Cell Traits by Inhibiting Wnt Signaling in Colorectal Cancer." *Cancer Cell*, vol. 28, no. 6, Dec. 2015, pp. 815–29. PubMed, <https://doi.org/10.1016/j.ccell.2015.11.001>.

Pellegrinet, Luca, et al. "DII1- and DII4-Mediated Notch Signaling Are Required for Homeostasis of Intestinal Stem Cells." *Gastroenterology*, vol. 140, no. 4, Apr. 2011, pp. 1230-1240.e1-7. PubMed, <https://doi.org/10.1053/j.gastro.2011.01.005>.

Petersen, O. W., et al. "Interaction with Basement Membrane Serves to Rapidly Distinguish Growth and Differentiation Pattern of Normal and Malignant Human Breast Epithelial Cells." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 19, Oct. 1992, pp. 9064–68. PubMed, <https://doi.org/10.1073/pnas.89.19.9064>.

Plotnikov, Alexander, et al. "PRMT1 Inhibition Induces Differentiation of Colon Cancer Cells." *Scientific Reports*, vol. 10, no. 1, Nov. 2020, p. 20030. PubMed, <https://doi.org/10.1038/s41598-020-77028-8>.

Porter, E. M., et al. "Localization of Human Intestinal Defensin 5 in Paneth Cell Granules." *Infection and Immunity*, vol. 65, no. 6, June 1997, pp. 2389–95. PubMed, <https://doi.org/10.1128/iai.65.6.2389-2395.1997>.

Potten, C. S., et al. "Continuous Labelling Studies on Mouse Skin and Intestine." *Cell and Tissue Kinetics*, vol. 7, no. 3, May 1974, pp. 271–83. PubMed, <https://doi.org/10.1111/j.1365-2184.1974.tb00907.x>.

Poynter J.N. et al., "Statins and the risk of colorectal cancer." *The New England journal of medicine*. 2005;352:2184–2192. DOI: 10.1056/NEJMoa043792

---. "Extreme Sensitivity of Some Intestinal Crypt Cells to X and Gamma Irradiation." *Nature*, vol. 269, no. 5628, Oct. 1977, pp. 518–21. PubMed, <https://doi.org/10.1038/269518a0>.

---. "Kinetics and Possible Regulation of Crypt Cell Populations under Normal and Stress Conditions." *Bulletin Du Cancer*, vol. 62, no. 4, Dec. 1975, pp. 419–30.

Powell, Anne E., et al. "The Pan-ErbB Negative Regulator Lrig1 Is an Intestinal Stem Cell Marker That Functions as a Tumor Suppressor." *Cell*, vol. 149, no. 1, Mar. 2012, pp. 146–58. PubMed, <https://doi.org/10.1016/j.cell.2012.02.042>.

Rad, Roland, et al. "A Genetic Progression Model of Braf(V600E)-Induced Intestinal Tumorigenesis Reveals Targets for Therapeutic Intervention." *Cancer Cell*, vol. 24, no. 1, July 2013, pp. 15–29. PubMed, <https://doi.org/10.1016/j.ccr.2013.05.014>.

Ragusa, Simone, et al. "PROX1 Promotes Metabolic Adaptation and Fuels Outgrowth of Wnt(High) Metastatic Colon Cancer Cells." *Cell Reports*, vol. 8, no. 6, Sept. 2014, pp. 1957–73. PubMed, <https://doi.org/10.1016/j.celrep.2014.08.041>.

Ringel, Till, et al. "Genome-Scale CRISPR Screening in Human Intestinal Organoids Identifies Drivers of TGF- $\beta$  Resistance." *Cell Stem Cell*, vol. 26, no. 3, Mar. 2020, pp. 431-440.e8. DOI.org (Crossref), <https://doi.org/10.1016/j.stem.2020.02.007>.

Roepman, Paul, et al. "Colorectal Cancer Intrinsic Subtypes Predict Chemotherapy Benefit, Deficient Mismatch Repair and Epithelial-to-Mesenchymal Transition." *International Journal of Cancer*, vol. 134, no. 3, Feb. 2014, pp. 552–62. PubMed, <https://doi.org/10.1002/ijc.28387>.

Roerink, Sophie F., et al. "Intra-Tumour Diversification in Colorectal Cancer at the Single-Cell Level." *Nature*, vol. 556, no. 7702, Apr. 2018, pp. 457–62. PubMed, <https://doi.org/10.1038/s41586-018-0024-3>.

Roper, Jatin, et al. "In Vivo Genome Editing and Organoid Transplantation Models of Colorectal Cancer and Metastasis." *Nature Biotechnology*, vol. 35, no. 6, June 2017, pp. 569–76. PubMed, <https://doi.org/10.1038/nbt.3836>.

Saam, J. R., and J. I. Gordon. "Inducible Gene Knockouts in the Small Intestinal and Colonic Epithelium." *The Journal of Biological Chemistry*, vol. 274, no. 53, Dec. 1999, pp. 38071–82. PubMed, <https://doi.org/10.1074/jbc.274.53.38071>.

Sadanandam, Anguraj, et al. "A Colorectal Cancer Classification System That Associates Cellular Phenotype and Responses to Therapy." *Nature Medicine*, vol. 19, no. 5, May 2013, pp. 619–25. PubMed, <https://doi.org/10.1038/nm.3175>.

Sangiorgi, Eugenio, and Mario R. Capecchi. "Bmi1 Is Expressed in Vivo in Intestinal Stem Cells." *Nature Genetics*, vol. 40, no. 7, July 2008, pp. 915–20. PubMed, <https://doi.org/10.1038/ng.165>.

Santos, António J. M., et al. "The Intestinal Stem Cell Niche: Homeostasis and Adaptations." *Trends in Cell Biology*, vol. 28, no. 12, Dec. 2018, pp. 1062–78. PubMed, <https://doi.org/10.1016/j.tcb.2018.08.001>.

Sasaki N. et al., "Reg4+ deep crypt secretory cells function as epithelial niche for Lgr5+ stem cells in colon." *PNAS*, September 2016, 113 (37) E5399-E5407, <https://doi.org/10.1073/pnas.1607327113>

Sato, Toshiro, Johan H. van Es, et al. "Paneth Cells Constitute the Niche for Lgr5 Stem Cells in Intestinal Crypts." *Nature*, vol. 469, no. 7330, Jan. 2011, pp. 415–18. PubMed, <https://doi.org/10.1038/nature09637>.

Sato, Toshiro, Robert G. Vries, et al. "Single Lgr5 Stem Cells Build Crypt-Villus Structures in Vitro without a Mesenchymal Niche." *Nature*, vol. 459, no. 7244, May 2009, pp. 262–65. DOI.org (Crossref), <https://doi.org/10.1038/nature07935>.

Schütte, Moritz, et al. "Molecular Dissection of Colorectal Cancer in Pre-Clinical Models Identifies Biomarkers Predicting Sensitivity to EGFR Inhibitors." *Nature Communications*, vol. 8, Feb. 2017, p. 14262. PubMed, <https://doi.org/10.1038/ncomms14262>.

Shimokawa, Mariko, et al. "Visualization and Targeting of LGR5+ Human Colon Cancer Stem Cells." *Nature*, vol. 545, no. 7653, May 2017, pp. 187–92. PubMed, <https://doi.org/10.1038/nature22081>.

Simian, Marina, and Mina J. Bissell. "Organoids: A Historical Perspective of Thinking in Three Dimensions." *The Journal of Cell Biology*, vol. 216, no. 1, Jan. 2017, pp. 31–40. PubMed, <https://doi.org/10.1083/jcb.201610056>.

Simon, Jeremy M., et al. "High-Throughput Screening and Classification of Chemicals and Their Effects on Neuronal Gene Expression Using RASL-Seq." *Scientific Reports*, vol. 9, no. 1, Dec. 2019, p. 4529. DOI.org (Crossref), <https://doi.org/10.1038/s41598-019-39016-5>.

Stintzing, S., et al. "Consensus Molecular Subgroups (CMS) of Colorectal Cancer (CRC) and First-Line Efficacy of FOLFIRI plus Cetuximab or Bevacizumab in the FIRE3 (AIO KRK-0306) Trial." *Annals of Oncology: Official Journal of the European Society for Medical Oncology*, vol. 30, no. 11, Nov. 2019, pp. 1796–803. PubMed, <https://doi.org/10.1093/annonc/mdz387>.

Tabernero, Josep, et al. "Ramucirumab versus Placebo in Combination with Second-Line FOLFIRI in Patients with Metastatic Colorectal Carcinoma That Progressed during or after First-Line Therapy with Bevacizumab, Oxaliplatin, and a Fluoropyrimidine (RAISE): A Randomised, Double-Blind, Multicentre, Phase 3 Study." *The Lancet Oncology*, vol. 16, no. 5, May 2015, pp. 499–508. DOI.org (Crossref), [https://doi.org/10.1016/S1470-2045\(15\)70127-0](https://doi.org/10.1016/S1470-2045(15)70127-0).

Taieb, Julien, et al. "Exploring the Best Treatment Options for BRAF-Mutant Metastatic Colon Cancer." *British Journal of Cancer*, vol. 121, no. 6, Sept. 2019, pp. 434–42. PubMed, <https://doi.org/10.1038/s41416-019-0526-2>.

Teder, Hindrek, et al. "TAC-Seq: Targeted DNA and RNA Sequencing for Precise Biomarker Molecule Counting." *Npj Genomic Medicine*, vol. 3, no. 1, Dec. 2018, p. 34. DOI.org (Crossref), <https://doi.org/10.1038/s41525-018-0072-5>.

Tetteh, Paul W., et al. "Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters." *Cell Stem Cell*, vol. 18, no. 2, Feb. 2016, pp. 203–13. PubMed, <https://doi.org/10.1016/j.stem.2016.01.001>.

Tian, Hua, et al. "A Reserve Stem Cell Population in Small Intestine Renders Lgr5-Positive Cells Dispensable." *Nature*, vol. 478, no. 7368, Sept. 2011, pp. 255–59. PubMed, <https://doi.org/10.1038/nature10408>.

Ting, Hung-An, and Jakob von Moltke. "The Immune Function of Tuft Cells at Gut Mucosal Surfaces and Beyond." *Journal of Immunology (Baltimore, Md.: 1950)*, vol. 202, no. 5, Mar. 2019, pp. 1321–29. PubMed, <https://doi.org/10.4049/jimmunol.1801069>.

Van Cutsem, Eric, Josep Tabernero, et al. "Addition of Aflibercept to Fluorouracil, Leucovorin, and Irinotecan Improves Survival in a Phase III Randomized Trial in Patients With Metastatic Colorectal Cancer Previously Treated With an Oxaliplatin-Based Regimen." *Journal of Clinical Oncology*, vol. 30, no. 28, Oct. 2012, pp. 3499–506. DOI.org (Crossref), <https://doi.org/10.1200/JCO.2012.42.8201>.

Van Cutsem, Eric, Claus-Henning Köhne, et al. "Cetuximab and Chemotherapy as Initial Treatment for Metastatic Colorectal Cancer." *The New England Journal of Medicine*, vol. 360, no. 14, Apr. 2009, pp. 1408–17. PubMed, <https://doi.org/10.1056/NEJMoa0805019>.

van de Wetering, Marc, et al. "Prospective Derivation of a Living Organoid Biobank of Colorectal Cancer Patients." *Cell*, vol. 161, no. 4, May 2015, pp. 933–45. PubMed, <https://doi.org/10.1016/j.cell.2015.03.053>.

van der Flier, Laurens G., Andrea Haegebarth, et al. "OLFM4 Is a Robust Marker for Stem Cells in Human Intestine and Marks a Subset of Colorectal Cancer Cells." *Gastroenterology*, vol. 137, no. 1, July 2009, pp. 15–17. DOI.org (Crossref), <https://doi.org/10.1053/j.gastro.2009.05.035>.

van der Flier, Laurens G., Marielle E. van Gijn, et al. "Transcription Factor Achaete Scute-Like 2 Controls Intestinal Stem Cell Fate." *Cell*, vol. 136, no. 5, Mar. 2009, pp. 903–12. DOI.org (Crossref), <https://doi.org/10.1016/j.cell.2009.01.031>.



van Es, Johan H., et al. "Dll1+ Secretory Progenitor Cells Revert to Stem Cells upon Crypt Damage." *Nature Cell Biology*, vol. 14, no. 10, Oct. 2012, pp. 1099–104. PubMed, <https://doi.org/10.1038/ncb2581>.

Verissimo, Carla S., et al. "Targeting Mutant RAS in Patient-Derived Colorectal Cancer Organoids by Combinatorial Drug Screening." *ELife*, vol. 5, Nov. 2016, p. e18489. PubMed, <https://doi.org/10.7554/eLife.18489>.

Vlachogiannis, Georgios, et al. "Patient-Derived Organoids Model Treatment Response of Metastatic Gastrointestinal Cancers." *Science (New York, N.Y.)*, vol. 359, no. 6378, Feb. 2018, pp. 920–26. PubMed, <https://doi.org/10.1126/science.aao2774>.

von Furstenberg, Richard J., et al. "Side Population Sorting Separates Subfractions of Cycling and Non-Cycling Intestinal Stem Cells." *Stem Cell Research*, vol. 12, no. 2, Mar. 2014, pp. 364–75. DOI.org (Crossref), <https://doi.org/10.1016/j.scr.2013.10.012>.

von Kleist, S., et al. "Immunohistology of the Antigenic Pattern of a Continuous Cell Line from a Human Colon Tumor." *Journal of the National Cancer Institute*, vol. 55, no. 3, Sept. 1975, pp. 555–60. PubMed, <https://doi.org/10.1093/jnci/55.3.555>.

Xi, Yue, and Pengfei Xu. "Global Colorectal Cancer Burden in 2020 and Projections to 2040." *Translational Oncology*, vol. 14, no. 10, Oct. 2021, p. 101174. DOI.org (Crossref), <https://doi.org/10.1016/j.tranon.2021.101174>.

Xiao, Haitao, et al. "Poorly Differentiated Colorectal Cancers: Correlation of Microsatellite Instability with Clinicopathologic Features and Survival." *American Journal of Clinical Pathology*, vol. 140, no. 3, Sept. 2013, pp. 341–47. PubMed, <https://doi.org/10.1309/AJCP8P2DYNKGRBVI>.

Xie, Yuan-Hong, et al. "Comprehensive Review of Targeted Therapy for Colorectal Cancer." *Signal Transduction and Targeted Therapy*, vol. 5, no. 1, Mar. 2020, p. 22. PubMed, <https://doi.org/10.1038/s41392-020-0116-z>.

Yan, Kelley S., et al. "The Intestinal Stem Cell Markers Bmi1 and Lgr5 Identify Two Functionally Distinct Populations." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 2, Jan. 2012, pp. 466–71. PubMed, <https://doi.org/10.1073/pnas.1118857109>.

Ye, Chaoyang, et al. "DRUG-Seq for Miniaturized High-Throughput Transcriptome Profiling in Drug Discovery." *Nature Communications*, vol. 9, no. 1, Dec. 2018, p. 4307. DOI.org (Crossref), <https://doi.org/10.1038/s41467-018-06500-x>.

Yilmaz, Ömer H., et al. "MTORC1 in the Paneth Cell Niche Couples Intestinal Stem-Cell Function to Calorie Intake." *Nature*, vol. 486, no. 7404, June 2012, pp. 490–95. PubMed, <https://doi.org/10.1038/nature11163>.

Yin, Xiaolei, et al. "Niche-Independent High-Purity Cultures of Lgr5+ Intestinal Stem Cells and Their Progeny." *Nature Methods*, vol. 11, no. 1, Jan. 2014, pp. 106–12. DOI.org (Crossref), <https://doi.org/10.1038/nmeth.2737>.

Yu, Shiyan, et al. "Paneth Cell Multipotency Induced by Notch Activation Following Injury." *Cell Stem Cell*, vol. 23, no. 1, July 2018, pp. 46-59.e5. PubMed, <https://doi.org/10.1016/j.stem.2018.05.002>.

Zeineldin, Maged, and Kristi L. Neufeld. "Understanding Phenotypic Variation in Rodent Models with Germline Apc Mutations." *Cancer Research*, vol. 73, no. 8, Apr. 2013, pp. 2389–99. PubMed, <https://doi.org/10.1158/0008-5472.CAN-12-4607>.

Zhan, Tianzuo, et al. "MEK Inhibitors Activate Wnt Signalling and Induce Stem Cell Plasticity in Colorectal Cancer." *Nature Communications*, vol. 10, no. 1, Dec. 2019, p. 2197. DOI.org (Crossref), <https://doi.org/10.1038/s41467-019-09898-0>.

Zhang, Lu, and Jerry W. Shay. "Multiple Roles of APC and Its Therapeutic Implications in Colorectal Cancer." *Journal of the National Cancer Institute*, vol. 109, no. 8, Aug. 2017. PubMed, <https://doi.org/10.1093/jnci/djw332>.

## Chapter 5: Annexes

### 5.1. Single-Cell Studies of Intestinal Stem Cell Heterogeneity During Homeostasis and Regeneration

#### Authors:

Maxim Norkin<sup>1\*</sup>, Claudia Capdevila<sup>23\*</sup>, Ruben I Calderon<sup>23</sup>, Tianhong Su<sup>23</sup>, Maria Trifas<sup>23</sup>, Paloma Ordóñez-Morán<sup>4</sup>, Kelley S Yan<sup>56</sup>

#### Authors' affiliation:

<sup>1</sup> Swiss Institute for Experimental Cancer Research, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

<sup>2</sup> Division of Digestive and Liver Diseases, Department of Medicine, Columbia Center for Human Development, Columbia Stem Cell Initiative, New York, NY, USA.

<sup>3</sup> Department of Genetics and Development, Columbia University Irving Medical Center, New York, NY, USA.

<sup>4</sup> Department of Cancer and Stem Cells, School of Medicine, Biodiscovery Institute, University of Nottingham, Nottingham, UK. paloma.ordonezmoran@nottingham.ac.uk.

<sup>5</sup> Division of Digestive and Liver Diseases, Department of Medicine, Columbia Center for Human Development, Columbia Stem Cell Initiative, New York, NY, USA. ky2004@cumc.columbia.edu.

<sup>6</sup> Department of Genetics and Development, Columbia University Irving Medical Center, New York, NY, USA. ky2004@cumc.columbia.edu.

\*these authors contributed equally

My contributions in this manuscript are the following: I collected analyzed the data for marker expression in epithelial cells and wrote a draft version of chapter “Cellular Heterogeneity During Intestinal Homeostasis”. Further, participated in the writing of the whole manuscript. Additionally, I made Figure 1.



## Single-Cell Studies of Intestinal Stem Cell Heterogeneity During Homeostasis and Regeneration

Maxim Norkin, Claudia Capdevila, Ruben I. Calderon, Tianhong Su, Maria Trifas, Paloma Ordóñez-Morán, and Kelley S. Yan

### Abstract

Single-cell RNA-sequencing (scRNA-seq) provides a unique opportunity to study heterogeneous cell populations within tissues, including the intestinal epithelium, to gain detailed molecular insights into their biology. Many new putative markers of intestinal stem cells and their progeny have been described using single-cell transcriptomics, which has contributed to the identification of novel subpopulations of mature cell types and insight into their developmental trajectories. This approach has revealed tremendous cellular heterogeneity within the intestinal epithelium that is concordant with its diverse and multifaceted functions. We discuss the function of these subpopulations during tissue homeostasis, as well as putative subpopulations with inducible regenerative potential following tissue injury.

**Key words** Single-cell RNA-sequencing, Stem cell, Intestine, Epithelium, Lineage, Homeostasis, Regeneration, Differentiation, Plasticity

---

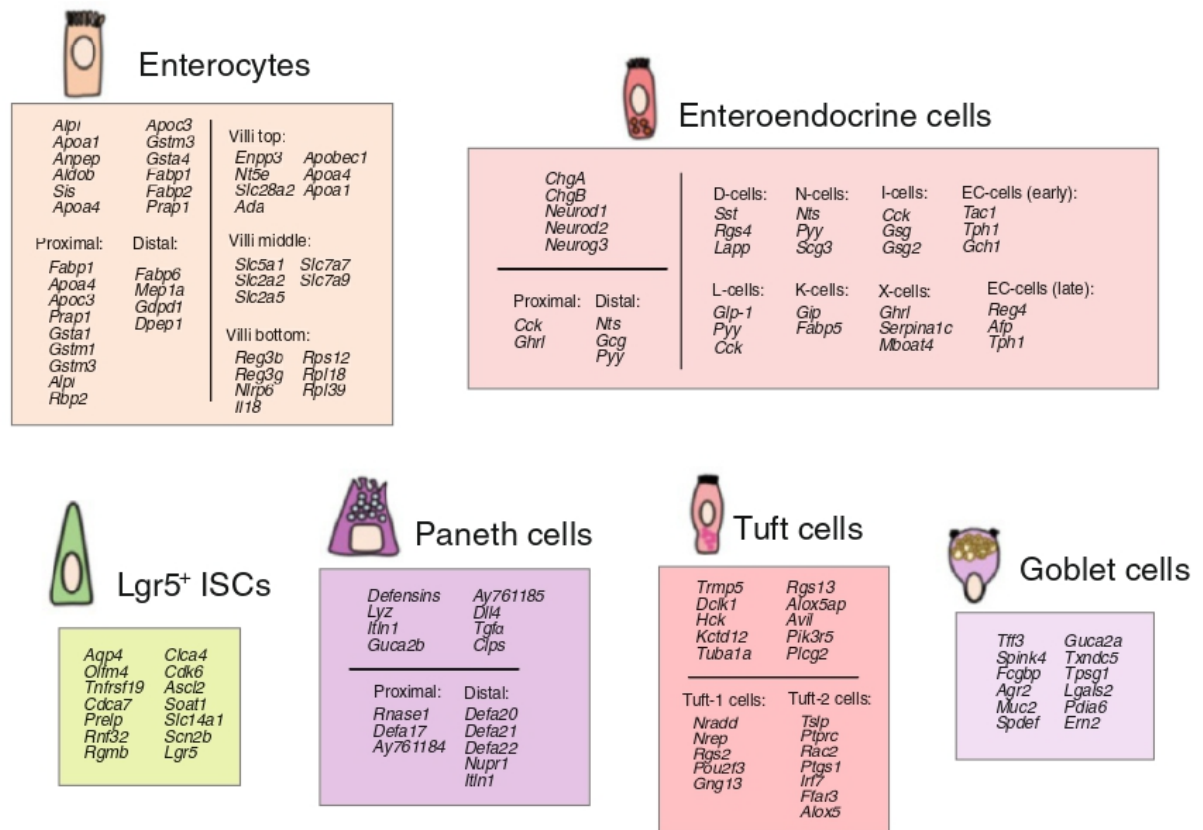
### 1 Introduction

The renewal of the intestinal epithelium is characterized by both tight regulation and an immense turnover capacity. This rapid renewal is important given the relatively harsh conditions to which the epithelial cells are exposed during their lifetime. In addition to its remarkable proliferative activity, the intestinal epithelium is extremely diverse in its cellular composition, reflecting a high degree of heterogeneity within its major lineages [1, 2] (Fig. 1). For example, the enteroendocrine cell (EEC) lineage is thought to be comprised of at least ten different cell types [1, 3]. Additionally, there is controversy over the heterogeneity of its stem cell compartment that supports tissue turnover and regeneration. Multiple studies highlight the regenerative function of

---

Maxim Norkin and Claudia Capdevila contributed equally to this work.

Paloma Ordóñez-Morán (ed.), *Intestinal Stem Cells: Methods and Protocols*, Methods in Molecular Biology, vol. 2171, [https://doi.org/10.1007/978-1-0716-0747-3\\_9](https://doi.org/10.1007/978-1-0716-0747-3_9), © Springer Science+Business Media, LLC, part of Springer Nature 2020



**Fig. 1** Markers of intestinal stem cells and their progeny. Summary of putative marker genes of different intestinal cell types based on transcriptional profiling. Enterocytes: left upper corner—enterocyte specific markers, left bottom corner—proximal and distal markers of enterocytes, right—genes with variable expression in enterocytes along the intestinal villus axis. EECs: left upper corner—EEC specific markers, left bottom corner—proximal and distal markers of EECs, right—specific markers of EEC subpopulations. Lgr5<sup>+</sup> ISCs: stem cell specific markers. Paneth cells: top—specific genes, bottom—proximal and distal markers of corresponding populations. Tuft cells: up—specific markers, bottom—specific markers of tuft subpopulations. Goblet cells: specific markers

reserve stem cells that are deployed during times of tissue injury using distinct mechanisms from those employed during homeostasis [1, 4–14]. Moreover, numerous progenitor cells and intermediates along the developmental trajectories of individual lineages remain to be established.

Single-cell RNA-sequencing (scRNA-seq) is a rapidly evolving method that is becoming an indispensable tool to investigate the cellular composition of tissues. Continuous changes in cellular identity can also be analyzed by real time-resolved single-cell transcriptomics, which can be used for cellular processes that display transient activation of a marker gene [3]. Early studies using scRNA-seq in the intestine were performed on mouse epithelial cells [15, 16]. More recent reports have delineated the mRNA expression in human epithelial cells. The cell types identified in human tissues largely correlated with the previous scRNA-seq

analyses performed on the mouse [17]. Overall, scRNA-seq studies have revealed new insights into the cell types and mechanisms that underlie intestinal homeostasis and epithelial regeneration. Here, we summarize recent findings using this technology as well as new questions raised by these studies (Figs. 1 and 2).

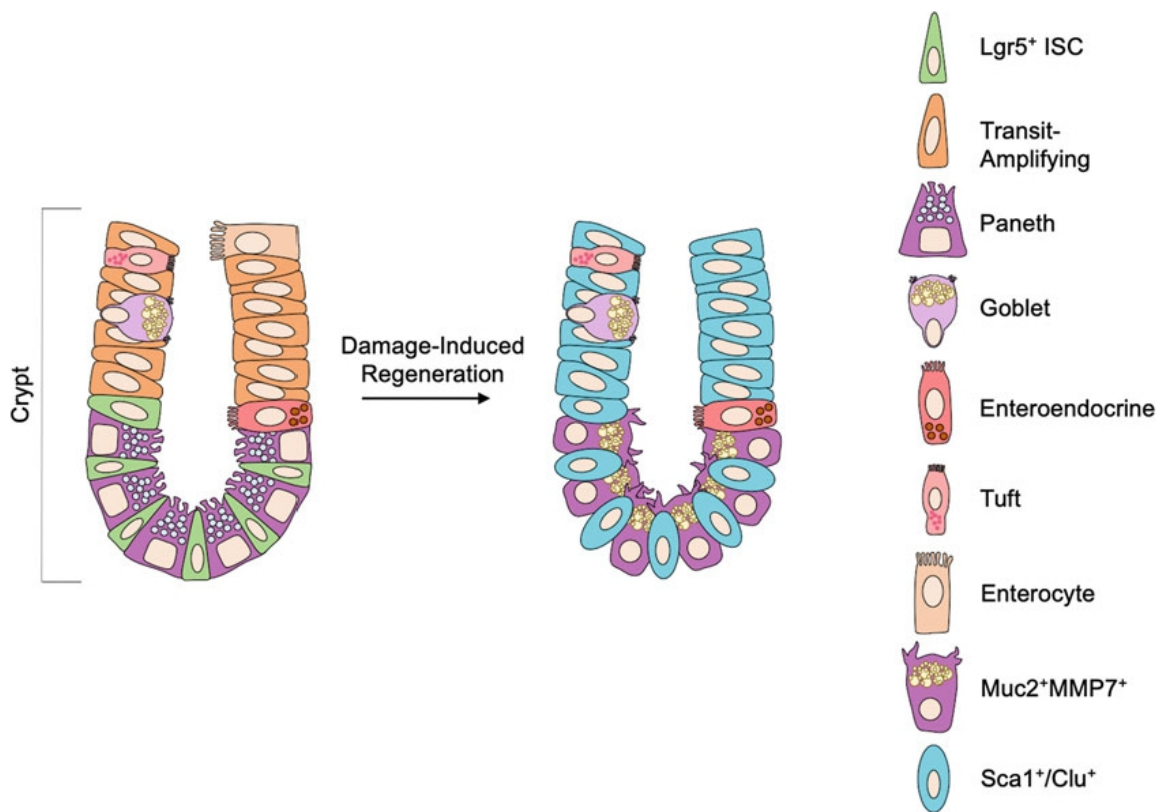
## 2 Cellular Heterogeneity During Intestinal Homeostasis

The constant tissue renewal of the intestinal epithelium during homeostasis is fueled by continuously dividing intestinal stem cells (ISCs) that reside at the base of the crypts [18–20]. These crypt-base columnar ISCs have been molecularly defined by the expression of the R-spondin receptor *Lgr5*, a leucine rich repeat containing G protein-coupled receptor [18]. Single-cell transcriptomic analysis reveals that *Lgr5*-eGFP<sup>+</sup> ISCs isolated from *Lgr5-eGFP-IRES-CreERT2* mice are composed of distinct clusters of both cycling and non-cycling Wnt- and R-spondin-dependent ISC as well as nascent transit-amplifying (TA) cell populations, consistent with multiple reports suggesting heterogeneity within the ISC pool [10, 21, 22]. Many markers of ISCs were initially identified based on microarray and bulk RNA-seq profiling of cells isolated using *Lgr5* reporter expression. Subsequently, transcriptional profiling using scRNA-seq technology helped to support these markers and also allowed the identification of new ones, with varying degrees of characterization and validation (Fig. 1): *Aqp4*, *Olfm4*, *Tnfrsf19*, *Cdca7*, *Prepl*, *Rnf32*, *Cdk6*, *Rgmb*, *Clca4*, *Scn2b*, *Slc14a1*, *Ascl2*, and *Soat1* [2, 15, 16, 23–25].

Under homeostatic conditions, *Lgr5*<sup>+</sup> ISCs regenerate the intestinal epithelium [18]. In contrast, a distinct, label-retaining cell at the +4 cell position in intestinal crypts has been reported and historically proposed to serve as a quiescent/slowly cycling, reserve stem cell population that comes into play upon loss of *Lgr5*<sup>+</sup> ISCs [4–6, 8, 26–28]. Even though multiple markers have been described for the +4 cell [4–8], more recent studies show that many of these overlap with *Lgr5*<sup>+</sup> ISCs, or are expressed very broadly throughout the crypt [29–31]. While it remains unclear if there are other dedicated reserve stem cell populations that coexist within the crypt alongside the *Lgr5*<sup>+</sup> ISCs, there is increasing evidence of diverse cell types that are capable of exhibiting cellular plasticity to become activated upon injury conditions to rapidly regenerate the epithelium following loss of *Lgr5*<sup>+</sup> ISCs [32]. Indeed, some recent observations suggest that it is not a single type of quiescent stem cell but rather a functionally distinct, non-*Lgr5*-expressing cell type that reverts to a stem cell state in an injury-inducible fashion [1, 4, 8].

Globally, along the crypt–villus axis, the progenitor cells arising from *Lgr5*<sup>+</sup> ISCs and their immediate TA progeny reside in the





**Fig. 2** Injury-induced regeneration of the intestinal epithelium. Injury-induced regeneration of the intestinal epithelium is mediated by Sca1<sup>+</sup> and/or Clu<sup>+</sup> cells (blue), which are rare under homeostatic conditions. Sca1<sup>+</sup>/Clu<sup>+</sup> cells repopulate the crypts of the small intestine and regenerate the epithelium following damage-induced loss of Lgr5<sup>+</sup> ISCs

crypts close to the stem cells and move into the villi upon differentiation. Absorptive enterocytes are the most abundant cell lineage in the intestinal epithelium. Additionally, there are multiple secretory cell lineages including mucus-producing goblet cells, Paneth cells, EECs, and rare tuft cells [20]. All these differentiated cell lineages contribute to specific functions of the intestine. The advantage of scRNA-seq methodology is that it can be applied to study specific gene expression in individual subpopulations and it can help to infer the hierarchical relationships of individual lineages (Fig. 1).

Tuft cells are a rare cell type involved in chemosensory function. Tuft cells express proteins known to be involved in taste signal transduction and also in an immune response against parasite infection [33, 34]. Tuft cells express such markers as *Trpm5*, *Dclk1*, *Rgs13*, *Alox5ap*, *Avil*, *Hck*, *Kctd12*, *Tuba1a*, *Pik3r5*, and *Pleg2* [2, 15, 16, 35, 36]. Recent findings using scRNA-seq demonstrated the existence of two distinct subtypes of tuft cells: tuft-1 cells are highly enriched in genes related to neuronal development, while tuft-2 cells show upregulation of genes related to immunity. Specifically, tuft-1 cells are enriched for *Nradd*, *Gng13*, *Nrep*, *Rgs2*,

and *Pou2f3* [2] (Fig. 1). Conversely, tuft-2 cells showed higher expression levels of the TH2-promoting cytokine thymic stromal lymphopoietin (*Tslp*) and *Ptprc* (pan-immune cell marker CD45), as well as enrichment in transcripts for *Rac2*, *Ptgs1*, *Irf7*, *Ffar3*, and *Alox5*.

Enterocytes are the most abundant cell type in the intestinal epithelium, making up to 80% of intestinal epithelial cells [37]. Enterocytes are predominantly located in the villus region and function in the hydrolysis, absorption, and transport of nutrients [38]. Many enterocyte markers were identified using bulk RNA-seq [39] and scRNA-seq [2, 15, 16, 41]; among them are *Alpi*, *Apoa1*, *Anpep*, *Aldob*, *Sis*, *Apoa4*, *Prap1*, *Apoc3*, *Gstm3*, *Gsta4*, *Fabp1*, and *Fabp2*. Recently, scRNA-seq was used to identify enterocyte progenitor populations [15, 16], to examine enterocyte regional diversity throughout the gut [2], and to investigate the spatial allocation of distinct functional classes of enterocytes along the intestinal crypt–villus axis [40, 41]. It has been shown that earlier enterocyte progenitors express high transcript levels of ribosomal proteins (*Rn45s*, *Rps19*, and *Rps12*), *Dmbt1*, *Reg3g*, *Ube2c* and low levels of those for enterocyte-specific genes (*Prap1*, *Apoa1*, *Apoa4*, *Apoc3*, etc.) [2]. Late progenitor cells lose the expression of ribosomal proteins with concurrent elevation of enterocyte-specific transcripts. Finally, mature enterocytes are characterized by further upregulation of enterocyte-specific mRNAs.

A study of regional enterocyte markers reveals that *Fabp1*, *Apoa4*, *Apoc3*, *Gsta1*, *Gstm3*, *Gstm1*, *Alpi*, *Prap1*, and *Rbp2* are highly expressed in the proximal part of the intestine, while *Fabp6*, *Mep1a*, *Dpep1*, and *Gdgd1* are predominantly expressed in the distal small intestine [2]. This is consistent with differential absorptive functions along the longitudinal proximal-to-distal gut axis. Moreover, a combination of high-throughput scRNA-seq and bulk RNA-seq of laser-microdissected crypt–villus sections revealed that enterocyte transcriptional signatures differ along the crypt–villus axis consistent with their zonation [41]. Enterocytes at the bottom of the crypt showed enrichment in ribosomal/proliferation signatures (*Rps12*, *Rpl18*, and *Rpl39*) and antimicrobial program peptides (*Reg3b*, *Reg3g*, *Nlrp6*, and *Il18*) [41]. Enterocytes in the middle of the crypt–villus axis are enriched in the genes responsible for the processing and absorption of various nutrients, especially amino acids (*Slc7a7* and *Slc7a9*) and carbohydrates (*Slc5a1*, *Slc2a2*, and *Slc2a5*) [41]. Enterocytes at the top of the villus exhibited a distinct expression program: enrichment in cell adhesion signature genes (*Egfr*, *Klf4*, *Fos*, *Junb*), purine catabolism genes (*Enpp3*, *Nt5e*, *Slc28a2*, and *Ada*), and apolipoproteins/cholesterol processing genes (*Apobec1*, *Apoa4*, and *Apoa1*) (Fig. 1) [41].

Goblet cells are secretory cells present in both the small intestine and colon. They produce and secrete mucus into the intestinal lumen, which facilitates the migration of chyme through the gut



and contributes to a physical barrier that prevents microorganisms and toxins from direct contact with the mucosa. In addition, goblet cells are also shown to be expanded in response to parasite infection [2, 42, 43]. Several scRNA-seq experiments revealed multiple markers of goblet cells: *Tff3*, *Spink4*, *Fcgbp*, *Agr2*, *Muc2*, *Txnrc5*, *Tpsg1*, *Spdef*, *Guca2a*, *Lgals2*, *Pdia6*, and *Ern2* [2, 15, 16]. *Spdef* has been shown to have an essential role in goblet cell differentiation [44, 45] (Fig. 1).

Paneth cells are intermingled with *Lgr5*<sup>+</sup> ISCs at the crypt base and function in antimicrobial defense and metabolic regulation of ISCs. They interact with ISCs via signaling pathways such as Notch, Wnt, and EGF [46–48]. While the first Paneth cell markers (lysozyme and defensins) were reported decades ago [49, 50], recent RNA-seq studies on sorted CD24<sup>+</sup> Paneth cells reveal additional transcripts differentially enriched in Paneth cells relative to ISCs: *Dll4*, *Tgfa*, *Wnt11*, *Clps*, *AY761185*, *Itln1*, and *Guca2b* [46]. Recent scRNA-seq studies additionally showed that Paneth cells also exhibit regional diversity, and the expression profiles of these cells are different in distal and proximal part of the small intestine: for example, *Rnase1*, *Defa17*, and *AY761184* are highly expressed in the duodenum, while *Defa20*, *Defa21*, *Defa22*, *Nupr1*, and *Itln1* show the highest expression in ileum [2]. Other Paneth-specific markers that show uniform expression throughout the gut include *Lyz1*, *Ang4*, *Clps*, and *Habp2* (Fig. 1) [2]. It still remains unclear how and why Paneth cells are heterogeneous, and the functional implications on their interactions with ISCs and on the other Paneth cell functions.

EECs produce hormones that regulate digestion and metabolism, and participate in chemosensation, including nutrient detection. EECs represent a very small fraction of epithelial cells that are widely dispersed throughout the intestinal epithelium and have been elusive to molecular characterization due to their marked cellular heterogeneity. While EECs are located throughout the entire intestinal epithelium, they also exhibit regionalized patterns of gene expression. Recently, scRNA-seq studies have identified new putative markers for the EECs subpopulations and further underscored the heterogeneity of this lineage. At least ten different EEC subtypes have been identified either in vivo or via in vitro intestinal organoid culture [2, 3, 15, 16, 19, 51]. *Neurog3*, *Neurod1*, and *Neurod2* are known EEC markers specific to EEC progenitor populations [2]. In addition, the EEC markers *ChgA* and *ChgB* were shown to be predominantly expressed in enterochromaffin (EC) cells [2, 3]. Many more putative novel markers were identified for each of the EEC subpopulations, using sequencing technology, including scRNA-seq: *Gip*, *Fabp5* (K-cells); *Reg4*, *Afp*, *Tph1* (EC-cells (late)); *Sst*, *Lapp*, *Rgs4* (D-cells); *Ghrl*, *Serpina1c*, *Mboat4* (X-cells); *Cck*, *Gsg*, *Gsg-2* (I-cells); *Glp-1*, *Pyy*, *Cck* (L-cells); *Nts*, *Pyy*, *Scg3* (N-cells); *Tac1*, *Tph1*, *Gch1* (EC cells (early)) [2, 3,

15, 16, 19, 51, 52]. Some of the markers such as *Cck* and *Gcg* are shown to be expressed in several EEC subtypes simultaneously. Furthermore, while some markers are predominantly expressed in the proximal small intestine (*Cck* and *Ghr1*), others are more localized in the distal region (*Nts*, *Gcg*, and *Pyy*). *Sct* is expressed in all subtypes of EECs and is detected in both proximal and distal parts of the gastrointestinal tract [3, 53] (Fig. 1). These EEC lineage markers and their distribution need further confirmatory experimental validation and biological interpretation. The diverse repertoire of EEC subsets found by scRNA-seq likely reflects the multifaceted functions of these cells that orchestrate chemosensation, digestion, metabolism, and communication with other organs via their hormone products.

---

### 3 Cellular Heterogeneity During Tissue Repair

Multiple distinct types of cells capable of regenerating the intestinal epithelium after tissue injury or *Lgr5*<sup>+</sup> ISC loss have been described [32]. Importantly, these include progenitor cells of both secretory and absorptive cell lineages, as well as more differentiated cells that reside within the intestinal crypt. Functional studies using lineage tracing in mice using *Alpi*<sup>+</sup> absorptive enterocytes [54] and secretory lineage markers [8, 9, 55, 56] support the possibility that there are multiple populations capable of reconstituting the intestinal epithelium following injury, suggesting a model of cellular plasticity in which perturbation of homeostasis enables even lineage-committed cells to regain regenerative capability. For instance, recent comparative transcriptomics using scRNA-seq to examine cells isolated from various reporter mice revealed that lineage-committed EECs also exhibit regenerative potential in the context of irradiation injury [1]. In this setting, *Lgr5*<sup>+</sup> ISCs are rapidly ablated, and *Prox1*<sup>+</sup> crypt cells of the enteroendocrine lineage can come into play to regenerate the epithelium. Notably, *Prox1* is expressed in quiescent crypt cells near the +4 cell position and identifies cells of the tuft and enteroendocrine lineages, suggesting a shared lineage progenitor [1]. Thus, despite being already-committed secretory populations, these cells are capable of regenerating the epithelium following tissue injury.

scRNA-seq has recently provided additional insight into potential mechanisms for injury-induced regeneration of the intestinal epithelium. Recent reports of injury-induced regeneration have identified and characterized two additional subpopulations of epithelial cells that occupy the crypt compartment following tissue damage. Various injury models, including helminth infection, result in expansion of crypt cells that adopt a fetal-like transcriptional program, in which the crypt cells are functionally Wnt-independent and express the protein *Sca1* [12, 14, 57,

58]. In another report, a rare population that selectively expands upon damage is characterized by high expression of clusterin (Clu) (Fig. 2) [13]. The inter-relationship between these populations remains undefined.

In contrast to Wnt pathway-dependent Lgr5<sup>+</sup> ISCs, Sca1<sup>+</sup> cells rise, peak, and return to basal expression levels in different injury/perturbation models including parasitic helminth infection, irradiation, and Lgr5<sup>+</sup> ISC ablation by diphtheria toxin, in a manner that is inversely proportional to the loss of Lgr5<sup>+</sup> ISCs [12]. These Sca1<sup>+</sup> cells are capable of forming regenerative, primarily undifferentiated, fetal-like proliferative crypts devoid of canonical adult Lgr5<sup>+</sup> ISC markers shortly after injury [12]. By the time Lgr5<sup>+</sup> ISCs re-emerge and repopulate the regenerating epithelium, Sca1 expression decreases. Notably, Sca1<sup>+</sup> cells in the adult small intestine are ultimately derived from Lgr5<sup>+</sup> ISCs; yet, they can arise and regenerate the epithelium independent of Lgr5<sup>+</sup> ISCs themselves [12]. Importantly, this report provides evidence of epithelial cells co-opting a fetal-like transcriptional program for repair of damaged adult tissue [12].

Cells characterized by high expression levels of clusterin (Clu) and referred to as “revival” stem cells (revSCs) were recently reported to selectively expand following irradiation damage [13]. RevSCs were identified by scRNA-seq analysis of the regenerating intestinal epithelium 3 days after irradiation, and were characterized as a preexisting Lgr5<sup>-</sup>, Clu<sup>+</sup>, YAP-dependent injury-induced quiescent cell type that is most prominently found following tissue damage [13]. Immunofluorescent analysis of transgenic Clu reporter mice confirmed limited numbers of Clu<sup>+</sup> cells within the epithelium during homeostatic conditions. Indeed, crypts populated by Clu<sup>+</sup> cells did not harbor Lgr5<sup>+</sup> ISCs [13]. However, lineage tracing experiments initiated under homeostatic conditions demonstrated that those rare Clu<sup>+</sup> cells were multipotent and capable of giving rise to Lgr5<sup>+</sup> ISCs [13]. Following irradiation damage, Clu<sup>+</sup> cells repopulated the small intestine and colon epithelium [13]. Consistent with these data, ablation of Clu<sup>+</sup> cells displayed no detrimental phenotype under homeostasis; however, impaired regeneration of the epithelium was observed in the context of irradiation, acute and chronic colitis [13]. Similarly, both Sca1<sup>+</sup> and Clu<sup>+</sup> cells are derived from Lgr5<sup>+</sup> progeny, although the degree of overlap between these and other regenerative populations remains unknown. Is this repair capacity restricted to certain discrete cell populations, or do these gene signatures represent a common intermediate state where lineage-committed cells transiently converge as they go backward in the lineage hierarchy and reacquire regenerative potential (Fig. 2)? Further studies are needed in order to address these questions.



## 4 Conclusion

scRNA-seq has enabled the description of intestinal epithelial cell populations and their interrelationships with unprecedented granularity. One caveat to using this technology to study cellular heterogeneity is that cellular categorization ultimately requires experimental functional validation, which currently remains a bottleneck. In addition to providing more accurate definitions for mature populations and describing potential inferred lineage trajectories, scRNA-seq has demonstrated regional differences in cell type prevalence, as well as substantial compartmentalization along the crypt–villus axis. Taken together, these studies highlight the high degree of cellular diversity in intestinal tissue architecture, which is concordant with the great diversity of its biological functions. Moreover, these studies underscore the notion that cellular identity is malleable and influenced by environmental cues, as cells can adopt different functions in different biological contexts, further expanding the potential cellular heterogeneity within the tissue.

## Acknowledgments

K.S.Y. is supported by NIH R03DK114656 and U01DK103155, BWF CAMS, Louis V. Gerstner, Jr. Scholars Award, Lisa Dean Moseley Foundation, and the Irma T. Hirschl Career Award. C. C. is supported by a NYSTEM predoctoral training grant and T.S. is supported by a Berrie Foundation fellowship.

## References

1. Yan KS, Gevaert O, Zheng GXY, Anchang B, Probert CS, Larkin KA, Davies PS, Cheng ZF, Kaddis JS, Han A, Roelf K, Calderon RI, Cynn E, Hu X, Mandleywala K, Wilhelmy J, Grimes SM, Corney DC, Boutet SC, Terry JM, Belgrader P, Ziraldo SB, Mikkelsen TS, Wang F, von Furstenberg RJ, Smith NR, Chandrakesan P, May R, Chrissy MAS, Jain R, Cartwright CA, Niland JC, Hong YK, Carrington J, Breault DT, Epstein J, Houchen CW, Lynch JP, Martin MG, Plevritis SK, Curtis C, Ji HP, Li L, Henning SJ, Wong MH, Kuo CJ (2017) Intestinal enteroendocrine lineage cells possess homeostatic and injury-inducible stem cell activity. *Cell Stem Cell* 21(1):78–90 e76. <https://doi.org/10.1016/j.stem.2017.06.014>
2. Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K, Smillie C, Burgin G, Delorey TM, Howitt MR, Katz Y, Tirosh I, Beyaz S, Dionne D, Zhang M, Raychowdhury R, Garrett WS, Rozenblatt-Rosen O, Shi HN, Yilmaz O, Xavier RJ, Regev A (2017) A single-cell survey of the small intestinal epithelium. *Nature* 551(7680):333–339. <https://doi.org/10.1038/nature24489>
3. Gehart H, van Es JH, Hamer K, Beumer J, Kretzschmar K, Dekkers JF, Rios A, Clevers H (2019) Identification of enteroendocrine regulators by real-time single-cell differentiation mapping. *Cell* 176(5):1158–1173 e1116. <https://doi.org/10.1016/j.cell.2018.12.029>
4. Sangiorgi E, Capecchi MR (2008) Bmi1 is expressed in vivo in intestinal stem cells. *Nat Genet* 40(7):915–920. <https://doi.org/10.1038/ng.165>
5. Tian H, Biehs B, Warming S, Leong KG, Rangell L, Klein OD, de Sauvage FJ (2011) A

- reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. *Nature* 478(7368):255–259. <https://doi.org/10.1038/nature10408>
6. Montgomery RK, Carlone DL, Richmond CA, Farilla L, Kranendonk ME, Henderson DE, Baffour-Awuah NY, Ambruzs DM, Fogli LK, Algra S, Breault DT (2011) Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. *Proc Natl Acad Sci U S A* 108(1):179–184. <https://doi.org/10.1073/pnas.1013004108>
  7. Powell AE, Wang Y, Li Y, Poulin EJ, Means AL, Washington MK, Higginbotham JN, Juchheim A, Prasad N, Levy SE, Guo Y, Shyr Y, Aronow BJ, Haigis KM, Franklin JL, Coffey RJ (2012) The pan-ErbB negative regulator Lrig1 is an intestinal stem cell marker that functions as a tumor suppressor. *Cell* 149(1):146–158. <https://doi.org/10.1016/j.cell.2012.02.042>
  8. Yan KS, Chia LA, Li X, Ootani A, Su J, Lee JY, Su N, Luo Y, Heilshorn SC, Amieva MR, Sangiorgi E, Capecchi MR, Kuo CJ (2012) The intestinal stem cell markers Bmi1 and Lgr5 identify two functionally distinct populations. *Proc Natl Acad Sci U S A* 109(2):466–471. <https://doi.org/10.1073/pnas.1118857109>
  9. van Es JH, Sato T, van de Wetering M, Lyubimova A, Yee Nee AN, Gregorieff A, Sasaki N, Zeinstra L, van den Born M, Korving J, Martens ACM, Barker N, van Oudenaarden A, Clevers H (2012) Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. *Nat Cell Biol* 14(10):1099–1104. <https://doi.org/10.1038/ncb2581>
  10. Buczacki SJ, Zecchini HI, Nicholson AM, Russell R, Vermeulen L, Kemp R, Winton DJ (2013) Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* 495(7439):65–69. <https://doi.org/10.1038/nature11965>
  11. Asfaha S, Hayakawa Y, Muley A, Stokes S, Graham TA, Ericksen RE, Westphalen CB, von Burstin J, Mastracci TL, Worthley DL, Guha C, Quante M, Rustgi AK, Wang TC (2015) Krt19 (+)/Lgr5(-) cells are radioresistant cancer-initiating stem cells in the colon and intestine. *Cell Stem Cell* 16(6):627–638. <https://doi.org/10.1016/j.stem.2015.04.013>
  12. Nusse YM, Savage AK, Marangoni P, Rosendahl-Huber AKM, Landman TA, de Sauvage FJ, Locksley RM, Klein OD (2018) Parasitic helminths induce fetal-like reversion in the intestinal stem cell niche. *Nature* 559(7712):109–113. <https://doi.org/10.1038/s41586-018-0257-1>
  13. Ayyaz A, Kumar S, Sangiorgi B, Ghoshal B, Gosio J, Ouladan S, Fink M, Barutcu S, Trcka D, Shen J, Chan K, Wrana JL, Gregorieff A (2019) Single-cell transcriptomes of the regenerating intestine reveal a revival stem cell. *Nature* 569(7754):121–125. <https://doi.org/10.1038/s41586-019-1154-y>
  14. Pont AR, Yan KS (2018) Intestinal crypts assume the fetal position in response to injury. *Cell Stem Cell* 23(2):158–159. <https://doi.org/10.1016/j.stem.2018.07.013>
  15. Grun D, Lyubimova A, Kester L, Wiebrands K, Basak O, Sasaki N, Clevers H, van Oudenaarden A (2015) Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* 525(7568):251–255. <https://doi.org/10.1038/nature14966>
  16. Grun D, Muraro MJ, Boisset JC, Wiebrands K, Lyubimova A, Dharmadhikari G, van den Born M, van Es J, Jansen E, Clevers H, de Koning EJP, van Oudenaarden A (2016) De novo prediction of stem cell identity using single-cell transcriptome data. *Cell Stem Cell* 19(2):266–277. <https://doi.org/10.1016/j.stem.2016.05.010>
  17. Fujii M, Matano M, Toshimitsu K, Takano A, Mikami Y, Nishikori S, Sugimoto S, Sato T (2018) Human intestinal organoids maintain self-renewal capacity and cellular diversity in niche-inspired culture condition. *Cell Stem Cell* 23(6):787–793 e786. <https://doi.org/10.1016/j.stem.2018.11.016>
  18. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, Cozijnsen M, Haegebarth A, Korving J, Begthel H, Peters PJ, Clevers H (2007) Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* 449(7165):1003–1007. <https://doi.org/10.1038/nature06196>
  19. Gehart H, Clevers H (2019) Tales from the crypt: new insights into intestinal stem cells. *Nat Rev Gastroenterol Hepatol* 16(1):19–34. <https://doi.org/10.1038/s41575-018-0081-y>
  20. Clevers H (2013) The intestinal crypt, a prototype stem cell compartment. *Cell* 154(2):274–284. <https://doi.org/10.1016/j.cell.2013.07.004>
  21. Barriga FM, Montagni E, Mana M, Mendez-Lago M, Hernando-Momblona X, Sevillano M, Guillaumet-Adkins A, Rodriguez-Esteban G, Buczacki SJA, Gut M, Heyn H, Winton DJ, Yilmaz OH, Attolini CS, Gut I, Batlle E (2017) Mex3a marks a slowly dividing subpopulation of Lgr5+ intestinal stem cells. *Cell Stem Cell* 20(6):801–816



- e807. <https://doi.org/10.1016/j.stem.2017.02.007>
22. von Furstenberg RJ, Buczacki SJ, Smith BJ, Seiler KM, Winton DJ, Henning SJ (2014) Side population sorting separates subfractions of cycling and non-cycling intestinal stem cells. *Stem Cell Res* 12(2):364–375. <https://doi.org/10.1016/j.scr.2013.10.012>
  23. van der Flier LG, Haegebarth A, Stange DE, van de Wetering M, Clevers H (2009) OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* 137(1):15–17. <https://doi.org/10.1053/j.gastro.2009.05.035>
  24. van der Flier LG, van Gijn ME, Hatzis P, Kujala P, Haegebarth A, Stange DE, Begthel H, van den Born M, Guryev V, Oving I, van Es JH, Barker N, Peters PJ, van de Wetering M, Clevers H (2009) Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* 136(5):903–912. <https://doi.org/10.1016/j.cell.2009.01.031>
  25. Muñoz J, Stange DE, Schepers AG, van de Wetering M, Koo BK, Itzkovitz S, Volckmann R, Kung KS, Koster J, Radulescu S, Myant K, Versteeg R, Sansom OJ, van Es JH, Barker N, van Oudenaarden A, Mohammed S, Heck AJR, Clevers H (2012) The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent ‘+4’ cell markers. *EMBO J* 31(14):3079
  26. Potten CS, Kovacs L, Hamilton E (1974) Continuous labelling studies on mouse skin and intestine. *Cell Tissue Kinet* 7(3):271–283
  27. Potten CS (1975) Kinetics and possible regulation of crypt cell populations under normal and stress conditions. *Bull Cancer* 62(4):419–430
  28. Potten CS (1977) Extreme sensitivity of some intestinal crypt cells to X and gamma irradiation. *Nature* 269(5628):518–521. <https://doi.org/10.1038/269518a0>
  29. Munoz J, Stange DE, Schepers AG, van de Wetering M, Koo BK, Itzkovitz S, Volckmann R, Kung KS, Koster J, Radulescu S, Myant K, Versteeg R, Sansom OJ, van Es JH, Barker N, van Oudenaarden A, Mohammed S, Heck AJ, Clevers H (2012) The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent ‘+4’ cell markers. *EMBO J* 31(14):3079–3091. <https://doi.org/10.1038/emboj.2012.166>
  30. Li N, Yousefi M, Nakauka-Ddamba A, Jain R, Tobias J, Epstein JA, Jensen ST, Lengner CJ (2014) Single-cell analysis of proxy reporter allele-marked epithelial cells establishes intestinal stem cell hierarchy. *Stem Cell Reports* 3(5):876–891. <https://doi.org/10.1016/j.stemcr.2014.09.011>
  31. Itzkovitz S, Lyubimova A, Blat IC, Maynard M, van Es J, Lees J, Jacks T, Clevers H, van Oudenaarden A (2011) Single-molecule transcript counting of stem-cell markers in the mouse intestine. *Nat Cell Biol* 14(1):106–114. <https://doi.org/10.1038/ncb2384>
  32. Santos AJM, Lo YH, Mah AT, Kuo CJ (2018) The intestinal stem cell niche: homeostasis and adaptations. *Trends Cell Biol* 28(12):1062–1078. <https://doi.org/10.1016/j.tcb.2018.08.001>
  33. Gerbe F, Jay P (2016) Intestinal tuft cells: epithelial sentinels linking luminal cues to the immune system. *Mucosal Immunol* 9(6):1353–1359. <https://doi.org/10.1038/mi.2016.68>
  34. Howitt MR, Lavoie S, Michaud M, Blum AM, Tran SV, Weinstock JV, Gallini CA, Redding K, Margolskee RF, Osborne LC, Artis D, Garrett WS (2016) Tuft cells, taste-chemosensory cells, orchestrate parasite type 2 immunity in the gut. *Science* 351(6279):1329–1333. <https://doi.org/10.1126/science.aaf1648>
  35. Ting H-A, von Moltke J (2019) The immune function of tuft cells at gut mucosal surfaces and beyond. *J Immunol* 202(5):1321. <https://doi.org/10.4049/jimmunol.1801069>
  36. Gerbe F, Brulin B, Makrini L, Legerverend C, Jay P (2009) DCAMKL-1 expression identifies tuft cells rather than stem cells in the adult mouse intestinal epithelium. *Gastroenterology* 137(6):2179–2180. <https://doi.org/10.1053/j.gastro.2009.06.072>
  37. de Santa BP, van den Brink GR, Roberts DJ (2003) Development and differentiation of the intestinal epithelium. *Cell Mol Life Sci* 60(7):1322–1332. <https://doi.org/10.1007/s00018-003-2289-3>
  38. van der Flier LG, Clevers H (2009) Stem cells, self-renewal, and differentiation in the intestinal epithelium. *Annu Rev Physiol* 71:241–260. <https://doi.org/10.1146/annurev.physiol.010908.163145>
  39. Kazakevych J, Sayols S, Messner B, Krienke C, Soshnikova N (2017) Dynamic changes in chromatin states during specification and differentiation of adult intestinal stem cells. *Nucleic Acids Res* 45(10):5770–5784. <https://doi.org/10.1093/nar/gkx167>
  40. Gassler N, Newrzella D, Bohm C, Lyer S, Li L, Sorgenfrei O, van Laer L, Sido B, Mollenhauer J, Poustka A, Schirmacher P, Gretz N (2006) Molecular characterisation of non-absorptive and absorptive enterocytes in human small intestine. *Gut* 55

- (8):1084–1089. <https://doi.org/10.1136/gut.2005.073262>
41. Moor AE, Harnik Y, Ben-Moshe S, Massasa EE, Rozenberg M, Eilam R, Bahar Halpern K, Itzkovitz S (2018) Spatial reconstruction of single enterocytes uncovers broad zonation along the intestinal villus axis. *Cell* 175(4):1156–1167 e1115. <https://doi.org/10.1016/j.cell.2018.08.063>
  42. Artis D, Wang ML, Keilbaugh SA, He W, Brenes M, Swain GP, Knight PA, Donaldson DD, Lazar MA, Miller HR, Schad GA, Scott P, Wu GD (2004) RELMbeta/FIZZ2 is a goblet cell-specific immune-effector molecule in the gastrointestinal tract. *Proc Natl Acad Sci U S A* 101(37):13596–13600. <https://doi.org/10.1073/pnas.0404034101>
  43. Biton M, Levin A, Slyper M, Alkalay I, Horwitz E, Mor H, Kredo-Russo S, Avnit-Sagi T, Cojocaru G, Zreik F, Bentwich Z, Poy MN, Artis D, Walker MD, Hornstein E, Pikarsky E, Ben-Neriah Y (2011) Epithelial microRNAs regulate gut mucosal immunity via epithelium-T cell crosstalk. *Nat Immunol* 12(3):239–246. <https://doi.org/10.1038/ni.1994>
  44. Noah TK, Kazanjian A, Whitsett J, Shroyer NF (2010) SAM pointed domain ETS factor (SPDEF) regulates terminal differentiation and maturation of intestinal goblet cells. *Exp Cell Res* 316(3):452–465. <https://doi.org/10.1016/j.yexcr.2009.09.020>
  45. Gregorieff A, Stange DE, Kujala P, Begthel H, van den Born M, Korving J, Peters PJ, Clevers H (2009) The ets-domain transcription factor Spdef promotes maturation of goblet and paneth cells in the intestinal epithelium. *Gastroenterology* 137(4):1333–1345 e1331-1333. <https://doi.org/10.1053/j.gastro.2009.06.044>
  46. Sato T, van Es JH, Snippert HJ, Stange DE, Vries RG, van den Born M, Barker N, Shroyer NF, van de Wetering M, Clevers H (2011) Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* 469(7330):415–418. <https://doi.org/10.1038/nature09637>
  47. Pellegrinet L, Rodilla V, Liu Z, Chen S, Koch U, Espinosa L, Kaestner KH, Kopan R, Lewis J, Radtke F (2011) Dll1- and dll4-mediated notch signaling are required for homeostasis of intestinal stem cells. *Gastroenterology* 140(4):1230–1240 e1231-1237. <https://doi.org/10.1053/j.gastro.2011.01.005>
  48. Yilmaz OH, Katajisto P, Lamming DW, Gultekin Y, Bauer-Rowe KE, Sengupta S, Birsoy K, Dursun A, Yilmaz VO, Selig M, Nielsen GP, Mino-Kenudson M, Zukerberg LR, Bhan AK, Deshpande V, Sabatini DM (2012) mTORC1 in the Paneth cell niche couples intestinal stem-cell function to calorie intake. *Nature* 486(7404):490–495. <https://doi.org/10.1038/nature11163>
  49. Ghooos Y, Vantrappen G (1971) The cytochemical localization of lysozyme in Paneth cell granules. *Histochem J* 3(3):175–178. <https://doi.org/10.1007/bf01002560>
  50. Porter EM, Liu L, Oren A, Anton PA, Ganz T (1997) Localization of human intestinal defensin 5 in Paneth cell granules. *Infect Immun* 65(6):2389–2395
  51. Basak O, Beumer J, Wiebrands K, Seno H, van Oudenaarden A, Clevers H (2017) Induced quiescence of Lgr5+ stem cells in intestinal organoids enables differentiation of hormone-producing enteroendocrine cells. *Cell Stem Cell* 20(2):177–190 e174. <https://doi.org/10.1016/j.stem.2016.11.001>
  52. Habib AM, Richards P, Cairns LS, Rogers GJ, Bannon CA, Parker HE, Morley TC, Yeo GS, Reimann F, Gribble FM (2012) Overlap of endocrine hormone expression in the mouse intestine revealed by transcriptional profiling and flow cytometry. *Endocrinology* 153(7):3054–3065. <https://doi.org/10.1210/en.2011-2170>
  53. Egerod KL, Engelstoft MS, Grunddal KV, Nohr MK, Secher A, Sakata I, Pedersen J, Windelov JA, Fuchtbauer EM, Olsen J, Sundler F, Christensen JP, Wierup N, Olsen JV, Holst JJ, Zigman JM, Poulsen SS, Schwartz TW (2012) A major lineage of enteroendocrine cells coexpress CCK, secretin, GIP, GLP-1, PYY, and neurotensin but not somatostatin. *Endocrinology* 153(12):5782–5795. <https://doi.org/10.1210/en.2012-1595>
  54. Tetteh PW, Basak O, Farin HF, Wiebrands K, Kretzschmar K, Begthel H, van den Born M, Korving J, de Sauvage F, van Es JH, van Oudenaarden A, Clevers H (2016) Replacement of lost Lgr5-positive stem cells through plasticity of their enterocyte-lineage daughters. *Cell Stem Cell* 18(2):203–213. <https://doi.org/10.1016/j.stem.2016.01.001>
  55. Jadhav U, Saxena M, O'Neill NK, Saadatpour A, Yuan GC, Herbert Z, Murata K, Shivdasani RA (2017) Dynamic reorganization of chromatin accessibility signatures during dedifferentiation of secretory precursors into Lgr5+ intestinal stem cells. *Cell Stem Cell* 21(1):65–77 e65. <https://doi.org/10.1016/j.stem.2017.05.001>
  56. Yu S, Tong K, Zhao Y, Balasubramanian I, Yap GS, Ferraris RP, Bonder EM, Verzi MP, Gao N (2018) Paneth cell multipotency induced by notch activation following injury. *Cell Stem*

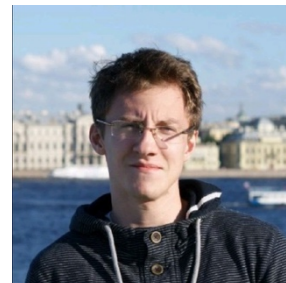
- Cell 23(1):46–59. e45. <https://doi.org/10.1016/j.stem.2018.05.002>
57. Huels DJ, Medema JP (2018) Think about the environment: cellular reprogramming by the extracellular matrix. *Cell Stem Cell* 22(1):7–9. <https://doi.org/10.1016/j.stem.2017.12.006>
58. Yui S, Azzolin L, Maimets M, Pedersen MT, Fordham RP, Hansen SL, Larsen HL, Guiu J, Alves MRP, Rundsten CF, Johansen JV, Li Y, Madsen CD, Nakamura T, Watanabe M, Nielsen OH, Schweiger PJ, Piccolo S, Jensen KB (2018) YAP/TAZ-dependent reprogramming of colonic epithelium links ECM remodeling to tissue regeneration. *Cell Stem Cell* 22(1):35–49 e37. <https://doi.org/10.1016/j.stem.2017.11.001>



# Curriculum vitae

## PERSONAL INFORMATION

Name Maxim Norkin  
Date and place of birth 24<sup>th</sup> September 1991, Zainsk,  
Russia Nationality Russian  
Address EPFL/SV, Station 19 (SV2521),  
CH-1015 Lausanne,  
Switzerland  
E-mail: [maxim.norkin@epfl.ch](mailto:maxim.norkin@epfl.ch)  
Phone: +41 21 693 0753



## PREVIOUS EDUCATION / ACADEMIC CAREER

**2017-2021** PhD at EPFL (École Polytechnique Fédéral de Lausanne), ISREC (Swiss Institute for Experimental Cancer Research)  
**2015-2017** Internship, Institute of Molecular Biology, Gutenberg University, Mainz, Germany  
**April-August 2014** Erasmus in biochemistry and molecular biology Eberhard Karls University, Tuebingen, Germany  
**2010-2015** Diploma in biochemistry. Lomonosov Moscow State University, Moscow, Russia.  
Final grade/GPA=3.7 (max.=4)

## RESEARCH EXPERIENCE

- PhD at EPFL (École Polytechnique Fédéral de Lausanne), ISREC (Swiss Institute for Experimental Cancer Research). Research in cancer stem cells, differentiation therapy, colorectal cancer, high-content high-throughput screening in organoids (April 2017 – 2021). Several courses on bioinformatics (including scRNA-seq analysis) and biostatistics. During my PhD I was in charge of designing TORNADO-seq bioinformatics pipeline for analyzing the NGS Illumina sequencing results and extracting cell identities and frequencies from bulk sample
- Research at the Institute of Molecular Biology gGmbH (IMB), Mainz, Germany. «Elucidating molecular mechanisms responsible for embryonic establishment of adult intestinal stem cells». Research in stem cells and developmental biology, mouse gut development (August 2015 – April 2017)
- Diploma research at the Lomonosov Moscow State University. «Determination of DNA bending angle induced by DNA methyltransferase SsoII» (January-May 2015)
- Internship in the Center for Plant Molecular Biology (ZMBP), Tuebingen, Germany. Research in the immune system of Arabidopsis thaliana (May-August, 2014)

## TEACHING EXPERIENCE

- 2018-2019 - Practical course at EPFL – BIO-204 Integrated laboratory in Life sciences II
- 2017 - Supervision of master student at the Institute of Molecular Biology, Mainz, Germany – Ana Patricia Jesus
- 2016 – Supervision of summer school student at the Institute of Molecular Biology, Mainz, Germany – Philipp Gusev
- 2015 – Supervision of master student at Moscow State University, Moscow, Russia – Julia Korepina

#### **PUBLICATION LIST FOR THE LAST 5 YEARS**

- Norkin M., P. Ordóñez-Morán, J. Huelsken (2021). High-content, targeted RNA-seq screening in organoids for drug discovery in colorectal cancer. **Cell Reports** 35:10902.
- Dionellis V.S., M. Norkin, A. Karamichali, G.G. Rossetti, J. Huelsken, P. Ordóñez-Morán, T.D. Halazonetis (2021). Genomic Instability Profiles at the Single Cell Level in Mouse Colorectal Cancers of Defined Genotypes. **Cancers** 13:1267.
- M. Norkin\*, C. Capdevila\*, R. I Calderon, T. Su, M. Trifas, P. Ordóñez-Morán, K. S Yan (2020). Single-Cell Studies of Intestinal Stem Cell Heterogeneity During Homeostasis and Regeneration. **Methods Mol Biol** 2171:155-167.
- L. Nigmatullina, M. Norkin, M. M. Dzama, B. Messner, S. Sayols, N. Soshnikova (2017). Id2 controls Lgr5+ intestinal stem cells progenitors specification during gut development. **EMBO J** 36:869-885.
- N. A. Timofeyeva, A. Yu. Ryazanova, M. V. Norkin, T. S. Oretskaya, O. S. Fedorova, E. A. Kubareva. (2018). Kinetic Basis of the Bifunctionality of SsoII DNA Methyltransferase. **Molecules** 2018, 23(5), 1192
- L. A. Abrosimova, E. A. Kubareva, A. Yu. Migur, A. V. Gavshina, A. Yu. Ryazanova, M. V. Norkin, T. A. Perevyazova, W. Wende, T. Hianik, L. A. Zheleznaya, T. S. Oretskaya. (2016). Peculiarities of the interaction of the restriction endonuclease BspD6I with DNA containing its recognition site. **Acta-Proteins and Proteomics**, 2016, v.1864, pp. 1072-1082.

#### **FELLOWSHIPS**

- Boehringer Ingelheim fellowship for the program “Gene Regulation, Epigenetics and Genome Stability” (August 2015 – April 2017)
- University of Tuebingen Erasmus fellowship (Germany, April-July 2014)

#### **CONFERENCES AND TALKS**

- Organoids in Cancer Research. Agora workshop, Lausanne, Switzerland, November 2021 (short talk)
- Stem Cells and Cancer, Barcelona, Spain, IRB Conference, October 2021 (short talk)
- Organoids: Modelling Organ Development and Disease in 3D Culture – Virtual. EMBO-EMBL Symposium, 2020 (e-poster)
- 8th Faculty & Staff Retreat of the Swiss Cancer Center Lausanne, Switzerland, 2019 (short talk)
- Horizons of Cancer Biology and Therapy, Lausanne, Switzerland, 2018 (poster)
- 7th Faculty & Staff Retreat of the Swiss Cancer Center Lausanne, Switzerland, 2018 (poster)
- Joint Meeting of the German and Japanese Societies of Developmental Biologists, Kiel,

- Germany, 2017 (poster)
- International young scientist conference “Mendeleev-2015” in section “bioorganic and medical chemistry”, Saint Petersburg, 2015 (talk)
  - International student, postgraduate and young scientist conference “Lomonosov”, Moscow. 2012, 2013, 2015 (talks)

## SCIENTIFIC INTERESTS

I consider the field of targeted personalized medicine as one of the most avantgarde and promising directions in the future of medicine. For my academic career, I strongly envisage translational medicine as a path. I am confident that future advances on tumor classification taking all possible factors into consideration will lead to more personalized and more effective approaches which will allow to improve the life of many cancer patients. Sophisticated computational analysis tools will be essential to reach these goals and I am grateful that I had the chance to learn these techniques and that I enjoy performing bioinformatics tasks and biostatistical analysis. For example, the whole bioinformatic pipeline in our recently published TORNADO-seq paper was done by myself.

Topics of interest:

Stem cells, cancer stem cells, cancer research, cancer metastasis, personalized medicine, PDOs, tumor microenvironment, Lgr5<sup>+</sup> cells, cancer heterogeneity, cancer resistance.