Aus dem Deutschen Krebsforschungszentrum
Abteilung Signalwege und Funktionelle Genomik
Leitung: Prof. Dr. Michael Boutros

Image-based profiling of colorectal cancer organoid models

Inauguraldissertation
zur Erlangung des medizinischen Doktorgrades

an der
Medizinischen Fakultät Heidelberg der
Ruprecht-Karls-Universität

vorgelegt von
Niklas Timon Rindtorff

aus Tübingen

2023

Dekan: Prof. Dr. Michael Boutros
Doktorvater: Prof. Dr. Michael Boutros

Contents

List of Figures

List of Tables

List of Abbreviations and Notation

Abbreviations

5FU	5-fluorouracil
ACN	Acetonitrile
ANOVA	Analysis of Variance
ARD	Automatic Relevance Determination
ATP	Adenosine triphosphate
AUC	Area under the curve
AUROC	Area under the receiver operating characteristic
BCA	Bicinchoninic Acid Assay
BMP	Bone Morphogenic Protein
CE	Cholesterol Ester
Cer	Ceramide
Chol	Cholesterol
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
CMS	Consensus Molecular Subtype
CRC	Colorectal cancer
CRIS	CRC Intrinsic Subtypes
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTG	Cell Titer Glo Viability Assay
DAG	Diacylglycerol
DAPI	4 ,6-Diamidino-2-Phenylindole
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DPBS	Dulbecco's Phosphate Buffered Saline
DTT	Dithiothreitol
ENA	Organoid Medium containing Egf, Noggin, and A83-01
ENAS	Organoid Medium containing Egf, Noggin, A83-01, and SB202190
FDR	False discovery rate
FOLFIRI	Folinic Acid, 5-Fluorouracil, and Irinotecan

FOLFOX	Folinic Acid, 5-Fluorouracil, and Oxaliplatin
FOLFOXIRI	Folinic Acid, 5-Fluorouracil, Oxaliplatin, and Irinotecan
GSEA	Gene Set Enrichment Analysis
HDF5	Hierarchical Data Format Version 5
HEPES	4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid
Hex2Cer	Dihexosylceramide
HexCer	Hexosylceramide
HRP	Horse Radish Peroxidase
ICA	Independent Component Analysis
LPC	Lysophosphatidylcholine
MCR	Mutation Cluster Region
MOFA	Multi-omics Factor Analysis
MSI	Microsatellite instability
MSS	Microsatellite stability
NMF	Non-negative Matrix Factorization
PA	Phosphatidate
PAINS	Pan Assay INterference compoundS
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline containing 0.1% (v/v) Triton X-100
PC	Phosphatidylcholine
PCA	Principal Component Analysis
PDO	Patient Derived Organoid
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PS	Phosphatidylserine
RIPA	Radioimmunoprecipitation Assay Buffer
RMA	Robust Multiarray Average
sgRNA	Single guide RNA
siRNA	small interfering RNA
SM	Sphingomyelin
TAG	Triacylglycerol
TRITC	Tetramethylrhodamine
UICC	Union for International Cancer Control
UMAP	Uniform Manifold Approximation and Projection
WENRAS	Organoid Medium containing Wnt3a, Egf, Noggin, R-Spondin, A83-01, and SB202190

Notation

ε_m	Error terms of modality m
$\mathrm{W_m}$	Factor weight matrix (also called loading matrix) of modality m
X	Matrix
X^+	Pseudoinverse of a Matrix
X^T	Matrix Transpose
$\mathrm{Y_m}$	Measured data matrix of modality m
Z	Factor score matrix

1 Introduction

1.1 Disclosure

Parts of this introduction, have been adapted from joint first-author publications, and were originally written by myself: *Wnt signaling in cancer* (T. Zhan et al. 2017) and *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). In particular, the section ”canonical Wnt signaling during colorectal cancer pathogenesis” was adapted from (T. Zhan et al. 2017) and the overview figure as well as its figure legend in the ”thesis aims” section was taken from (Betge et al. 2022).

1.2 Colorectal Cancer

Colorectal Cancer is among the three most common and lethal forms of cancer in the developed world, accounting for ca. 10% of all cancer incidence and mortality (Sung et al. 2021). The majority of colorectal cancer cases are diagnosed after the age of 50, are located in a region spanning the rectum and sigmoid colon, and develop from macroscopic precursor lesions referred to as adenomas over the course of ca. 10 years (Cho et al. 1992). Next to patient age, a sedentary lifestyle including obesity, diabetes, consumption of red and processed meats, alcohol, and smoking constitute risk factors (Sung et al. 2021). The disease is classified by the Union for International Cancer Control (UICC) into five stages ranging from *carcinoma-in-situ* (patch of malignant cells

that has not yet breached the basal lamina of the intestinal mucosa), to stages 1 (malignant cells have breached the basal lamina), stage 2 (malignant cells have spread beyond the large muscle layer surrounding the colon mucosa), stage 3 (malignant cells have spread to loco-regional lymph nodes), and stage 4 (metastatic disease) (Van Cutsem et al. 2016).

Among the most effective interventions to reduce the incidence and mortality rate of colorectal cancer is the removal of macroscopic lesions during preventative colonoscopy (Nishihara et al. 2013). In case the disease has further progressed, therapeutic options include oncological resection coupled with adjuvant chemotherapy (available for UICC 2, recommended for UICC 3) based on the DNA replication inhibitors 5-Fluoruracil (5FU) and Oxaliplatin (FOLFOX). Additional treatment options exist for patient subgroups such as rectal cancer patients (neoadjuvant radiochemotherapy) and frail patients (5FU instead of FOLFOX for adjuvant chemotherapy). In case of metastatic disease (UICC 4) the treatment depends on the tumor's molecular characteristics: Options for *KRAS* wild type disease include anti-Egfr antibodies (i.e., Cetuximab) in combination with chemotherapy (FOLFOX or FOLFIRI) (Van Cutsem et al. 2016). Options for *BRAF* mutant disease include triple therapy (FOLFOXIRI (Van Cutsem et al. 2016)), or combined Egfr-, Mek- and Braf-inhibitor treatment (Kopetz et al. 2019). Options for tumors with microsatellite instability (MSI+) include PD-1 immune checkpoint inhibition (i.e., Pembrolizumab, anti-PD1) (André et al. 2020). Most therapeutic options can be combined with angiogenesis inhibiting antibodies (Bevacizumab, anti-VEGFR), especially when patients cannot receive a full treatment protocol. Additional lines of therapy can also include agents like Regorafenib and Triflouridin/Tipiracil (Van Cutsem et al. 2016).

1.3 Colorectal Cancer Pathogenesis

1.3.1 The colon crypt

The sequence of molecular events leading to colorectal cancer are governed by the tissue's stem cell biology (Clevers 2011). The colon stem cell niche, or crypt, is the source of all epithelial cells lining the colon. Similar to the small intestine, Lgr5+ intestinal stem cells are located at the bottom of the crypt and continuously renew the epithelium by proliferating and, through displacement, pushing newly formed cells towards the colon's lumen. This architecture serves multiple purposes, including protection of stem cells and the control of cell fate decisions across the epithelium (Clevers 2013). Multiple developmental pathways, especially Wnt signaling, Notch, Bmp and ERK-MAPK signaling, govern cell identity and growth rate in the intestinal niche (Gehart et al. 2019). The concentration of ligands for most of these signaling pathways is organized in gradients along the crypt-lumen axis. For example, the concentration of stem cell identity maintaining Wnt ligands (Wnt signaling) and growth-rate controlling Egf ligands (ERK-MAPK signaling), decreases as cells are pushed outside the crypt (Sasaki et al. 2016). In contrast, the effect of Bmp ligands, which promote cell differentiation, increases as the concentration of BMP signaling inhibitors (i.e., Noggin) derived from basal mesenchymal cells decreases (He et al. 2004). As a

cell is pushed outside the crypt by a continuous stream of newly formed, proliferating cells, the signals it receives change -leading to cellular differentiation. Similarly, if enterocyte-progenitors are moved back into the crypt, the ambient signaling leads to a dedifferentiation towards an intestinal stem cell state (Tetteh et al. 2016).

Given the spacial confinement of growth- and stemness-stimulating signals, the crypt architecture also leads to a protection against tissue damage. At the bottom of the crypt, a neutral competition of proliferating intestinal stem cells leads to the removal of damaged cells that show a reduced proliferation rate relative to the surrounding stem cells (Snippert et al. 2010). Given this neutral competition and the dependence on external signals to proliferate, dysfunctional or transformed cells have a higher likelihood of being removed from the niche unless they acquire a set of molecular alterations that render them independent of niche signals and increase their growth rate.

1.3.2 The adenoma carcinoma sequence

Colorectal cancer is a disease marked by the accumulation of genetic events in somatic cells, which over time lead to uncontrolled growth and spread beyond the tissue of origin. Both structural and functional genomics experiments have helped identify functional genetic events, so-called drivers, that cause the transition from controlled to malignant cellular state. A classic model for the sequence of functional events leading to colorectal cancer is known as the adenoma-carcinoma sequence (Vogelstein et al. 1988). It starts with the activation of the Wnt signaling pathway through loss of the tumor suppressor *APC* and is followed by the activation of the ERK-MAPK signaling pathway through mutations of the *KRAS* or *BRAF* oncogene. Further mutations activating PI3K signaling (i.e., *PI3K* oncogene, *PTEN* tumor suppressor) and inhibiting TGF-beta (i.e., *SMAD4* tumor suppressor) and p53-signaling (i.e., TP53 tumor suppressor) complete the transformation from adenoma to invasive carcinoma (Fearon 2011).

When simplified, colorectal cancer forms along the adenoma-carcinoma sequence through (I) a common chromosomal instability (ca 80%) or (II) a DNA-mismatch repair deficiency (ca. 20%) associated mechanism (Markowitz et al. 2009) and (Pancione et al. 2012). These two trajectories of tumor development have been associated with characteristic clinical, pathological, and molecular properties.

Tumors of the common chromosomal instability phenotype present themselves early as common, non-serrated polyps in the rectum and sigmoid colon (Markowitz et al. 2009). Most tumors are microsatellite stable and have frequent *APC* truncating mutations coupled with *KRAS* activating mutations.

In contrast, tumors of the DNA-mismatch repair phenotype are frequently present as serrated polyps and more likely to be located in the right colon (Markowitz et al. 2009). These tumors are often marked by a CPG-island methylation phenotype (CIMP) that can lead to the hypermethylation of tumor suppressors and DNA-mismatch repair genes (Ogino et al. 2009). Consequently, these tumors have a higher proportion of microsatellite instability, higher mutational burden, higher immunogenicity, and thus a higher immune-cell infiltration. APC mutations are less frequent and if present more likely to be missense than truncating (Borowsky

et al. 2018). Mutations of the BRAF oncogene are more frequently found than *KRAS* mutations.

Independent of the mechanism, the sequential accumulation of functional genetic events renders colorectal cancer cells progressively independent of the organism's tissue control mechanisms and thus enables the uncontrolled, independent growth of cells which defines late-stage cancer.

1.3.3 The role of APC and KRAS during colorectal cancer pathogenesis

Genetic alterations in the tumor suppressor *APC* and the proto-oncogene *KRAS* are among the two most frequent mutations found in across colorectal cancer cases (Markowitz et al. 2009). Alterations in the two genes are highly correlated and found already at early stages (Mina et al. 2017). What defines the interaction between the two colorectal cancer drivers, is, however, still an open question (Parsons et al. 2021). In the following section the mechanism of Wnt signaling and ERK-MAPK signaling are briefly introduced, as these two pathways are directly affected by Apc loss and Kras activation, respectively, and both have a regulatory function within the colon stem cell niche (Gehart et al. 2019).

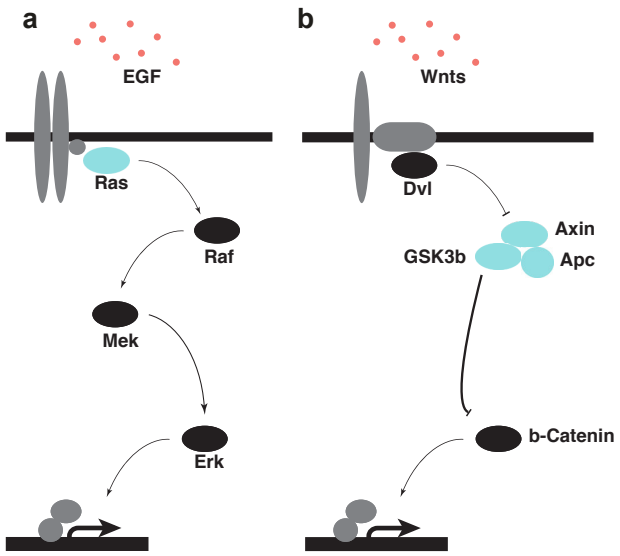

Figure 1.1: Simplified illustration of Egf dependent ERK-MAPK signaling and canonical Wnt signaling a Egf dependent signaling cascade. Ras is highlighted in blue. **b** canonical Wnt signaling cascade. The destruction complex, including Apc, is highlighted in blue.

canonical Wnt signaling during colorectal cancer pathogenesis

Members of the Wnt signaling pathway were first identified as regulators of embryonic development (R. P. Sharma et al. 1976; Nüsslein-Volhard et al. 1980) and are controlling adult stem cell identity across tissues. As a result, members of the pathway are frequently altered during cancer initiation (T. Zhan et al. 2017). In canonical Wnt signaling, absence of Wnt ligands leads to phosphorylation of beta-catenin by the destruction complex, which contains the scaffold protein Axin, the large disordered protein Apc (Adenomatous polyposis

coli, *APC*) and the kinases Gsk3-beta as well as Casein kinase (T. Zhan et al. 2017) (Figure 4.11 b). In this state, beta-catenin is phosphorylated by Gsk3-beta, ubiquitinated by beta-TrCP and subsequently targeted for proteasomal degradation. In the absence of nuclear beta-catenin, the transcriptional repressive complex containing TCF/LEF and transducing-like enhancer protein (TLE/Groucho) recruits histone deacetylases to repress target genes.

The role of Wnt signaling during colorectal cancer development is well established (Polakis 2007). Here, loss of *APC* is the most frequent driver of Wnt signaling in colorectal cancer and can be found in about 80% of colorectal cancer patients (Fearon 2011; T. Zhan et al. 2017). Both genetic mouse models and *in-vitro* experiments with human colon organoids (introduced below) have demonstrated that loss of function of *APC* (*Apc* in mice) is sufficient to cause adenoma formation (Moser et al. 1995), which is marked by Wnt-ligand independent proliferation (Matano et al. 2015; Drost et al. 2015). Further, in the context of additional functional events in *Kras* and *Trp53*, *Apc* loss is necessary to sustain invasive tumor growth of *in-vivo* cancer models (Dow et al. 2015; Sakai et al. 2018).

ERK-MAPK signaling during colorectal cancer pathogenesis

The extracellular-signal-regulated (ERK) mitogen-activated protein kinase family (MAPK) is one of three major MAPK families, together with the JNK (c-jun N-terminal kinase or stress-activated protein kinases) and MAPK14 group of protein kinases (W. Zhang et al. 2002). These kinases and the signaling cascades they form play a major role in (I) integrating external proliferative signals, (II) reacting to stress or ambient cytokines and (III) protecting cells from apoptosis, respectively (Fang et al. 2005). The ERK-MAPK signaling cascade itself is constituted by a set of protein kinase subfamilies, Ras (encoded by i.e., *KRAS*, *NRAS*, *HRAS*), Raf (encoded by *BRAF*, *CRAF*, *ARAF*), Mek (encoded by *MEK1*, *MEK2*) and Erk (*ERK1*, *ERK2*) which transmit extracellular mitogenic stimuli in a step-wise phosphorylation-based signaling cascade towards downstream transcription factors (Figure 4.11 a).

In the colon crypt, the ERK-MAPK signaling cascade and its members are key regulators of cell proliferation and tissue growth rate (Gehart et al. 2019). Blockade or removal of Egf, a canonical ligand of the signaling cascade's upstream receptor (encoded by *EGFR*), leads to a cell cycle arrest of proliferating Lgr5+ cells while not affecting their cellular state (Basak et al. 2017). As a result, Lgr5+ cells can re-enter proliferation as soon as ERK-MAPK signaling is restored.

In colorectal cancer, the Ras kinase subfamily members are mutated in about 36% of cases (Fang et al. 2005). Next to Ras, Raf mutations can also be found in around 10% of colorectal cancers (Fang et al. 2005). Of note, mutations of *KRAS* and *BRAF* occur mostly in a mutually exclusive pattern (introduced above) (Fang et al. 2005; Mina et al. 2017). According to the Vogelstein model of colorectal cancer initiation (Fearon 2011), activating mutations of *KRAS* takes place early during cancer development, more specifically, after loss of *APC*. *in-vivo* mouse models of the oncogenic $Kras^{G12D/+}$ show that this functional event is not sufficient to cause adenoma formation in isolation, but instead leads to *Cdkn2a*-dependent oncogene-induced senescence (Bennecke

et al. 2010). In the context of *in-vivo Apc* loss of function ($Apc^{\Delta 716}$), however, the $Kras^{G12D/+}$ genotype leads to an increased number of adenomas and reduced survival (Sakai et al. 2018). Of note, the combination of $Apc^{\Delta 716}$ and $Kras^{G12D/+}$ in colon epithelial cells is, however, not sufficient to cause significant invasive and metastatic disease *in-vivo* (Sakai et al. 2018; Boutin et al. 2017), which is only seen in triple mutant genotypes (i.e., $Apc^{\Delta 716}$ / $Kras^{G12D/+}$ / $Tgfbr2^{-/-}$, or $Apc^{-/-}$ / $Kras^{G12D/+}$ / $Trp53^{-/-}$). Here, in the context of an $Apc^{-/-}$ and $Trp53^{-/-}$ genotype, $Kras^{G12D/+}$ is necessary and sufficient to lead to invasive and metastatic disease *in-vivo*. Conversely, removal of the $Kras^{G12D/+}$ in fully formed tumors *in-vivo* leads to a regression back towards an adenoma phenotype (Boutin et al. 2017).

1.4 Organoids

Organoids are three-dimensional cell culture models from primary adult tissue. Intestinal organoids develop from Lgr5+ adult stem cells and were first isolated from the small intestine of mice (Sato, Vries, et al. 2009). Subsequently, further organoid models across tissue-types and species have been developed. These include organoids from healthy and cancerous human large intestine (Sato, Stange, et al. 2011), pancreas (Driehuis et al. 2019), mammary epithelium (L. Zhang et al. 2017; Sachs et al. 2018) and the hepatobiliary system (Huch et al. 2013). The number of tissues that can be modeled through organoids has been continuously increasing since the first publication of this methodology.

1.4.1 Isolation and Culture of Organoids

Culturing organoids from primary cells requires the addition of specific tissue-dependent growth factors and the embedding of cells in 3D hydrogels (Merker et al. 2016). In the case of colon organoids, the necessary growth factors are inspired by signaling cues available in the intestinal stem cell niche (discussed above): Wnt and R-spondin ligands to activate and maintain canonical Wnt signaling; EGF ligands to stimulate ERK-MAPK signaling, and Noggin ligands to inhibit the differentiating effects of the BMP signaling cascade (Sato and Clevers 2013). When combined with inhibitors of TGF-beta and p38 mediated signaling, these growth factors enable organoid formation and continuous proliferation *in-vitro*. In line with the niche-focused model of colon cancer formation outlined previously, patient derived colorectal cancer organoids proliferate in simpler conditions in a niche-independent manner that corresponds to their genetic state (i.e., *APC* mutant colon organoids grow in conditions lacking Wnt and R-spondin ligands) (Fujii, Shimokawa, et al. 2016). Not only the high isolation efficiency of up to 80%, but also the high genetic and transcriptomic correlation with their tissue of origin, have made organoids an increasingly popular translational research model (Pauli et al. 2017). Both the high isolation efficiency, and the molecular representation of the tissue of origin are likely a result of the culture conditions which are emulating the tissue-specific stem cell environment.

1.4.2 Diagnostic use of Colorectal Cancer Organoids

Given the benefits of the methodology, patient derived organoids are being trialed as predictive models for personalized treatment recommendation in colorectal cancer care (Wetering et al. 2015; Vlachogiannis et al. 2018; Ganesh et al. 2019; Salo N. Ooft et al. 2019; Yao et al. 2020). In these assays, personalized organoid models are generated and then treated with therapeutic options *in-vitro*. While robust reports on the predictive validity of organoids for progression-free survival are still missing, current results for response rate models show noteworthy predictive validity for Irinotecan, FOLFIRI, and neoadjuvant radiation therapy. A pooled analysis of the previously published literature recently summarized the aggregate sensitivity and specificity for organoid based treatment response rate prediction at 0.81 and 0.74, respectively (conservative single-point AUROC estimate: 0.78) (Wensink et al. 2021; J. Zhang et al. 2005). Despite the potential of patient derived organoids to guide clinical decision-making, the method is not ready for use in a clinical context, given a high biopsy-to-evaluation dropout rate and a long turnaround time being the primary technical hurdles. A recent failed single-arm prospective clinical trial for organoid-based last-line treatment recommendations reported only 6 treated patients out of 61 patients who were initially included in the study (>90% dropout rate, 23 participants lost due to organoid establishment failure, 11 participants lost due to disease progression under standard-of-care before a recommendation could be made) (S. N. Ooft et al. 2021). Further improvements of reagents, isolation procedures, and *in-vitro* protocols might lead to lower dropout rates and turnaround times to enable further investigation of organoid-based diagnostics for personalized colorectal cancer care.

1.4.3 Therapeutic Discovery using Colorectal Cancer Organoids

Next to diagnostics, organoid have been evaluated as models for therapeutic discovery. Patient derived organoids can be processed in high-throughput small molecule screens with, ATP-based, cell viability readouts (Wetering et al. 2015; Boehnke et al. 2016). In addition, organoids are amenable to efficient genetic engineering using CRISPR (Matano et al. 2015; Drost et al. 2015). This opens up the possibility of evaluating the effect of therapeutic candidates against precisely defined, engineered genetic disease states. Next to cell viability measurements, readouts such as systematic microscopy have been used to study the self-organization of healthy intestinal organoids (Lukonin et al. 2020). However, a large morphological heterogeneity of organoids as well as their three-dimensional, multicellular architecture present a challenge to these methods.

1.5 Profiling Experiments

Profiling experiments are biological experiments in which high dimensional phenotypes are observed of *in-vitro* models that are subjected to multiple treatments. Examples of such methods include Transcriptome Profiling Experiments (i.e., Perturb-Seq (Dixit et al. 2016), L1000 Profiling (Subramanian et al. 2017)) as well as Image-based Profiling Experiments (Juan C Caicedo et al. 2016) (described below). These experiments generate

data containing (1) multiple features (i.e., transcript count or morphological feature) for (2) multiple treatments (i.e., small molecules or CRISPR perturbations) and (3) multiple experimental units (i.e., multiple genetically defined cell states, multiple organoids). Of note, the presence of high dimensional features for multiple treated states is what separates profiling experiments from observational (i.e., many features, such as single-cell RNA-Seq, across multiple clinical samples, but no prospective treatment) or focused screening projects (i.e., one feature, such as cell viability, across multiple treatments).

1.5.1 Image-based Profiling

Image-based profiling is an unbiased high-throughput microscopy based profiling method in which *in-vitro* models are systematically treated with small molecules or genetic perturbations (i.e., siRNA) to observe their treatment-induced phenotype (Carpenter 2007; Juan C Caicedo et al. 2016). A hallmark of image-based profiling is the emphasis on capturing unbiased morphological information from treated cells. Different from ATP-based cell viability readouts (where every well on a plate corresponds to one scalar measurement), every experimental unit (i.e., individual cell, organoid) is detected and multiple descriptive features (i.e., size, staining intensity, texture) are measured. As a result, every treatment condition (well on a plate) corresponds to a table with multiple biological units as rows and multiple features as columns, and thereby enables the ability to investigate biological heterogeneity within a given treatment condition. In contrast to screening experiments which are focused on one particular phenotype of interest, image-based profiling experiments are using general purpose staining protocols (i.e., DNA, Actin, Mitochondria, and the cellular membrane). Given the large amount of collected information from such experiments, analyzing their results is a common challenge and subject of continuous method development (Chandrasekaran et al. 2021).

The image-based profiling method is commonly used in functional genomics (Laufer et al. 2013; Billmann et al. 2016) and translational biomedical research (Gibson et al. 2015; Moshkov et al. 2023). While the exact approaches differ, a set of common steps in image-based profiling analysis exist: (1) Image Analysis, (2) Phenotype Modeling and (3) Novelty Detection. During Image Analysis, raw microscopy images are processed using manual or learned feature extraction methods to generate a low-dimensional representation of each treatment-induced phenotype. Ideally, such representations are comparable across contexts, capture biological differences, and suppresses technical artifacts. Examples for different methods include classic texture feature calculations followed by Principal Component Analysis (PCA) (Juan C. Caicedo et al. 2017) as well as modern approaches such as end-to-end self-supervised learning with a contrastive-loss (Perakis et al. 2021). During modeling, a set of conditions that are well annotated are observed and the low-dimensional representations of their treatment-induced phenotypes are modeled as a function of these treatment conditions. For example, the phenotype of siRNA treated Drosophila cells has been modeled as an effect of two combined siRNAs using a linear model (Billmann et al. 2016), and the phenotype of small-molecule treated human cells has been modeled as an effect of a single disease allele using a boosting-based model (Gibson et al. 2015). During Novelty

Detection, a set of treatment conditions of interest are identified based on their deviations from the learned model. For example, the phenotype of Drosophila cells might change more drastically than anticipated under a linear model after treatment with two siRNAs which target genes in the same signaling pathway (indicating a form of epistasis) (Billmann et al. 2016). Analogously, a small-molecule induced phenotype of human cells might be misclassified as wild type, although the treated cells are bearing a disease allele (indicating a potential small molecule candidate inhibitor of the disease associated molecular mechanism) (Gibson et al. 2015). Once identified, these novelties can be further validated in independent experiments to identify potential new biological mechanisms or therapeutic candidates.

1.5.2 Matrix Factorization

The natural representation of data generated by high-throughput experiments, including profiling experiments, are matrices containing the measured treatment condition (rows) and their observed features (columns). A common challenge in large biological data analysis is the identification of interpretable biological processes that underlie the observed data (Stein-O'Brien et al. 2018). The basic goal of matrix factorization methods is to identify a small number of underlying factors (i.e., biological processes) within high-dimensional data.

$$Y = ZW^T + \varepsilon \tag{1.1}$$

In this process, a data matrix Y, consisting of n samples and m features, is decomposed into two lower-dimensional matrices:

1. The factor score matrix Z, which has dimensions $n \times k$, where k represents the number of latent factors. Each row in Z corresponds to a sample, and each column corresponds to a latent factor.
2. The factor weight matrix W^T, which has dimensions $k \times m$. Here, each row corresponds to a latent factor, and each column corresponds to a feature. This matrix is also referred to as the loading matrix.

The factor score matrix Z can be understood as a learned, lower-dimensional representation of the data with k dimensions. The choice of k (the number of latent factors) ideally strikes a balance between simplicity and comprehensiveness and can be algorithmically determined.

Different matrix factorization methods exist, with Principal Component Analysis (PCA), non-negative matrix factorization (NMF) and Independent Component Analysis (ICA) being among the most common (Stein-O'Brien et al. 2018). Each of these methods performs the basic decomposition outlined above, but applies different constraints during the decomposition (Stein-O'Brien et al. 2018): In PCA, the identified factors are orthogonal to each other and are ranked by the amount of variance they explain within the data. In NMF, both the factor score and weight matrix must be non-negative, which aides interpretation. In ICA, the factors are statistical independent and non-Gaussian, which is useful for the analysis of time-course data, such as audio or electrical

activity. Each of these three popular methods do not include further assumptions, such as a dedicated measurement error ε.

In case the data generating process is better understood, additional assumptions can be introduced. For example, directly modeling the measurement error for each feature ε and applying sparsity constraints to the matrix factorization are options. If such additional assumptions are made, the method is more commonly referred to as a form of factor analysis (Klami et al. 2014).

1.5.3 Extending Matrix Factorization to multiple experimental Views

Many physical processes of scientific interest can be observed through more than one method or "view". For example, a patient tissue sample can be characterized through DNA sequencing, RNA sequencing and lipid abundance measurements. Finding a small number of underlying factors across high-dimensional data coming from multiple views is thus an increasingly frequent problem in biological data analysis and is the focus of the multi-view representation learning field (Li et al. 2019). In group factor analysis methods (Virtanen et al. 2012; Klami et al. 2014), which include multi-omics factor analysis (MOFA, introduced further below) (Argelaguet, Velten, et al. 2018; Argelaguet, Arnol, et al. 2020), the factor analysis described above is extended to multiple data views (also referred to as modalities): A set of m data matrices Y_m is decomposed into a single factor score matrix Z and a corresponding set of m weight matrices W_m^T and error terms ε_m.

$$Y_m = ZW_m^T + \varepsilon_m \tag{1.2}$$

Factors that are identified using this approach describe patterns within the data that can span multiple different data modalities. Next to simplifying the interpretation of complex data, factors that have been learned to represent the structure of measurements across multiple modalities with group factor analysis have been shown to enable more reliable similarity estimation and missing data prediction than alternative factor analysis methods or simply raw features taken from the measured data (Klami et al. 2014).

Multi-omics matrix factorization (MOFA) builds on group factor analysis to learn a multi-view representation from data generated by omics-methods (Argelaguet, Velten, et al. 2018; Argelaguet, Arnol, et al. 2020). While other methods, such as classical group factor analysis, can be used to learn a representation across multiple modalities, MOFA has been developed with design choices that make it particularly useful for complex biological data:

1. feature sparsity -the model reduces the number of features assigned to a given factor
2. factor sparsity -the model reduces the number of factors that are assigned to a given modality.
3. modality-specific priors -the model uses dedicated priors for continuous, binary and count-based modalities.

4. missing data handling -the model can be trained on incomplete observations.

5. speed -the model is fast to run compared to alternative methods, such as group factor analysis.

In the context of the work presented in this thesis, multi-omics factor analysis (MOFA) was used to learn multi-modality representations for both patient derived and genetically engineered colon organoids. The modeled data comprised modalities such as organoid morphology, transcript expression, protein-, and lipid-abundance. The identified factors corresponded to major axes of biological variation and were used to interpret results from image-based profiling experiments as outlined below.

1.5.4 Using Matrix Factorization to interpret observations

While matrix factorization methods are commonly used to learn the intrinsic structure of high-dimensional data within an exploratory context, it can also be used to guide decision-making by predicting similarities between observations. In the latter context, a well-annotated subset of the data, here referred to as the support set, is first used to learn factors. In a second step, another set of observations which are to be investigated, here referred to as the query set, are projected into the learned representation to estimate their properties. The ”support set” and ”query set” terminology is adapted from the field of meta-learning -a part of the machine learning literature focused on scenarios when only few well-annotated observations are available (Hospedales et al. 2020), a common situation in biomedical research and therapeutic discovery.

To estimate the factor scores for new observations in the query set, the previously identified factor weight matrix is used.

$$Z_{\text{new}} = Y_{\text{new}} W^{+} \tag{1.3}$$

Instead of decomposing the query data matrix Y_{new} into a factor score matrix and factor weight matrix, as done during the initial learning, the query data matrix Y_{new} is multiplied with an approximated inverse of the factor weight matrix W^{+} to estimate the corresponding factor scores Z_{new}.

In practice, scenarios in which such a multistep approach is applied include, for example, recommender systems (Koren et al. 2009). Here, a new observation is made and put in context with existing observations to estimate its properties and guide decision-making (i.e., which experiment to run, or a what product to recommend). In this thesis, a related approach is used for multi-view profiling of *in-vitro* organoid models. This approach is useful, especially when cost differences between data modalities exist and treatment conditions with an effect of interest are rare: Instead of observing and modeling all possible treatments with multiple views at once (i.e., transcript expression, morphology, protein abundance), a representative subset, the support set, is first selected and fully observed to learn a multi-view model. This subset includes well annotated conditions, such as untreated *in-vitro* models of healthy and diseased tissue. In a second step, numerous treatment conditions, the query set, are observed with only one modality. This modality is affordable and sufficiently informative (i.e.,

microscopy images). In the third and final step, the single-view observations are projected using the previously defined multi-view model. Provided the support set was representative enough, treatment conditions that show an effect of interest (i.e., a phenotype deviates from the expectation) are identified as novel and can be further characterized by making observations from additional views.

1.6 Aims of Thesis

Organoids are high fidelity three-dimensional *in-vitro* models that can be generated from cancer and healthy tissue. Genetically engineered organoid models demonstrate strong morphological differences. Similarly, patient derived organoids show pronounced morphological heterogeneity between donors, as well as within donors *in-vitro*. In both cases, the factors determining this heterogeneity are not well understood. In addition, it is not clear whether these morphological differences are associated with differences in small molecule treatment sensitivity *in-vitro* and whether morphological states can be manipulated through small molecule treatment.

The aim of this thesis is to understand sources of morphological variation in colorectal cancer organoid models. To this end, organoid models were generated and image-based profiling was performed to identify multi-view factors explaining both morphological and molecular heterogeneity (Figure 4.8). In addition, the relationship between factors and small molecule treatment sensitivity was investigated. Here, besides measuring relative differences in treatment sensitivity, the aim was to also identify non-lethal treatments that changed organoid molecular and morphological state along these newly identified axes of variation.

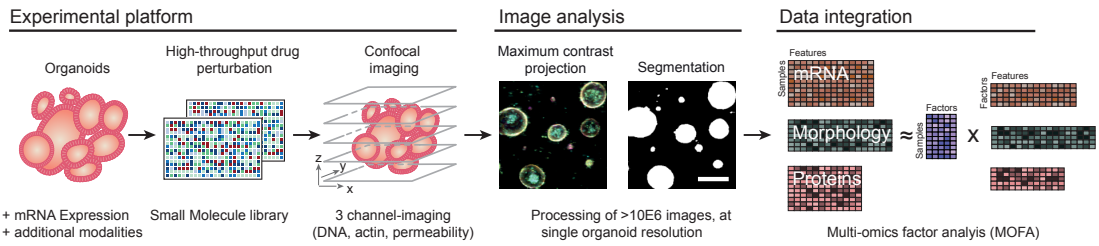

Figure 1.2: Overview of organoid multi-view profiling experiments Organoids were a) isolated from endoscopic biopsies from patients with colorectal cancer or b) derived from mouse tissue and modified using CRISPR and conditional allele expression. Organoids were dissociated and evenly seeded in 384-well plates before perturbation with small molecules (ca. 800 and 1700 conditions for patient derived and mouse organoids, respectively). After treatment, high-throughput fluorescence microscopy was used to capture the morphology of organoids. The multichannel (DNA, beta-actin, cell permeability) 3D imaging data was projected, segmented, and descriptive features were extracted to quantify potential drug-induced phenotypes. Untreated organoid morphology, organoid size and drug activity scores were integrated with additional modalities, such as mRNA expression and DNA sequencing data in a Multi-Omics Factor Analysis (MOFA) to increase interpretability of organoid variation. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022)

To understand sources of heterogeneity for both models of fully developed and precursor states of colorectal cancer, the thesis is separated into two parts. In the first part, 11 patients derived colorectal cancer organoids were isolated and image-based profiling was performed. In the second part, healthy colon organoids were genetically engineered to generate models of wild type, $Apc^{-/-}$, $Kras^{G12D/+}$, and $Apc^{-/-}$ / $Kras^{G12D/+}$ genotype and profiled similarly. In both cases, multi-view factors and their characteristic treatment sensitivities were identified and small molecules with non-toxic, factor-specific effects were investigated.

To summarize, the three central questions this thesis aims to answer are:

1. What are the molecular factors determining organoid morphology in colorectal cancer models?
2. Are these axes of variation associated with differences in small molecule treatment sensitivity?

3. Can small molecules shift organoids in their molecular state along these factors? If so, what are the mechanisms through which such small molecules exert their effect?

These questions were addressed by (1) establishing image-based profiling for colorectal cancer organoid models, (2) multi-view modeling of organoid morphology together with their molecular state, and (3) using the resulting multi-view factors to interpret differences as well as to identify small molecule treatments that caused molecular state changes along them.

2

Materials and Methods

2.1 Disclosure

The materials and methods chapter, including text and figures, has been adapted from a joint first-author publication, *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). Descriptions of core facility services are quoted and indicated as such.

2.2 Patients

All patients were identified at the University Hospital Mannheim, Mannheim, Germany. Untreated patients with a new diagnosis of colorectal cancer were included in this study. Biopsies were obtained from their primary tumors and adjacent normal tissue via forceps-based endoscopy. Exclusion criteria were active HIV, HBV or HCV infections. Biopsies were transported in phosphate buffered saline (PBS) on ice. Clinical data, tumor characteristics and molecular tumor data were pseudonymized. The study was approved by the Medical Ethics Committee II of the Medical Faculty Mannheim, Heidelberg University (Reference no. 2014-633N-MA and 2016-607N-MA). All patients gave written informed consent before tumor biopsy was performed. The initial cohort consisted of organoids from 25 patients with colorectal cancer, 10 of them female, 15 male, with a mean age of 66 years.

Table 2.1: Patient-derived organoid lines

Line	Optimal medium	Qualitative growth	Patient sex	Location	UICC stage
D004T	ENA	good	female	rectum	3
D007T	ENA	good	male	rectum	3
D010T	ENA	good	female	sigmoid	3
D013T	ENAS	good	male	rectum	1
D015T	ENA	good	male	descending	2
D018T	ENA	good	female	sigmoid	1
D019T	ENA	good	male	sigmoid	2
D020T	ENA	good	male	rectum	1
D021T	ENA	medium	male	rectum	1
D022T	ENA	good	male	ascending	4
D027T	ENA	good	female	rectum	2
D030T	ENA	good	female	ascending	2
D046T	ENA	good	female	rectum	3

2.3 Organoid Culture

2.3.1 Patient-Derived Organoid Culture

Patient-derived organoid (PDO) cultures were extracted from biopsies as reported by Sato et al. (Sato, Stange, et al. 2011) with slight modifications. Tissue fragments were washed in DPBS (Life technologies) and digested with Liberase TH (Roche) before embedding into BME R1 (Trevigen). The medium, termed ENA, contained Advanced DMEM/F12 (Life technologies) medium with 1% v/v penicillin/streptomycin (Life Technologies), Glutamax and HEPES (basal medium) supplemented with 100 ng/ml Noggin (Peprotech), B27 (Life technologies), 1.25 mM n-Acetyl Cysteine (Sigma), 10 mM Nicotinamide (Sigma), 50 ng/ml human EGF (Peprotech), 10 nM Gastrin (Peprotech), 500 nM A83-01 (Biocat), 10 nM Prostaglandin E2 (Santa Cruz Biotechnology), 10 µM Y-27632 (Selleck chemicals) and 100 mg/ml Primocin (Invivogen). After isolation, cells were kept in 2 conditions including medium as described (ENA), or supplemented with additional 3 µM SB202190 (Biomol) (ENAS) as described by Fujii et al. (Fujii, Shimokawa, et al. 2016). The tumor niche was determined after 14 days, and organoids were subsequently cultured in the condition with the best growth. Organoids were passaged every 7 days and the medium was changed every 2-3 days.

Table 2.2: Basal Medium Components

Component	Concentration	Manufacturer
Advanced DMEM/F12	97% v/v	Life technologies; Carlsbad, California, USA
GlutaMAX (100x)	1% v/v	Life technologies; Carlsbad, California, USA
Pen/Strep (100x)	1% v/v	Life technologies; Carlsbad, California, USA
1 M HEPES (100x)	1% v/v	Sigma-Aldrich Life Science/Merck, St. Louis, Missouri, USA

Table 2.3: ENA Medium Components

Supplement	Concentration	Manufacturer
Basal Medium (described above)	98% v/v	Produced In-House
B27 (50x)	2% v/v	Life technologies; Carlsbad, California, USA
N-Acetylcysteine	1.25 mM	Sigma/Merck; St. Louis, Missouri, USA
Human EGF	50 ng/ml	PeproTech; Hamburg, Germany
Noggin	100 ng/ml	PeproTech; Hamburg, Germany
Y-27632	10 µM	Selleck chemicals; Houston, Texas, USA
A83-01	500 nM	BioCat; Heidelberg, Germany
Prostaglandin E2	10 nM	Santa Cruz Biotechnology
Gastrin	10 nM	PeproTech; Hamburg, Germany
Nicotinamide (Vitamin B3)	10 mM	Sigma/Merck; St. Louis, Missouri, USA
Primocin	100 µg/ml	Invivogen; San Diego, California, USA

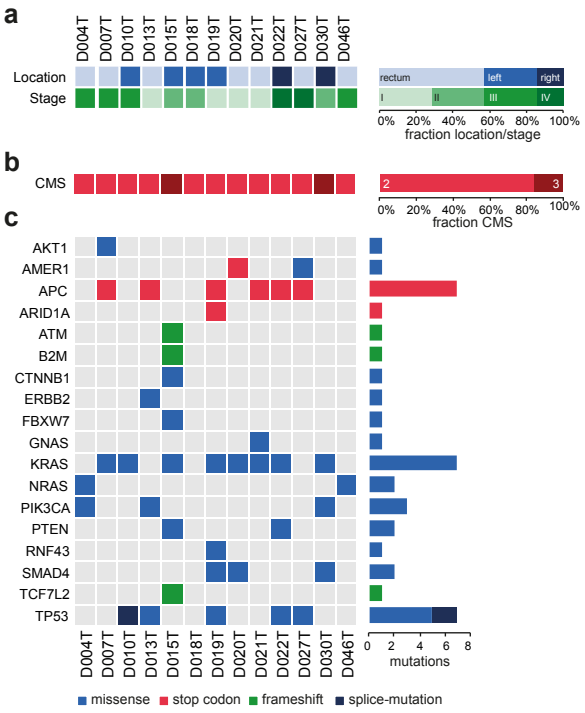

Figure 2.1: Patient-derived organoid cohort overview a Tumor location (right/left/rectum) and AJCC/UICC stage of colorectal cancers that Patient-derived organoids were derived from. **b** Consensus molecular subtypes of organoids determined by RNA expression analysis. **c** Mutation status in Patient-derived organoids, as analyzed by amplicon sequencing. Figure created by Johannes Betge (graphical presentation), Erica Valentini (sequencing data analysis) and Benedikt Rauscher (CMS type inference). Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022)

2.3.2 Mouse Organoid Culture

A heterozygous *LSL-Kras G12D (B6.129S4-Krastm4Tyj/J)* female mouse (Jackson et al. 2001) was crossed with a homozygous *Rosa26-CreERT2 (B6.129-Gt(ROSA)26Sortm1(cre/ERT2)Tyj/J)* male to generate offspring with a Tamoxifen activatable $Kras^{G12D/+}$ allele. A single healthy $LSL\text{-}Kras^{G12D}$ *CreERT2* mouse (male, 8 weeks) was sacrificed for organoid generation.

Mouse colon organoids were isolated based on work by Sato et al. (Sato, Vries, et al. 2009). After cervical dislocation of the sacrificed mouse, the colon was prepared and excised between cecum and rectum. The tissue

was stored on ice in cooled DPBS (Life technologies), cut open lengthwise and washed three times with DPBS. After thorough washing, colon fragments were cut into 2 mm pieces and incubated in a 5mM EDTA/DPBS (Sigma) solution for 60 minutes on a rocking table at 4C. Digested fragments were allowed to settle and resuspended in DMEM/F12 (Life technologies) by repeated up- and down-pipetting with a serological pipette. Here, care was taken to pre-wet the serological pipette to avoid loss of isolated cells during liquid handling. The resulting crypt suspension was filtered with a 70ul filter (Falcon), crypts were counted and centrifuged at 150g, 10 min, 4C. The resulting pellet was resuspended in 10 mg/ml Matrigel (Corning) and plated on prewarmed 6-well suspension plates (Greiner). After 30-60 minutes of solidification, droplets were overlaid with complete organoid growth medium and incubated at 37C, 5% CO_2 in atmospheric air.

Complete colon organoid medium, termed WENRAS, contained 30% advanced DMEM/F12 (Life Technologies) supplemented with 1% v/v penicillin/streptomycin solution (Life Technologies), 1% v/v HEPES buffer (Life Technologies) and 1% v/v Glutamax (Life Technologies), 50% Wnt3A conditioned medium, and 20% R-spondin1-FC conditioned medium. The medium was further supplemented with recombinant Noggin (100 ng/ml), 1x B27 (1x), n-Acetyl-cysteine (1.25 mM), Nicotinamide (10 mM), EGF (50 ng/ml), 500 nM A83-01 (Tocris), SB202190 (3 µM), Y-27632 (10 µM) and Primocin (100 µg/ml). All small molecule inhibitors were dissolved in DMSO.

Table 2.4: WENRAS Medium Components

Component	Concentration	Manufacturer
Wnt3A Conditioned Medium	50% v/v	Produced In-House
Advanced DMEM/F12	25% v/v	Life Technologies; Carlsbad, California, USA
R-spondin1-FC Conditioned Medium	20% v/v	Produced In-House
B27 Supplement (50x)	2% v/v	Life Technologies; Carlsbad, California, USA
Penicillin/Streptomycin Solution	1% v/v	Life Technologies; Carlsbad, California, USA
HEPES Buffer	1% v/v	Life Technologies; Carlsbad, California, USA
GlutaMAX	1% v/v	Life Technologies; Carlsbad, California, USA
Noggin	100 ng/ml	PeproTech; Hamburg, Germany
N-Acetyl-cysteine	1.25 mM	Sigma-Aldrich; St. Louis, Missouri, USA
Nicotinamide (Vitamin B3)	10 mM	Sigma-Aldrich; St. Louis, Missouri, USA
EGF	50 ng/ml	PeproTech; Hamburg, Germany
A83-01	500 nM	Tocris; Bristol, UK
SB202190	3 µM	Selleck Chemicals; Houston, Texas, USA
Y-27632	10 µM	Selleck Chemicals; Houston, Texas, USA
Primocin	100 µg/ml	Invivogen; San Diego, California, USA

Table 2.5: Cell Lines for Conditioned Media Production

Reagent or Resource	Source
R-Spondin1 expressing 293T Cell Line	Sigma/Merck; St. Louis, Missouri, USA
L Wnt-3a Cell Line	Clevers laboratory, Utrecht, Netherlands
HEK-293 Noggin-Fc Cell Line	Clevers laboratory, Utrecht, Netherlands

After isolation, colon organoids were cultured in solidified BME R1 (Trevigen) droplets for normal passaging and overlaid with genotype and experiment dependent growth medium. The medium was exchanged every 48-72 h.

Table 2.6: Hydrogels

Reagent or Resource	Manufacturer
Matrigel	Corning, Corning, NY 14831 USA
BME R1	Trevigen, Gaithersburg, Maryland, USA
BME V2	Trevigen, Gaithersburg, Maryland, USA

APC mutant colon organoid lines were cultured without Wnt and R-spondin conditioned medium, which was replaced by basal medium instead. Organoids were passaged weekly by digestion with TrypLE (Gibco) and resuspension in BME R1 (10 mg/ml). Organoids were regularly tested for Mycoplasma contamination.

2.3.3 Genetic editing of organoids

An sgRNA targeting the murine ortholog of the *APC* mutation cluster region (MCR) was designed using E-CRISP (Heigwer et al. 2014). The *Apc* targeting sgRNA was cloned into the one-vector plasmid pSpCas9(BB)-2A-Puro (PX459) V2.0 according to a protocol by Ran et al. (Ran et al. 2013). Briefly, the vector was digested with Bbs1-HF (Thermo Fischer Scientific) and the phosphorylated and annealed oligonucleotides for sgApc1 (sgApc1 F and -R) were ligated using T4-Ligase (Thermo Fischer Scientific). The construct was transformed into chemically competent bacterial cells (Stellar, Clontech) and plated on Carbenicillin agar. Individual colonies were isolated and sequencing of plasmid DNA from cultured colonies confirmed successful molecular cloning. Extracted organoids (abbreviated, "wild type" or "WT") were cultured for multiple passages before transfection of the plasmid with Lipofectamine 2000 (Thermo Fischer Scientific). For this step, grown organoids were digested with TrypLE (Gibco) and treated with Lipofectamine and plasmid DNA according to the manufacturer's protocol. Transfected organoids were seeded in BME R1 (10 mg/ml) and Wnt3A/R-Spondin1-Fc withdrawal was started 7 days after transfection. Surviving organoids were cultured continuously without Wnt3A and R-Spondin1-Fc conditioned medium. To activate oncogenic Kras, wild type and *Apc* mutant organoid lines were treated for 7 days with 0.5uM 4-Hydroxytamoxifen (Sigma) without EGF in the medium. 4-Hydroxytamoxifen was dissolved in Ethanol. After treatment, organoids were cultured with EGF containing media thereafter.

Table 2.7: List of Recombinant DNA

Reagent or Resource	Source	Identifier
pSpCas9(BB)-2A-Puro (PX459) V2.0	F. Zhang via Addgene	# 62988

Table 2.8: List of sgRNAs

Reagent or Resource	Sequence	Source
sgAPC1	GGCACTCAAAACGCTTTTGA	GATC Biotech

2.4 Image-based profiling

2.4.1 Patient-derived organoid seeding during compound testing

Patient-derived organoid drug profiling followed a standardized protocol with comprehensive documentation of all procedures. Organoids were collected and digested in TrypLE Express (Life technologies). Fragments were collected in basal medium with 300 U/ml DNAse and strained through a 40µm filter to achieve a homogeneous cell suspension with single cells and small clusters of cells, but without large organoid fragments. 384 well microclear assay plates (Greiner) were coated with 10µL BME V2 (Trevigen) at a concentration of 6.3 mg/ml in basal medium, centrifuged and incubated for >20 min at 300G and 37° C to allow solidification of the gel. Organoid cell clusters together with culture medium (ENA) and 0.8 mg/ml BME V2 were added in a volume of 50µl per well using a Multidrop dispenser (Thermo Fisher Scientific). Plates were sealed with a plate-loc (Agilent) and centrifuged for an additional 20 min, allowing cells to settle on the pre-dispensed gel. Cell number was normalized before seeding by measuring ATP levels in a 1:2 dilution series of digested organoids with CellTiter-Glo (Promega). The number of cells matching 10,000 units were seeded in each well. After seeding of organoid fragments, plates were incubated for three days at 37 °C to allow organoid formation before addition of compounds. Two biological replicates of each organoid line from different passages were profiled at different time points.

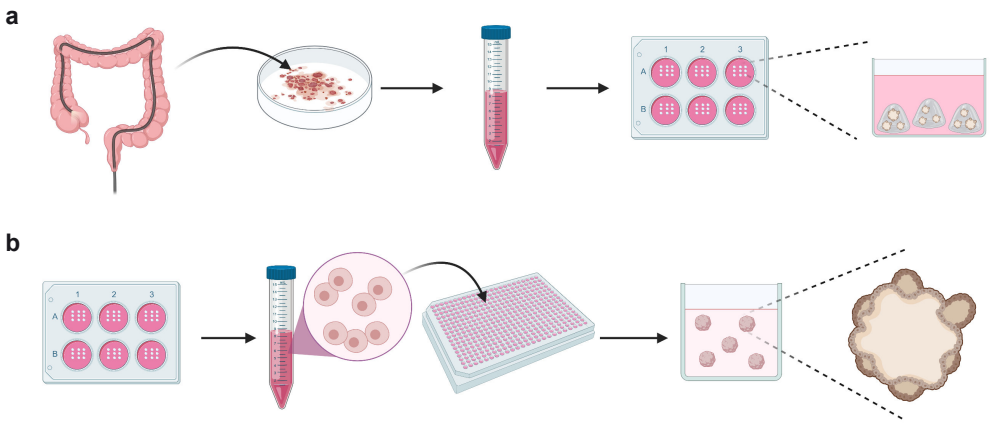

Figure 2.2: Core organoid liquid handling methods a Organoid isolation procedure. Colorectal cancer tissue biopsies were collected via endoscopy, enzymatically removed from extracellular matrix proteins, washed and resuspended in basal membrane extract hydrogel. After solidification of hydrogel domes, organoids were overlaid with growth factor rich culture medium. **b** Organoid high-throughput experimentation. Colorectal cancer organoids were harvested, partially digested, seeded in hydrogel-coated 384-well plates. Figure produced with Biorender

2.4.2 Mouse Organoid seeding during compound testing

Mouse organoid screening was performed as described above with slight modifications. Clotting of organoid fragments was avoided by adding 10 U/ml of bovine DNAse1 to the medium during filtration. The cell viability of

digested fragment suspensions was estimated using Cell-Titer-Glo (Promega). 40µl of cell suspension was mixed with 40µl of undiluted reagent and measured after 30 minutes on a Mithras plate reader (Berthold). Cell fragments with a viability corresponding to 5000 units were seeded per well on pre-coated 384 well microclear assay plates (Greiner) using a Multidrop peristaltic pump robot (Thermo Fischer Scientific) analogous to the human organoid seeding protocol outlined above.

2.4.3 Compound Libraries and Treatment for Patient-derived Organoids

Two compound libraries were used for profiling: A library containing 63 clinically relevant drugs (clinical cancer library) and a large library of 464 compounds targeting kinases and stem cell or developmental pathways associated genes (KiStem library). The clinical cancer library was manually curated by relevance for current (colorectal) cancer therapy, mechanism of action and potential clinical applicability. Compounds of this library are in clinical use or at least in phase I/II clinical trials. Five concentrations per compound were screened (five-fold dilutions). The concentrations were determined by analysis of literature data from previous 3D and 2D drug screens and own experiments. The KiStem library includes 464 compounds targeting a diverse set of kinases and stem cell relevant pathways. All compounds in this library were used at a concentration of 7.5 µM. All compounds were obtained from Selleck chemicals. Compounds of both libraries were arranged in an optimized random layout. Compound libraries were stored in DMSO at -80 C.

After 72 hours of expansion at 37C, medium was aspirated from all organoid screening plates and replaced with fresh ENA medium devoid of Y-27632, resulting in 45µl volume per well. Drug libraries were diluted in basal medium and subsequently 5µl of each compound was distributed to screening plates. All liquid handling steps were performed using a Biomek FX robotic system (Beckman Coulter). Plates were sealed and incubated with the compounds for four days. All PDO lines underwent profiling with the clinical cancer library, while the KiStem library was used with 13 PDO lines.

2.4.4 Compound Libraries and Treatment for Mouse Organoids

Two compound libraries were used for profiling: The KiStem library (see above) as well as a library of FDA-approved small molecules. For both the KiStem library and the FDA-approved library, one single concentration was used. All compounds in this library were used in a concentration of 7.5µM. All compounds were obtained from Selleck chemicals. Compounds of both libraries were arranged in an optimized random layout and complete compound libraries were stored in DMSO at -80 C.

Mouse Organoids were treated similar to Patient-derived Organoids with slight modifications: After 72 hours of organoid expansion in WENRAS, the medium was changed to ENR and compound libraries were added using a BiomekFX (Beckman Coulter). All Mouse Organoids were profiled with the KiStem library and FDA-approved compound library.

2.4.5 Automated Microscopy

After 96h of small molecule treatment, Image-IT DeadGreen (Thermo Fisher) was added to the cultures with a Multidrop dispenser (Thermo Fisher) in 100 nM final concentration and incubated for four hours. Afterward, the medium was removed and organoid cultures were fixed with 3% PFA in PBS with 1% BSA. Fixed plates were stored at 4° C for up to three days before permeabilization and staining. On the day of imaging, organoids were permeabilized with 0.3% Triton-X-100 and 0.05% Tween in PBS with 1% BSA and stained with 0.1µg/ml TRITC-Phalloidin (Sigma) and 2µg/ml DAPI (Sigma). All liquid handling steps were performed with a BiomekFX. Screening plates were imaged with an Incell Analyzer 6000 (GE Healthcare) line-scanning confocal fluorescent microscope. We acquired 4 fields per well with z-stacks of 16 slices at 10x magnification. The z-steps between the 16 slices had a distance of 5µm, the depth of field of each slice was 3.9µm.

2.4.6 Luminescence Viability Read Out of Patient-derived Organoids

Organoid screening plates undergoing ATP-based viability testing were cultured in solid, white plates (Greiner) treated with 30µl CellTiter-Glo reagent after medium aspiration with a Biomek FX. After incubation for 30 minutes, luminescence levels were measured with a Mithras reader (Berthold technologies).
Raw luminescence data of each plate was first normalized using the Loess-fit method to correct for edge effect if increased luminescence intensity was observed along the edges of each plate. Subsequently, each plate was normalized by division with the median luminescence intensity of the DMSO controls. Drug response Hill curves (DRC) were fitted and area under the curve values were calculated for each DRC using the 'PharmacoGx' (Smirnov et al. 2016) R/Bioconductor package.

2.4.7 Luminescence Viability Read Out of Patient-derived Organoids of Mouse Organoids

For selected compounds, before compound addition, organoid viability was measured using Cell-Titer-Glo (Promega). Cell viability was measured after compound exposure as described above. The pre-treatment viability of organoids was used to estimate growth-rate controlled dose-response curves according to (Hafner et al. 2016). Measurements of GR metrics were not robust for slow proliferating wild type lines. Therefore, these lines were omitted in the analysis.

2.5 Image analysis

2.5.1 Image Processing

Microscopic image z-stacks were compressed to HDF5 format for archival and underwent maximum contrast projection using the R/Bioconductor package MaxContrastProjection developed by Jan Sauer for further processing of the images. Standard image features, including shape, moment, intensity, and Haralick texture features on multiple scales, were extracted using the R/Bioconductor package EBImage (Pau et al. 2010). Of

note, the strong diversity of unperturbed organoid phenotypes between organoid lines did not allow the definition of a core set of individual reproducible descriptive features across all screened organoids. Therefore, no correlation-based filtering of features was done, allowing comparisons between different lines. Out-of-focus objects were manually and programmatically removed from the dataset using a custom classifier developed by Jan Sauer.

2.5.2 Feature Processing and Treatment-Induced Phenotypes

The dimensionality of calculated single-organoid morphological features was reduced using a principal component analysis (PCA) that was performed on the entire datasets (Patient-derived and mouse organoids separately) to reduce the dimensionality. 25 principal components were selected, explaining ca. 80% of the total variance within both datasets. These 25 principal components were used for further analysis of single-organoid morphology and were averaged on a per-well level for the MOFA analysis.

For treatment activity estimation, a logistic regression classifier was trained per line and treatment (and per concentration where applicable) to differentiate individual treated organoids from negative controls based on the PCA-transformed features. Drugs were categorized as either active or inactive based on the accuracy of the model. The direction of the decision plane's normal vector was used as the treatment effect vector. Drugs were clustered based on the cosine similarity of their normal vectors.

2.6 Multi-omics factor analysis

2.6.1 Model training of Patient-derived Organoids

A multi-omics factor analysis model (Argelaguet, Velten, et al. 2018) was trained based on a set of five modalities describing unperturbed organoid lines:

- organoid size estimated based on log-normal model fit of all DMSO treated organoids [22 replicates, 1 feature]
- organoid somatic mutations as determined by amplicon sequencing [20 replicates, 12 features]
- organoid transcript expression of the top 10% genes with the highest coefficient of variance after robust multi-array average normalization [22 replicates, 3222 features]
- organoid morphology as determined by averaging DMSO treated morphological profiles [22 replicates, 25 features]
- organoid drug activity as determined by AUROC score of logistic regression models for drugs that were defined as active in at least one observation [22 replicates, 252 features]

Input data was scaled and the MOFA model was trained with default MOFA2 training parameters and a number of 3 factors. The number of factors was chosen given the limited number of observations in the training data. The further analysis focused on the first two factors, which correlated with prominent visible organoid phenotypes. Gene set enrichment analysis and Reactome pathway enrichment of factor loadings was performed using the clusterProfiler R package (Yu et al. 2012). Enrichment of drug targets within factor loadings was tested using ANOVA by fitting a linear model, *lm(factor loading vs. target)*. Drug targets that were represented with at least three small molecule inhibitors were included in this analysis. The analysis was run using the MOFA docker container, available from https://hub.docker.com/r/gtca/mofa2.

2.6.2 Model training of Mouse Organoids

A multi-omics factor analysis model with k=4 factors was trained and results were analyzed as above, with a different set of input data:

- organoid size of all DMSO treated organoids (one replicate of Apc-/- organoids was removed from the analysis) [7 replicates, 1 feature]
- organoid transcript expression including the top 10% genes with the highest coefficient of variance after robust multi-array average normalization [8 replicates, 2727 features]
- organoid protein expression [12 replicates, 3906 features]
- organoid lipid abundance [12 replicates, 397 features]
- organoid genotype for the *Apc* and *Kras* gene [12 replicates, 2 features]
- organoid morphology as determined by averaging DMSO treated morphological profiles (one replicate of $Apc^{-/-}$ organoids was removed from the analysis) [7 replicates, 25 features]
- organoid drug activity as determined by AUROC score of logistic regression models for drugs [7 replicates, 1699 features]

2.6.3 Model projection

To estimate the factor scores for treatment-induced organoid morphologies, the morphology feature matrix was multiplied with the pseudoinverse of the previously learned model's factor weight matrix for the organoid morphology modality.

The resulting projected factor score matrix was used to estimate the drug-induced biological changes in both Patient-derived and mouse organoids. Associations between drug targets and projected factor scores of drug treated organoids were identified via ANOVA by fitting a linear model, *lm(projected factor score vs. target)*. For mouse organoids, the model was extended to account for line-wise effects in the modeling, *lm(projected factor*

score vs. target + organoid-line). Drug targets that were represented with at least three small molecule inhibitors were included in the analysis, except in the modeling of individual small molecule effects on mouse organoids.

2.7 Biochemical assays

2.7.1 Amplicon Sequencing of Mouse Organoids

Amplicon sequencing was performed to validate the CRISPR perturbation of *Apc*. DNA from *Apc* targeted and untargeted organoid lines was prepared using the DNA Blood and Tissue Kit (Qiagen), according to the manufacturer's tissue protocol, including an RNAse digestion. The targeted region was PCR amplified using primers F1 and R1. Libraries were sequenced on a MySeq (Illumina) using 100bp single end reads.

Table 2.9: List of Genomic PCR Primers for Amplicon Sequencing

Reagent or Resource	Sequence	Source
Primer F1	TCCCTACACGACGCTCTTCCGATCTGGAATGTCAGAAGGGAGACC	GATC Biotech
Primer R1	AGTTCAGACGTGTGCTCTTCCGATCTCCAACCAGAAATGCCAGTG	GATC Biotech

2.7.2 Genomic PCR of the KRAS G12D allele in Mouse Organoids

To confirm activation of oncogenic *Kras* in 4-Hydroxytamoxifen treated lines, genomic DNA was isolated from all 4 organoid lines as described above. Presence or absence of the Lox-STOP-Lox cassette was evaluated by PCR according to the $Kras^{G12D/+}$ conditional PCR protocol by Tyler Jacks' group (Jackson et al. 2001). Briefly, primers #2 and #3 were used for genotyping on genomic DNA using the Q5 PCR protocol (NEB).

Table 2.10: List of Primers for validating KRAS conditional PCR

Reagent or Resource	Sequence	Source
Primer #2	CTCTTGCCTACGCCACCAGCT	GATC Biotech
Primer #3	AGCTAGCCACCATGGCTTGAGTAAGTCTGCA	GATC Biotech

2.7.3 Western Blot of Patient-derived Organoids

Organoids seeded in 6-well plates were cultured in Matrigel (Corning). After 3-days incubation with WYE-132, organoids were collected, and cells were isolated using Matrisperse (Corning) for 40 minutes on a rocking table. Cells were subsequently lysed in RIPA buffer (Thermo Fisher Scientific) supplemented with Protease inhibitors (Complete Mini, Roche) and Phosphatase inhibitors 1 and 2 (Sigma), followed by sonication (Branson Sonifier, Heinemann). Protein concentrations of supernatants were measured using the Pierce BCA kit (Thermo Fisher Scientific) according to the manufacturers' protocol. Lysates were mixed with an SDS-loading buffer and heated to 99 C for 5 minutes. Proteins were separated by SDS–PAGE in MOPS running buffer and transferred to a

nitrocellulose membrane. Membranes were blocked with 5% (w/v) skim milk in PBS containing 0.1% (v/v) Triton X-100 (PBS-T). Western blotting was performed with the following antibodies (all in 5% (w/v) skim milk, PBS-T): anti-IRS1 (1:1000, 06-248, Sigma–Aldrich), anti-HSP-90 (1:1000, sc-13119, Santa Cruz), anti-Mouse-IgG-HRP (1:10000, Sigma–Aldrich). ECL Western Blotting W1001 (Promega) was used for visualization of bands.

Table 2.11: List of Antibodies used in experiments

Reagent or Resource	Source	Identifier
Anti-IRS1 (rabbit)	Sigma–Aldrich	06-248
Anti-Erk (p44/42 MAPK) (rabbit)	Cell Signaling	4695
Anti-phospho-Erk (phospho-p44/42 MAPK) (rabbit)	Cell Signaling	4370
Anti HSP-90 (mouse)	Santa Cruz	13119
Anti-Mouse-IgG-HRP (goat)	Sigma–Aldrich	12-349
Anti-beta-Actin-HRP (mouse)	Cell Signaling	12262

2.7.4 Western Blot of Mouse Organoids

Organoids were cultured in Matrigel (Corning). Organoids were collected, and cells were isolated using Matrisperse (Corning) for 40 minutes on a rocking table. Isolated organoids were lysed in RIPA buffer (Sigma) with Protease inhibitor (Sigma) and Phosphatase inhibitor 3 (Sigma). Protein concentration was measured using the Pierce BCA kit (Thermo Fischer Scientific) according to the manufacturers' protocol. Lysates were mixed with an SDS-loading buffer and heated to 99 C for 5 minutes. Samples were loaded onto NuPage gels (Thermo Fischer Scientific), separated in MOPS running buffer and transferred to a nitrocellulose membrane. Membranes were blocked with 5% (w/v) skim milk in PBS containing 0.1% (v/v) Triton X-100 (PBS-T). Western blotting was performed with the following antibodies (all in 5% (w/v) skim milk, PBS-T): anti-p(hospho)-Erk (1:2000, Cell Signaling, ID 4370), anti-Erk (1:1000, Cell Signaling, ID 4695) and anti-beta-Actin-HRP secondary antibody (1:150,000).

2.7.5 Mouse Organoid Growth Patterns

Organoids were passaged and seeded in 4 different growth media, with medium changes every 48h. Images were taken 120h after seeding with 4x magnification on a Zeis Axiovert (Zeis).

2.7.6 RT-qPCR of Patient-derived Organoids

Organoids seeding and treatment timing was performed in accordance with the image-based profiling protocol. After 120h of growth, total RNA was isolated using the RNAEasy Mini kit (Qiagen) without additives. cDNA synthesis was done with the Verso cDNA kit (Thermo Fisher Scientific), and RT-qPCR was performed using the SYBR Green Mix (Roche) on a LightCycler480 system (Roche). UBC expression levels were used as controls.

Table 2.12: List of RT-qPCR Primers

Reagent or Resource	Organism	Sequence	Source
LGR5 F	Human	TTCCCAGGGAGTGGATTCTAT	GATC Biotech
LGR5 F	Human	ACCAGACTATGCCTTTGGAAAC	GATC Biotech
UBC F	Human	CTGATCAGCAGAGGTTGATCTTT	GATC Biotech
UBC R	Human	TCTGGATGTTGTAGTCAGACAGG	GATC Biotech
Axin2 F	Mouse	GAGAGTGAGCGGCAGAGC	GATC Biotech
Axin2 R	Mouse	CGGCTGACTCGTTCTCCT	GATC Biotech
Ccnd F	Mouse	TTTCTTTCCAGAGTCATCAAGTGT	GATC Biotech
Ccnd R	Mouse	TGACTCCAGAAGGGCTTCAA	GATC Biotech
Sdha F	Mouse	TGTTCAGTTCCACCCCACA	GATC Biotech
Sdha R	Mouse	TCTCCACGACACCCTTCTGT	GATC Biotech
Hprt F	Mouse	CCTCCTCAGACCGCTTTTT	GATC Biotech
Hprt R	Mouse	CCTCCTCAGACCGCTTTTT	GATC Biotech

2.7.7 RT-qPCR of Mouse Organoids

Organoids were passaged and seeded in 4 different growth media, with medium changes every 48h. After 120h, organoid RNA was isolated using the RNAEasy Kit (Qiagen) with beta-Mercaptoethanol (Invitrogen) and a DNAse digestion step. cDNA was synthesized using Oligo-dT primers (Thermo Fisher Scientific), RiboLock Ribonuclease inhibitor (Thermo Fisher Scientific) and Revert Aid H Minus reverse transcriptase (Thermo Fisher Scientific). RT-qPCR was performed using the ROCHE UPL kit (Roche) on a LightCycler480 system (Roche). Sdha and Hprt expression levels were used as controls and averaged. Relative transcript abundance was measured using the ddCT method.

2.7.8 Proteomics Profiling of Mouse Organoids

Organoids were cultured according to the image-based profiling protocol. Samples were isolated with Matrisperse (Corning) as described above. Isolated organoids were lysed in Ammonium Bicarbonate lysis buffer (50mM, pH 8.2) with 2.5% w/v SDC and 25U/ml Benzonase.

Samples were handed over to the Proteomics core facility at the German Cancer Research Center. There, a ”1D-SDS-PAGE of the lysate was run followed by fractionation. Gel pieces were extracted, cysteins residues reduced by DTT and carbamidomethylated using iodoacetamide. The samples were digested with Trypsin overnight. Resulting peptides were loaded on a cartridge trap column, packed with Acclaim PepMap300 C18, 5 µm, 300Å wide pore (Thermo Fischer Scientific) and segregated in a 60-min gradient from 3% to 40% ACN on a nanoEase MZ Peptide analytical column (300Å, 1.7 µm, 75 µm x 200 mm, Waters). Eluted peptides were analyzed by an online coupled Q-Exactive-HF-X mass spectrometer” (*quoted description taken from direct correspondence with the core facility*).

Protein abundance data were processed using the DEP package (X. Zhang et al. 2018) accessed through the R/Bioconductor ecosystem (Huber, Carey, et al. 2015). After removing missing LFQ values, ca. 3000 protein

fragments were observed across all 12 conditions. Data were normalized using VSN variance stabilization (Huber, Heydebreck, et al. 2002) and missing values imputed. Batch effects were removed using Combat (Leek et al. 2007). The Pearson correlation between Protein and Transcript abundance ranged between 0.3-0.36 across four organoid lines.

2.7.9 Lipidomics Profiling of Mouse Organoids

Organoids were cultured according to the image-based profiling protocol and prepared for analysis according to the method developed for proteomics profiling (above). Isolated organoids were processed by the Lipidomics and Metabolomics core facility. "After lipid extraction, isolates were analyzed using a Qtrap 6500 mass spectrometer (SCIEX) coupled to a NanoMate (NanoMate) electrospray source" (*quoted description taken from direct correspondence with the core facility*).

Lipid abundance data was processed using the DEP package (X. Zhang et al. 2018) accessed through the R/Bioconductor ecosystem (Huber, Carey, et al. 2015). After removing missing values, ca. 350 lipid species were observed across all 12 conditions. Data were normalized using VSN variance stabilization (Huber, Heydebreck, et al. 2002) and missing values imputed. Batch effects were removed using "Combat" (Leek et al. 2007).

2.7.10 Transcript Expression Profiling of Patient-derived Organoids

Organoid RNA was isolated from 19 Patient-derived organoid lines with the RNeasy mini kit after snap freezing organoids on dry ice. Samples were hybridized on Affymetrix U133 plus 2.0 arrays.

Raw microarray data were normalized using the robust multi-array average (RMA) method (Irizarry et al. 2003) followed by quantile normalization as implemented in the "affy" (Gautier et al. 2004) R/Bioconductor (Huber, Carey, et al. 2015) package. Differential gene expression analyses were performed using a moderated t-test as implemented in the R/Bioconductor package "limma" (Ritchie et al. 2015). Gene set enrichment analyses were performed using ConsensusPathDB (Kamburov et al. 2011) for discrete gene sets or GSEA as implemented in the "fgsea" R/Bioconductor package for ranked gene lists.

2.7.11 Transcript Expression Profiling of Mouse Organoids

Mouse organoids were cultured according to the image-based profiling protocol. Briefly, organoid models were seeded and cultured for 72h in WENRAS and an additional 96h in ENR. Samples were harvested after 7 days and RNA was isolated using the RNAEasy Kit (Qiagen) as described above. Transcript expression levels were measured using MoGene-2_0-st chips (Affymetrix).

Raw microarray data processing was generally performed analogous to the analysis of Patient-derived Organoids. With a series of modifications, including the use of "Combat" (Leek et al. 2007) to remove batch after robust

multi-array averaging (RMA) (Irizarry et al. 2003), removing the lowest 5% of observed probes, and mapping of mouse genes to human orthologs. Gene set enrichment analyses were performed using methods described above, as well as using the ”clusterProfiler” package (Yu et al. 2012).

3

Profiling of colorectal cancer organoids identifies multi-view factors of cancer organoid architecture and plasticity

3.1 Disclosure

This chapter has been adapted from a joint first-author publication, *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). The adapted text was originally written by myself. The experiments presented in this chapter were conducted together with Johannes Betge. Jan Sauer performed the initial image feature extraction and treatment activity scoring (AUROC). Figures contributed by other co-authors are marked as such. Additional details are available in the *Eigenanteil* chapter of this thesis.

3.2 Establishing a Patient-derived organoid cohort for image-based Profiling

Patient-derived organoids can be established from a diverse set of healthy or malignant tissues and have been shown to represent their tissue of origin regarding morphological and molecular features, including gene

expression and somatic mutations (Fujii, Shimokawa, et al. 2016; Wetering et al. 2015; Sato, Stange, et al. 2011). To generate personalized cancer models for image-based profiling, a standardized laboratory workflow to generate patient-derived organoids from colorectal cancer samples via endoscopic biopsy was designed and implemented.

3.2.1 Establishing image-based profiling of organoids

To systematically study organoid morphology under different treatment conditions, I established a platform for high-throughput image-based profiling experiments. The three engineering problems that had to be solved were (1) control of organoid size and density in a 384-well plate format, (2) control of organoid location within the hydrogel for efficient microscopy and (3) maintenance of organoid integrity during automated fixation and permeabilization.

A standard protocol for cell-based assays, including image-based profiling, is seeding cells into microwell plates at a fixed cell number. To determine cell number, adherent cells are dissociated and counted using optical methods. Patient-derived organoids, however, demonstrated a low rate of organoid outgrowth when passaged by complete organoid dissociation down to the single cell level. To improve organoid outgrowth, the enzymatic dissociation process was stopped early, yielding cell clusters of ca. 1-10 cells. These organoid fragments showed an increased outgrowth rate, which could be further improved by treating cells with 10 µM of Rho-Kinase inhibitor Y-27632 (data not shown). To control organoid size and density, organoids were digested with a modified trypsin derivative, and filtered through a 40µm cell strainer to ensure an upper limit of organoid fragment size. To effectively estimate the cell number while maintaining organoid fragments, organoid fragments were titrated based on their ATP concentration, instead of cell count. The ATP concentration of the organoid fragment suspension was determined using an ATP-dependent luminescence readout, Cell-Titer-Glo (CTG). After controlling for ATP concentration, the organoid fragment suspension was seeded onto basal membrane extract covered 384 well plates.

Conventional image-based profiling of adherent cells is based on automatic microscopy of one 2D plane per field of view. Given the 3D growth patterns of organoids, more data has to be acquired to fully capture organoid phenotype. Acquiring multiple planes of imaging data per field of view, however, creates a technical data storage and processing burden. For a fixed 3D volume, the size of the collected data increases linearly with the number of acquired planes and quadratically with the target z-axis resolution. To reduce the observed 3D volume and thus the number of required imaging planes, the vertical distribution of organoid fragments within the basal membrane extract layer was controlled by centrifuging organoid fragments post-seeding at 500G for 20 minutes at 37 degrees Celsius. The centrifugal force led to an accelerated sedimentation of organoid fragments onto the same optical plane before the polymerization of the hydrogel was complete.

After three days of culture and four days of compound treatment, organoids were fixed and stained for actin (Phalloidin/TRITC), DNA (DAPI), and cell permeability (Image-IT DeadGreen/FITC). Subsequently, plates were

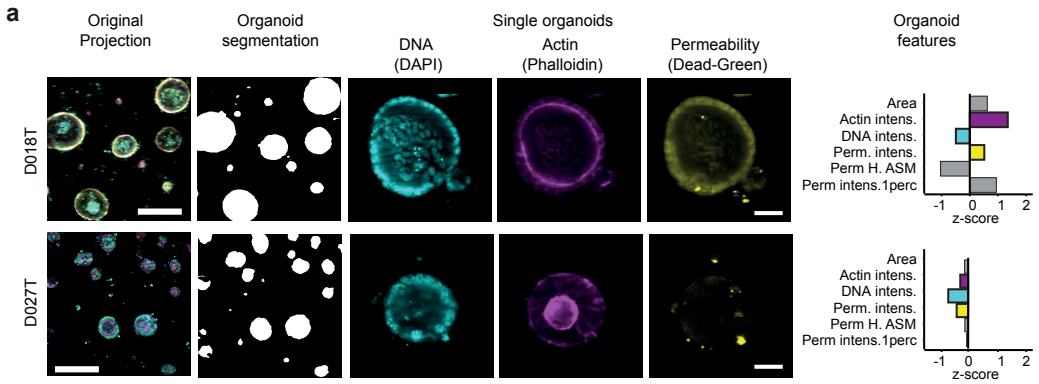

Figure 3.1: Image-based profiling method overview. a The image-processing pipeline illustrated with representative example images from two organoid lines: Organoids were imaged at multiple layers along the z-axis. Images were projected using a maximum contrast projection and segmented using a convolutional neural network, both designed and implemented by Jan Sauer. Descriptive features were extracted from all three channels to quantify phenotypes. Feature plots show the median phenotype of unperturbed organoids, six example features (Area, Phalloidin intensity, DAPI intensity, FITC intensity, FITC Haralick angular second moment (ASM) and FITC intensity 1-percentile) and their z-scores relative to all profiled organoid lines are shown. Figure created with support from Jan Sauer (data processing) and Johannes Betge (graphical presentation). Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022).

imaged at multiple z-positions by automated confocal microscopy. During treatment with the hyperosmolar fixative (3% para-formaldehyde in phosphate buffered saline) the protein rich hydrogel underwent an irreversible volume contraction (data not shown). To reduce this physical artifact, the fixative was supplemented with bovine serum albumin to a final concentration of 1% weight/volume. After image acquisition, the three-dimensional morphological information was projected onto a two-dimensional plane to further compress the size of the acquired data (Figure 3.1 a), organoids were detected, and classic phenotype features were calculated (implemented by Jan Sauer). These organoid-level morphological descriptors were used for further analysis.
To summarize, (1) seeding organoid fragments with controlled number and size-distribution instead of single cells, (2) centrifuging organoid fragments to reduce the imaged 3D volume, and (3) modifying liquid handling buffers to avoid hydrogel-driven artifacts, technically enabled high-throughput image-based profiling of organoid models.

To better understand the diversity of patient-derived organoid phenotypes and how their morphology was linked to biological state, image-based profiling at single organoid resolution was performed with 11 patient-derived cancer organoids using compounds targeting developmental pathways (464 compounds), as well as compounds in clinical use (63 compounds in 5 concentrations, Figure 3.2 a and b).

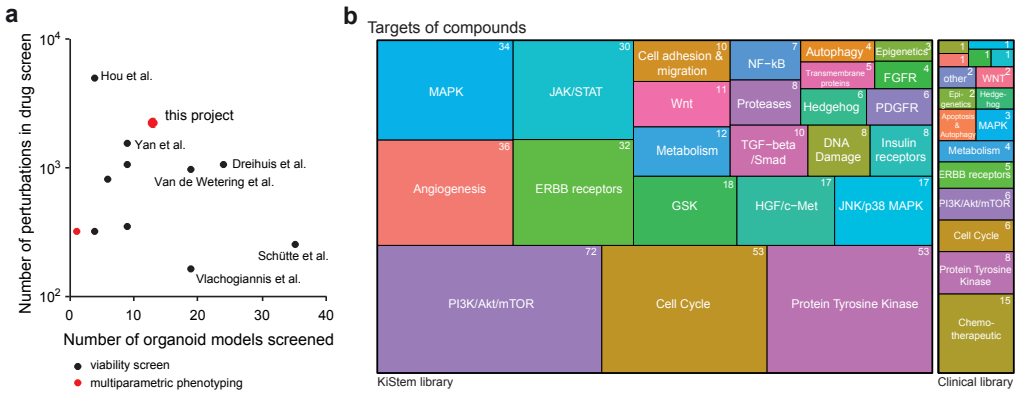

Figure 3.2: Dataset dimensions and compound library overview a Number of organoid models and number of perturbations in previous publications reporting high-throughput drug screenings with patient-derived cancer organoids, **b** Graphical representation of the compound libraries used for drug screening in this project: A library targeting kinases and stem cell pathways (KiStem library, 464 compounds) and a clinical library with 63 drugs in 5 concentrations. Figure created with support from Johannes Betge and adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022)

Additional transcript expression profiles of untreated organoids, and amplicon-based somatic mutation profiles were collected as supporting data. The resulting data comprised morphological profiles for each organoid with 528 phenotypic features that were subsequently processed using PCA. Based on an elbow method, 25 principal components, accounting for 81% of morphological variance, were retained as used as morphological features throughout this project.

To visualize the heterogeneity of colorectal cancer organoids and drug induced changes across and within cancer organoid lines, the 25 principal components of ca. 5.5 million profiled organoids were embedded using uniform manifold approximation and projection (UMAP) (Figure 4.7 a and 3.4 a-c). Most organoid lines showed characteristic bimodal log-normal distributions of organoid size with one component containing small organoids and another component made up of larger organoids with varying, line specific, average size (Figure 4.7 b, and 3.4 d-e). The log-normal-like size distribution likely resulted from intrinsic differences in cellular size and growth rate, compounding over time in multicellular organoids.

While DNA and Actin staining intensity were positively correlated with organoid size, cell permeability was negatively correlated and enriched in regions with relatively smaller organoids (Figure 3.4 a-c). Graph-based

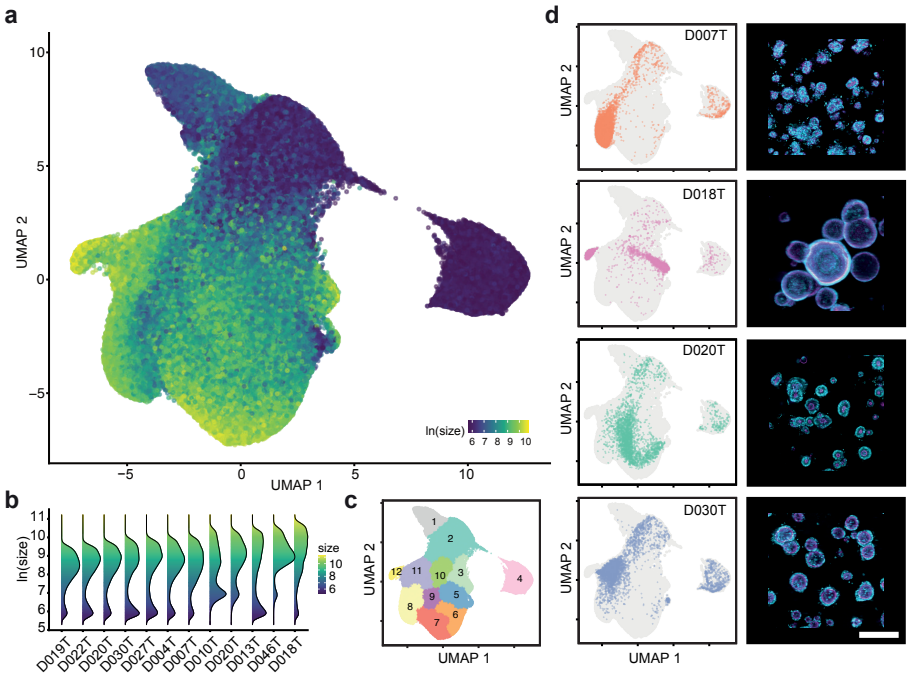

Figure 3.3: Image-based profiling captures the phenotype diversity of patient-derived cancer organoids a Uniform Manifold Approximation and Projection (UMAP) of 25 morphology principal components at single-organoid-resolution for a random 5% sample out of ca. 5.5 million organoids. The same sample is used for visualizations throughout the figure. Color corresponds to the log-scaled organoid area (dark blue: minimum size, yellow: maximum size). **b** organoid size distribution across lines. **c** UMAP representation of DMSO treated and drug treated organoids. Graph-based clustering of organoids by morphology. **d** UMAP embeddings of selected organoid lines (baseline state / 0.1% DMSO control-treated organoids) representing different morphological subsets, gray background consists of randomly sampled points. Depicted are representative example images for each line (right, cyan = DNA, magenta = Actin, scale-bar: 200 µm). Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). Johannes Betge retrieved the example images for part d.

clustering of this identified 12 regions within the embedding (Figure 4.7 c). When comparing drug-treated organoids to organoids treated with the negative control (DMSO), no clear separation of these two groups, except an increased presence of drug-treated organoids in region 3, was seen. This finding suggested that organoid morphology was distributed on a continuum of phenotypes spanning perturbed and unperturbed conditions of the experiment (Figure 3.4 f).

Different organoid lines within the embedding were located in characteristic regions, with organoid size and organoid architecture as primary organizing factors (Figure 4.7 b and d). For example, organoid line D018T had the largest median organoid size within the dataset and a cystic organoid architecture, while D020T organoids had a solid architecture and smaller median size. In most cases, organoid lines had two areas of main density, with one of them in regions 2, 3 or 4, reflecting the previously mentioned bimodal size distribution. In summary, image-based profiling of patient-derived colorectal cancer organoids showed strong morphological heterogeneity with line-dependent differences in size and organoid architecture.

Exploratory data analysis of the relationship between organoid morphology and experimental batch showed

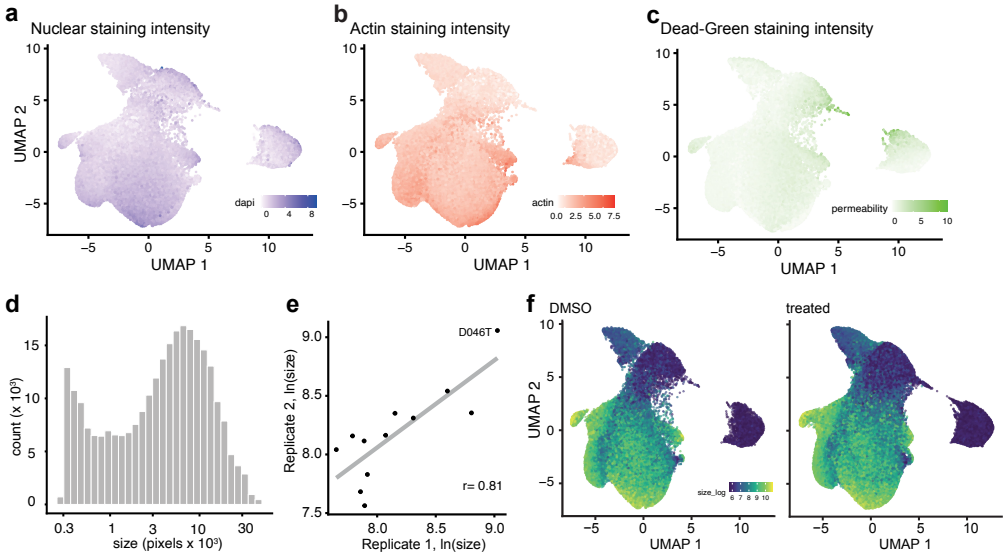

Figure 3.4: Basic image-based features and their role in organoid phenotype diversity. a-c Uniform Manifold Approximation and Projection (UMAP) of organoid-level features marked by DNA (DAPI) staining intensity (b), actin (Phalloidin/FITC) staining intensity (c) and permeability (Image-IT DeadGreen) staining intensity **d** Distribution of organoid size for all control (DMSO) treated organoids. **e** Replicate correlation of organoid size for control treated organoids. **f** UMAP representation of DMSO treated and drug treated organoids. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022)

overall reproducible measurements of organoid size across experiments (Figure 3.4 e). Similarly, further analysis of the UMAP showed overlapping morphology distributions between replicates for a given organoid line (Figure 3.5 a). Given that line differences were confounded by experimental batches (experimental batches and tested organoid lines were not independent), no procedure to reduce possible batch-dependent differences in organoid phenotype was performed.

3.3.1 Organoid phenotype-profiles capture organoid viability

Cell viability assays are common readouts in cancer drug discovery. Prompted by the observation that organoid size was a major factor determining the structure of the phenotype embedding (UMAP and factor 1 in MOFA analysis, see below), I hypothesized that low organoid size was at least partially the result of cell death within the organoid and, more broadly, that phenotype data could be used to estimate organoid viability. Bortezomib, a small molecule proteasome inhibitor with high *in-vitro* toxicity, led to dose dependent organoid death in all organoid lines, thus representing suitable positive controls (Figure 3.6 a). Analogous to pseudotime in single-cell gene expression analysis, dose-dependent trajectories of Bortezomib drug response could be fitted (Figure 3.6 b) using the non-parametric principal curve method. Starting from diverse baseline morphologies, increasing doses of Bortezomib led to a step-wise convergence on a final death-related phenotype, which corresponded to the areas with enrichment of small objects (regions 2, 3 and 4).

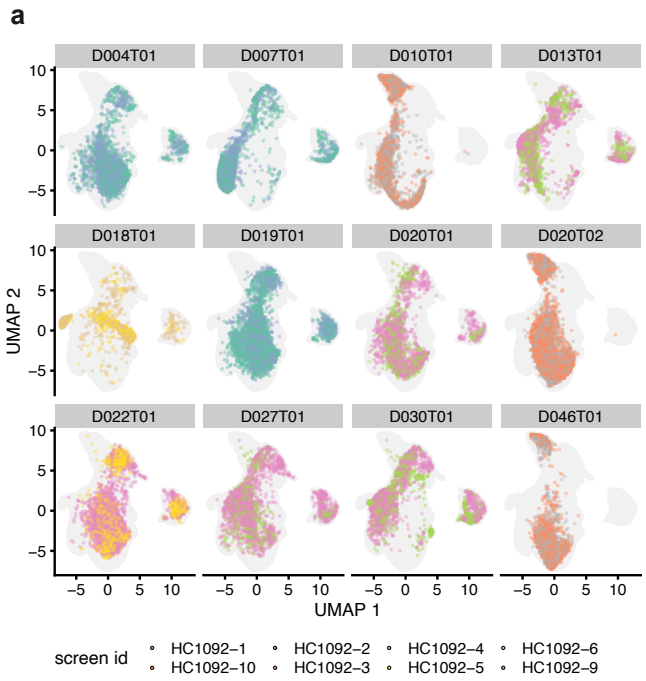

Figure 3.5: Experimental batches and their impact on organoid phenotype. a UMAP of organoid level features stratified by organoid line and colored by experimental batch.

Similarly, paclitaxel, a microtubule disassembly inhibitor, shifted the bimodal size distribution of organoids in a dose-dependent fashion (Figure 3.7 a), while organoid count remained largely unchanged (Figure 3.7 b). This effect, however, was organoid line-specific, as median organoid size in paclitaxel sensitive lines (e.g., D022T) decreased, while the size of other organoids remained unaffected (e.g., D046T, Figure 3.7 c-f). These observations suggested a link between organoid morphology, especially organoid size, with a loss of cell viability.

To further test the link between organoid morphology and cell viability, a luminescence-based, ATP dependent, cell viability assays (CTG) was performed in parallel with imaging as benchmark. A correlation of CTG viability with organoid size (r = 0.64) (Figure 3.8 a) was visible. Organoid size showed the most robust correlation with luminescence-based viability measurements across all profiled organoid lines (Figure 3.8 b), while other features, such as DAPI, Actin, and membrane permeability (DeadGreen) intensity were less suitable to predict the viability of organoids. In conclusion, organoid size is an informative metric to approximate organoid viability, but is biased by line-specific differences in organoid size.

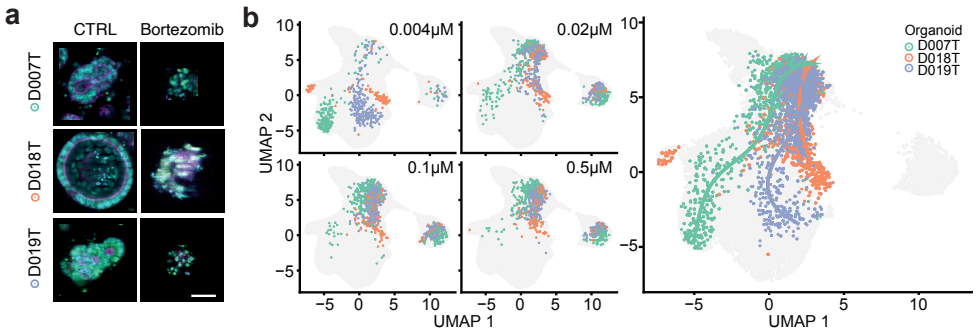

Figure 3.6: Organoid phenotype-profiles capture organoid viability. a Representative example images of negative- (0.1% DMSO) and positive control treated organoids (2.5µM bortezomib, cyan = DNA, magenta = actin, yellow = cell permeability; average images were selected and embedded in black background; scale bar: 50 µm). b, Dose-dependent-trajectory of bortezomib treatment effect. UMAP of organoid morphology at different bortezomib doses and (right panel) dose-dependent trajectory for three representative organoid lines. For visual purposes, trajectory inference was limited to partition 1, the left-hand set of measurements within the UMAP, representing ca. 95% of all imaging data. Panel a was created with support from Johannes Betge. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). Johannes Betge retrieved the example images for part a.

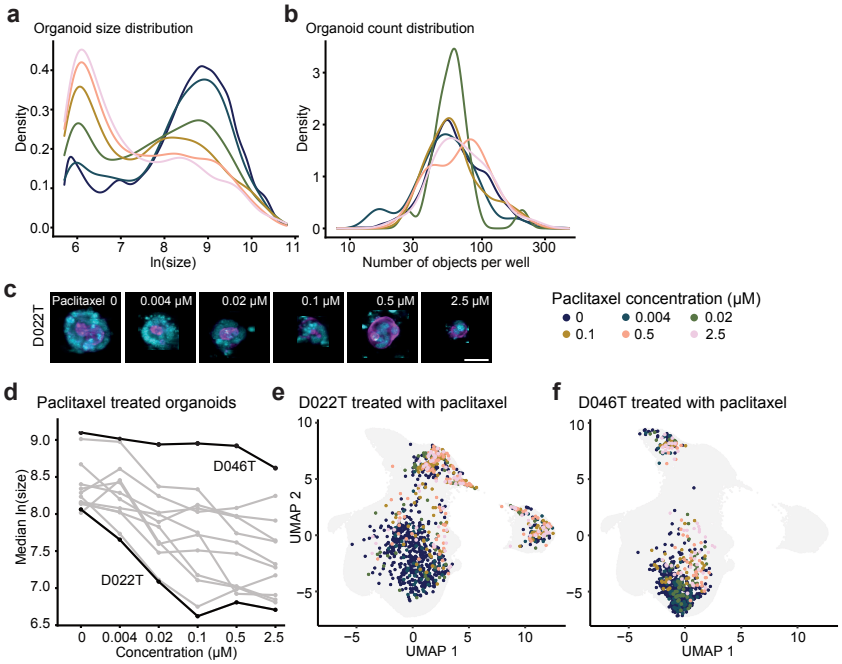

Figure 3.7: Organoid phenotype-profiles capture treatment-specific changes in organoid viability. a Distribution of organoid size at different concentrations of paclitaxel. Shown is a random sample of 30% of all paclitaxel treated organoids for this and following figures. **b** Distribution of organoid number per well at different concentrations of paclitaxel. **c** Example images of D022T organoids treated with paclitaxel. **d** Dose-response relationship of organoid size and paclitaxel dose. D022T and D046T are highlighted. **e** UMAP of organoid morphology, highlighting D022T organoids treated at different concentrations of paclitaxel. **f** UMAP of organoid morphology, highlighting D046T organoids treated at different concentrations of paclitaxel. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). Johannes Betge retrieved the example images for part c.

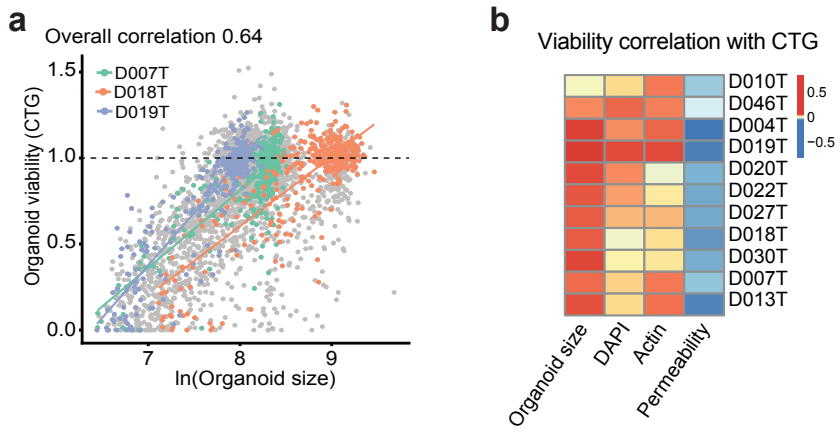

Figure 3.8: Organoid phenotype-profiles reflect ATP-dependent viability measurements. a Association of organoid size of selected example lines with organoid viability determined by luminescence-based, ATP-dependent viability profiling with CellTiter-Glo (CTG), which was performed in parallel with imaging on a subset of drug treatments. **b** Association of Organoid size and representative features (DNA, actin, and permeability dye intensities) with benchmark CTG viability read out. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022)

3.3.2 Drug induced organoid phenotypes correspond to drug mechanism of action

An advantage of image-based profiling over one-dimensional cell viability measurements in drug discovery is the ability to use the high dimensional drug-induced phenotype-profiles to identify small molecule mechanism of action based on similarity towards other, well-annotated, small molecules.

To test whether this approach could be used in cancer organoids, a weakly supervised learning approach to identify treatment effect vectors was developed with the support of Jan Sauer. Briefly, for every treatment and genotype, a logistic regression classifier was trained to distinguish DMSO-treated organoids from small molecule treated organoids. The resulting normal vector between control- and treated organoid profiles was referred to as the treatment effect vector. The classification performance, expressed as the AUROC, was used to determine the activity of a treatment, where the activity was defined (AUROC, mean from bootstrapped modeling with values ranging from 0.5 to 1). A high AUROC score (approaching 1) is observed for compounds that lead to a treatment-induced organoid morphology that is highly distinct from DMSO treated organoids. In contrast, a low AUROC (centered around 0.5) is observed for compounds where the model's classification performance approaches random guessing (Figure 3.9 a, b). I referred to the AUROC as the drug activity score. To test

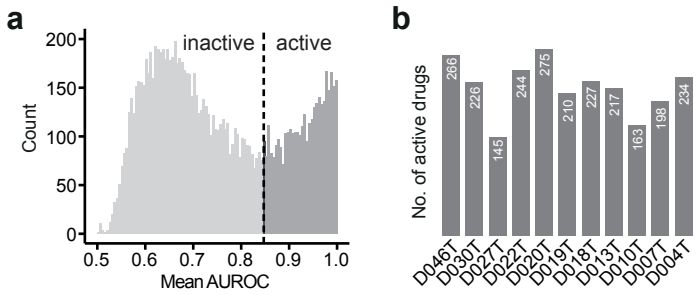

Figure 3.9: Distribution of AUROC drug activity scores. a Distribution of AUROC scores across all studied organoid models and treatments. **b** Number of active treatments per organoid line. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). Analysis was performed by Jan Sauer.

whether active drugs systematically lead to organoid phenotypes that are informative of mechanism of action, the cosine distance between concatenated treatment effect vectors was determined. This approach led to a clustering of specific mode-of-actions, including inhibitors of MEK, Aurora kinase, CDK, mTOR, AKT, EGFR or $GSK3$-beta (Figure 3.10 a).

Compounds with related, but not identical, targets also induced related phenotypes, for example MEK inhibitors clustered with specific RAF- and ERK inhibitors (Figure 3.10 c-f) or AKT and $PI3K$ inhibitors were part of a cluster mainly containing mTOR targeting compounds. Manual inspection of several phenotypes (Figure 3.10 g) revealed characteristic treatment-induced phenotypes. Most notably, MEK inhibitors led to reorganization towards more cystic organoid architecture. These target dependent phenotypes were observable across organoid lines and drugs.

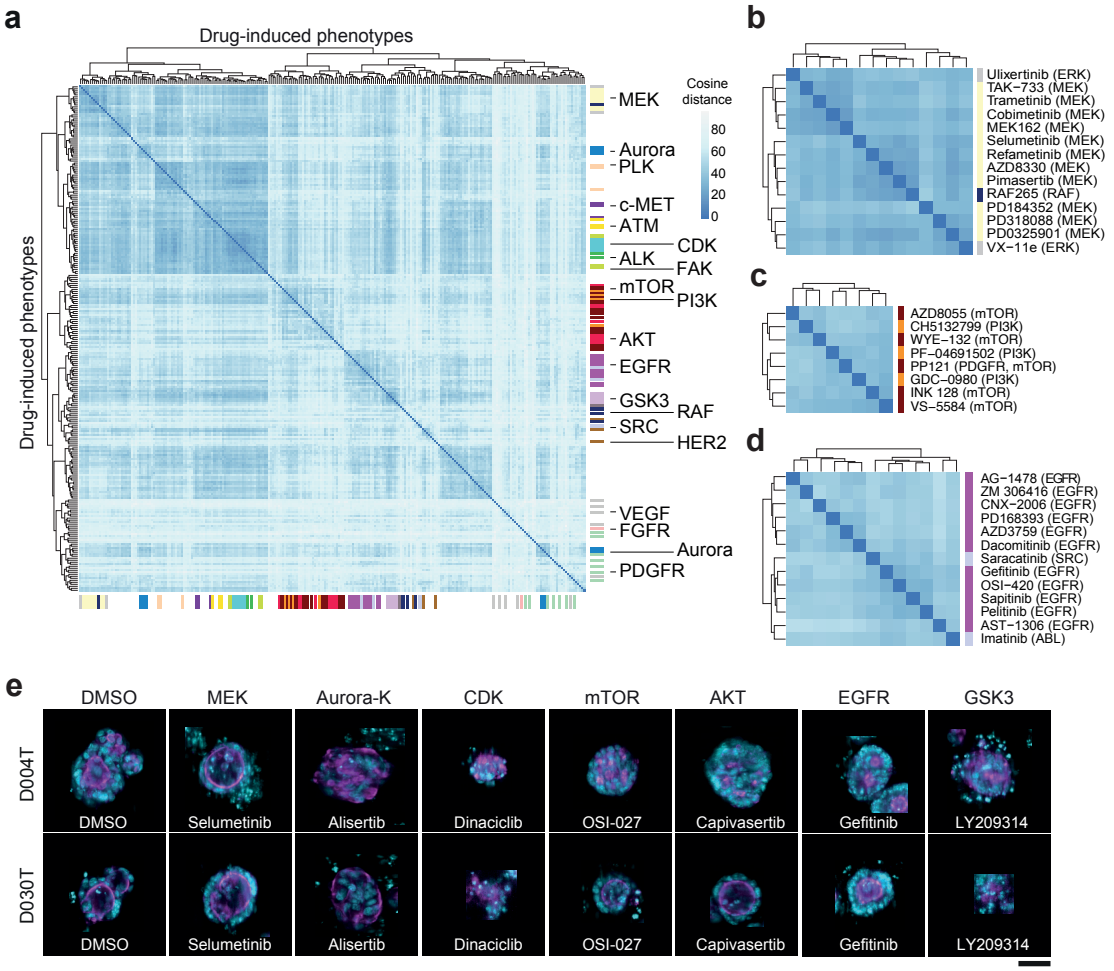

Figure 3.10: Clustering of treatment effects captures mechanism of action. a Hierarchical clustering of treatment effect vectors across all observed organoid models. Clustering is based on the cosine distance of concatenated treatment effect vectors. Sidebars are color-coded by the primary annotated target. **b-d** Magnified regions from panel a showing clusters of small molecule inhibitors targeting MEK, mTOR, and EGFR. **e** Representative images of treatment-induced organoid phenotypes for seven clusters of small molecules (cyan = DNA, magenta = actin, yellow = cell permeability; average images were selected and embedded in black background; scale bar: 50 µm). Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). Analysis was performed by Jan Sauer and retrieval of example images in part e was done in collaboration with Johannes Betge.

To validate the previous results with a second analysis approach, alternative methods to 1) quantify treatment activity and 2) cluster treatment effects by similarity were evaluated. The AUROC score, which was previously used to define active treatments, correlated with the Euclidean distance of phenotype profiles (Spearman correlation = 0.91). Similarly, evaluating an alternative clustering method based on phenotype profile averaging across lines (instead of concatenating) followed by clustering recovered an overall similar grouping of treatments (Figure 3.11 a).

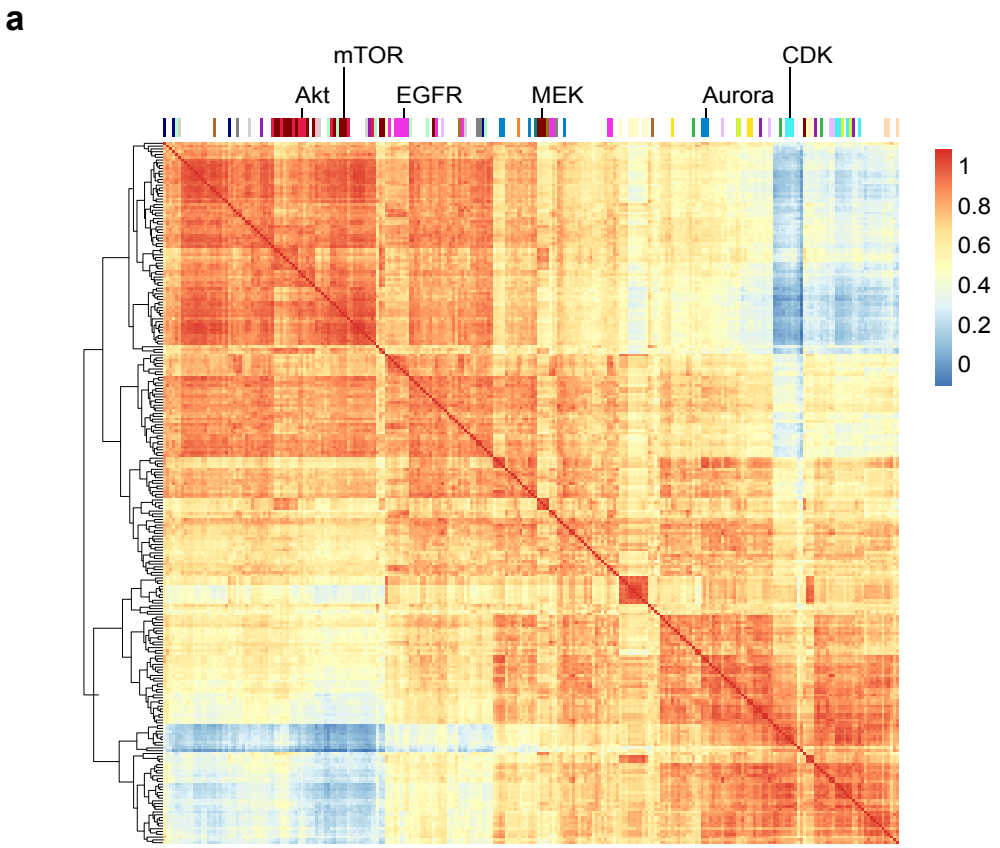

Figure 3.11: Validation of treatment activity and effect quantification method. a Hierarchical clustering of treatment effect vectors across all observed organoid models. Clustering is based on the Euclidean distance of average treatment effect vectors. Sidebars are color-coded by the primary annotated target.

3.4 Multi-omics factor analysis identifies factors linking morphology, genomic data and drug activity

A limitation of image-based profiling experiments is that both unperturbed and treatment-induced phenotypes are challenging to interpret. I hypothesized that, in the presence of multiple *in-vitro* models with both phenotype and transcriptome measurements, links between the two data modalities can be learned.

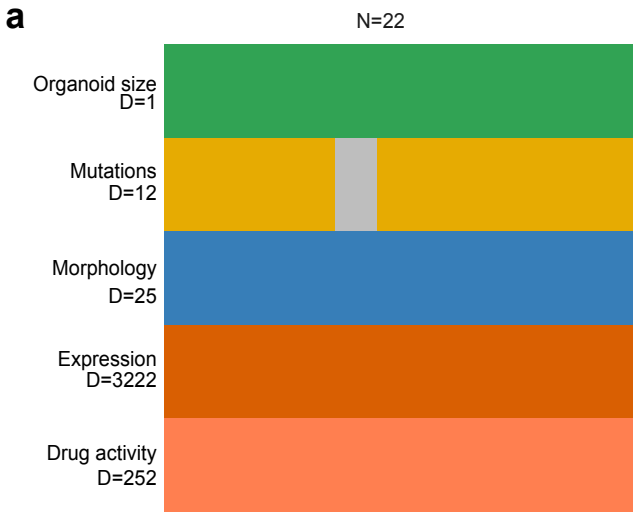

Figure 3.12: Overview of input data for MOFA modeling. a 22 observations of 11 models across 5 different views (size, mutation, morphology, gene expression and treatment activity) were integrated in the model. Missing data is shown in gray. "D" represents the vector length of the respective view. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022).

To learn a multi-view representation of unperturbed organoid morphology, unperturbed organoid size, transcript expression, somatic mutations, and drug activity, multi-omics factor analysis (MOFA) was performed. MOFA is a matrix factorization method built upon group factor analysis, with decomposes a set of different measurements into a shared matrix of factor scores for each observed sample and a set of corresponding matrices linking each factor to features in the set of original measurements through a weight or "loading" (Argelaguet, Velten, et al. 2018).

3.4.1 Learning a multi-view representation with MOFA

The support data used to fit the MOFA model consisted of a set of five matrices, each with 22 observations (rows) representing 11 organoid models with 2 replicates per model (Figure 3.14 a). Each matrix represented a modality (mutation, morphology, gene expression, etc.). When trained with a low number of k = 3 factors, MOFA recovered factors explaining between ca. 41 and 24% of variance across the different data modalities, with the

first two factors accounting for on average ca. 29 and 17% in aggregate, respectively (Figure 3.13 a). While gene expression, mutations and drug activity profiles for organoid lines contributed to all factors, factor 1 captured an exceptional amount of variation in median organoid size (ca. 39%). In contrast, factor 2 was primarily capturing variation within baseline organoid morphology (ca. 16%) (Figure 3.13 a).

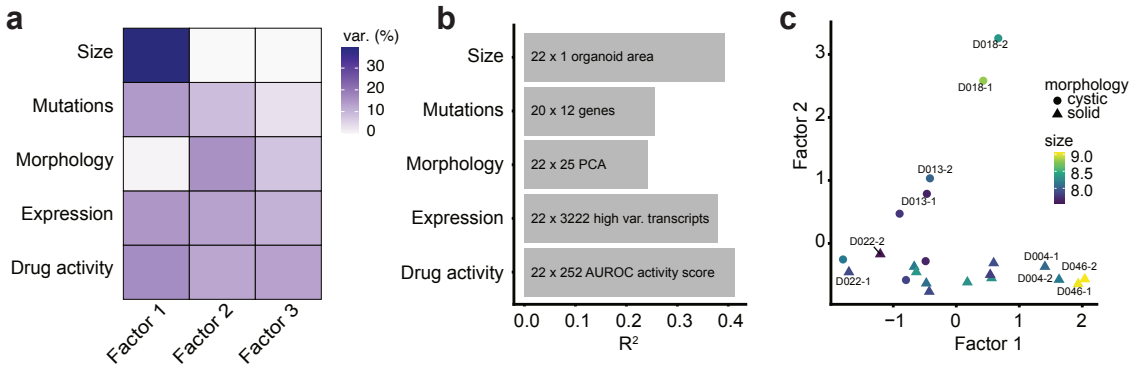

Figure 3.13: Multi-omics factor analysis of organoid profiles. a-b Variance decomposition of the MOFA model. Shown is the variance explained for every factor and data modality, as well as only by modality. **c** Factor scores for individual observations. The shapes represent the manually determined morphology label and color average organoid size. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022).

Overall, MOFA factors explained up to 40% of variance in median organoid size, drug activity and transcript expression, while less than 30% of variance in baseline organoid morphology was explained by the model (Figure 3.13 b). Organoid lines D046T and D004T stood out as lines with the strongest score for factor 1, while lines D018T and D013T had the strongest score in factor 2. Visual inspection of organoids revealed that organoid lines with a higher factor 1 score tended to be larger and organoids with high factor 2 score tended to have a more cystic organoid architecture based on manual classification (Figure 3.13 c). No interpretable morphological differences between factor 3 low and high organoids were identifiable, so the subsequent analysis was focused on the first two interpretable factors generated by MOFA.

To validate the MOFA finding and evaluate whether morphological differences could be seen in transcript expression data directly, I visualized organoid models by a set of previously determined manual morphology labels (Figure 3.14 a) and measured size (Figure 3.14 b) in a principal component analysis. The transcript expression data alone organized patient-derived organoid models by their manual morphology labels. To summarize, unsupervised multi-view representation learning with MOFA identified factors within the dataset that explained variation between organoid lines across different data modalities, including organoid morphology and median organoid size.

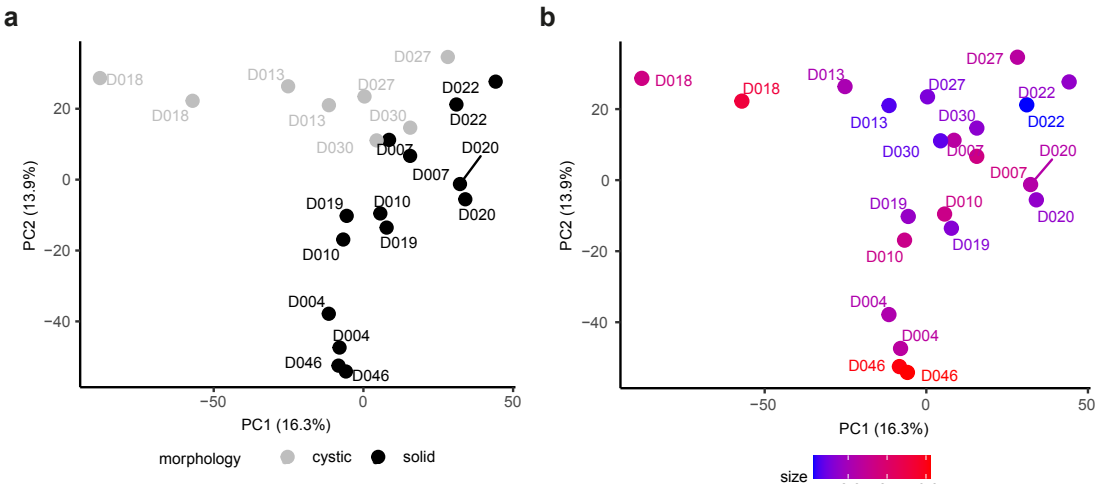

Figure 3.14: Validation of characteristic organoid phenotypes by transcript expression analysis. a PCA of transcript expression data for organoid models. Observations are color-coded by the manually determined morphology label. **b** Identical PCA of transcript expression data, color-coded by average organoid size.

3.4.2 An IGF1R signaling program is associated with increased organoid size, EGFR inhibitor resistance and can be induced by mTOR inhibition

Differences in organoid size are a contributing factor to intra- and inter-organoid line heterogeneity. Organoid size was influenced by both organoid line and drug treatments and was associated with factor 1 scores (Figure 3.15 a). An unsupervised gene set enrichment analysis (GSEA) for Reactome pathways across factor 1 loadings showed an enrichment for IGF1R signaling and mitogen-activated protein kinase signaling related genes (Figure 3.15 b).

In fact, the IGF signaling related transcripts H19 (rank 1) and IGF2 (rank 13) were among the strongest contributors to factor 1. This increase in proliferative signaling was confirmed by GSEA of a previously identified intestinal proliferation signature by Merlos-Suarez et al. (Merlos-Suárez et al. 2011) (Figure 3.15 c). To better understand clinical correlates to the identified gene expression patterns, I tested for an enrichment of molecular subtype profiles stemming from an analysis of cancer-cell intrinsic transcription (Isella et al. 2017a). Factor 1 showed an enrichment for CRIS-D, a molecular subtype linked to IGF2 overexpressing tumors, loss of IGF2 imprinting and resistance to EGFR inhibitor therapy (Figure 3.15 c). Conversely, I observed a depletion for CRIS-C, which has been linked to EGFR dependency (Figure 3.16 a and b). In fact, activity of EGFR inhibitors was the strongest contributor to a negative factor 1 score, while IGF1R and MEK inhibitor activity contributed to a positive factor 1 score (Figure 3.15 d-e and Figure 3.16 c-e).

To test whether small molecule treatments shifted organoids in factor space, the morphological features of treated organoids were projected into factor space. This projection was performed using the pseudoinverse of the loading matrix and multiplying it with the treatment-induced morphology features for treatments of interest. This approach enables interpreting treatment-induced phenotypes in a representation space that was previously learned from multi-view data of unperturbed organoids.

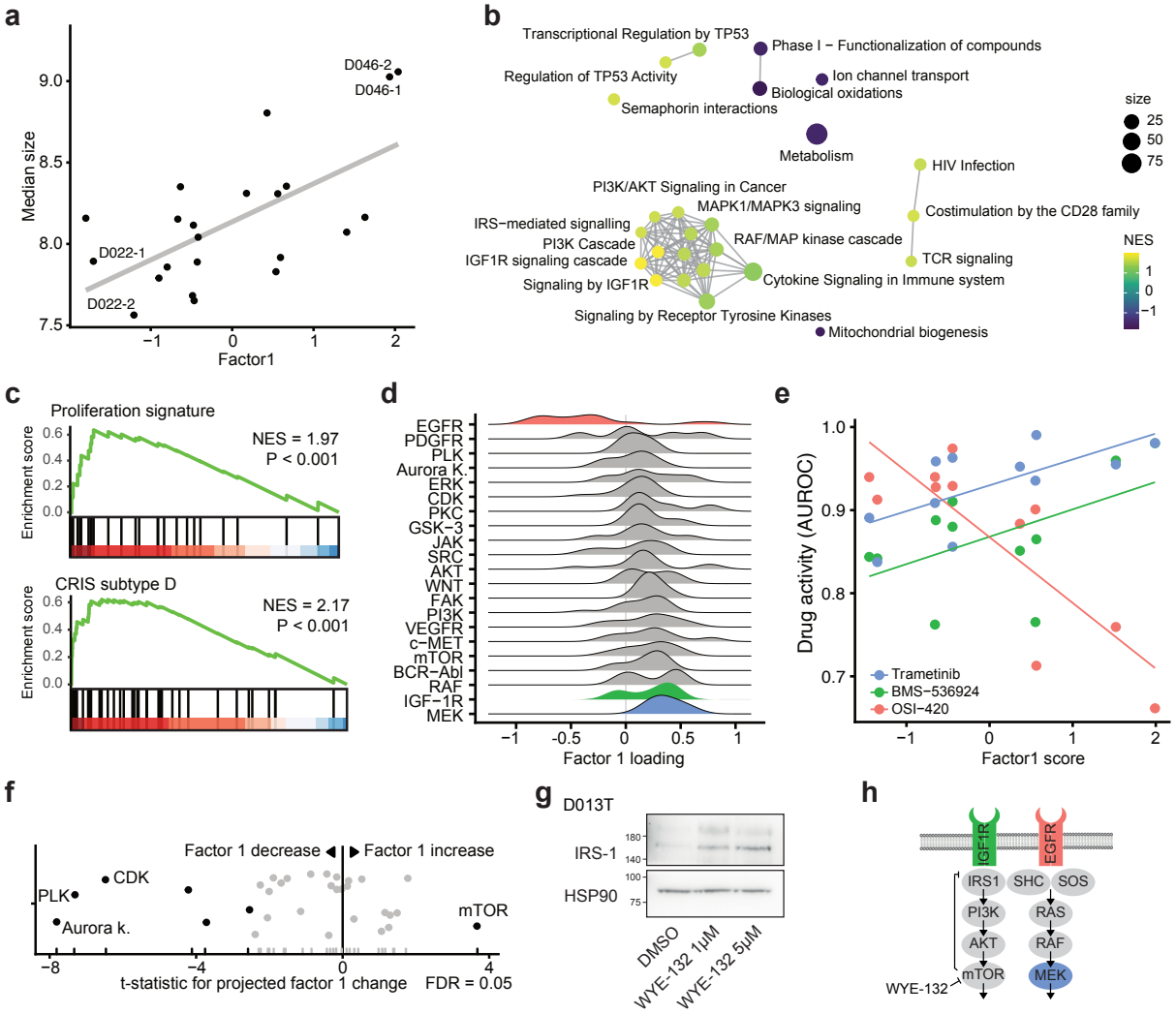

Figure 3.15: Factor 1 overview. **a** Association between organoid size and factor 1 score. **b** GSEA network of factor 1 gene expression loadings. An edge connects Reactome pathways with more than 20% overlap. **c** GSEA of the proliferation signature and the colorectal cancer CRIS-D subtype over ranked factor 1 gene expression loadings (ranking from high factor 1 to low factor 1 loading). **d** Distributions of treatment activity score loadings grouped by targets for factor 1. **e** Relationship of selected treatment activity scores (AUROC) with factor 1 score. **f** Projection of factor 1 scores for drug-induced phenotypes. Highlighted are drug targets leading to a significant change in projected factor scores across all organoid lines (ANOVA). **g** Western blot of IRS-1 protein abundance under mTOR inhibition for a representative organoid line (D013T). **h** Illustration of the IGF1R signaling pathway with highlighted drug targets. Shown is the disinhibition of mTOR mediated IRS-1 repression after treatment with mTOR inhibitors. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). Validating Western Blot and visual summary in parts g and h were contributed by Johannes Betge.

A group of cell cycle related kinase inhibitors targeting polo like kinases, Aurora kinases and cyclin-dependent kinases shifted organoids to a low factor 1 score (Figure 3.15 f). In contrast, mTOR inhibitor treatment increased factor 1 scores in cancer organoids (Figure 3.15 f). Given the observation that factor 1 was associated with IGF-1R signaling and mTOR inhibitor treatment led to an increase in factor 1 scores, I hypothesized that mTOR inhibition leads to a reactive upregulation of IGF1R signaling in cancer organoids. In fact, inhibition of mTOR signaling had previously been linked to transcriptional disinhibition of IRS-1 in a negative feedback loop (O'Reilly et al. 2006) and reactive induction of IGF1R signaling had previously been described as a resistance mechanism to small molecule mTOR inhibitors in cancer (S. V. Sharma et al. 2010). When testing this hypothesis in patient-derived organoids, a dose-dependent increase of IRS-1 protein abundance in organoids treated with the ATP competitive mTOR inhibitor WYE-132 was observable (Figure 3.15 g and h).

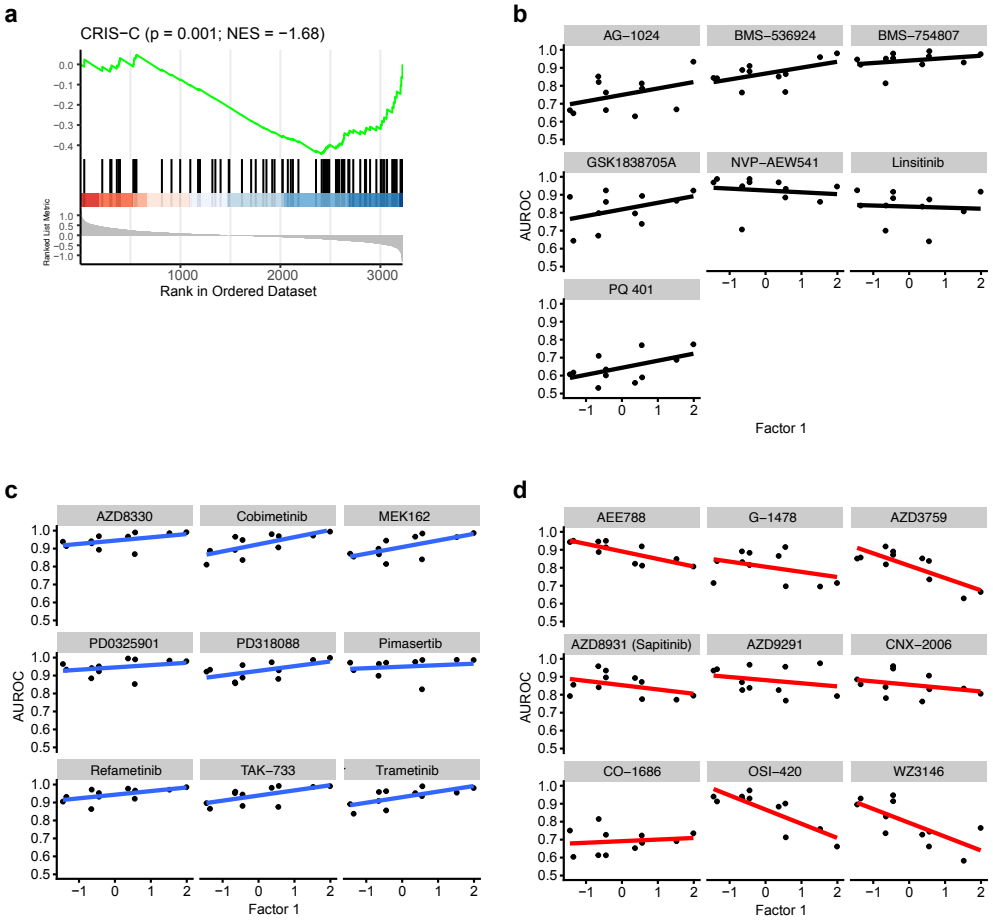

Figure 3.16: Factor 1 extended overview. a GSEA of the CRIS C subtype signature over ranked factor 1 gene expression loadings. **b-d** Association between treatment activity scores and factor 1 scores for three classes of compounds, targeting IGF1R, MEK and EGFR, respectively. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022).

To summarize, factor 1 described an organoid state with increased organoid size, elevated IGF1R dependent mitogenic signaling and increased resistance to EGFR inhibitors. This state could be induced by inhibiting mTOR, leading to a disinhibition of a negative feedback loop controlling IGF1R signaling.

3.4.3 An LGR5+ program is associated with cystic organoid architecture, Wnt signaling inhibitor sensitivity, and can be induced by inhibition of MEK

Besides size differences, a particularly strong recurring organoid phenotype was the presence of a cystic organoid architecture, seen for example in untreated Do18T organoids and organoids treated with MEK inhibitors (Figure 4.7 and 3.10). In the cystic state, which was observed in factor 2 high organoid lines (i.e., Do18T), organoids consisted of a monolayer of uniform cells lining a central spherical lumen with a distinct apico-basally oriented actin cytoskeleton (Figure 3.17 a and b). This phenotype was reminiscent of organoid morphologies previously seen in *APC*textsuperscript-/- or Wnt ligand treated human intestinal organoids (Matano et al. 2015).

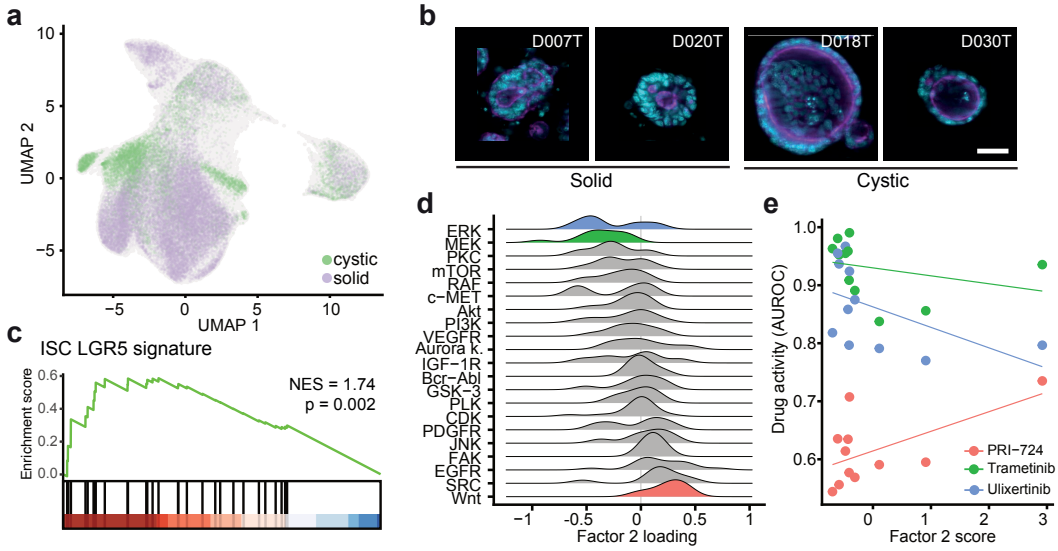

Figure 3.17: Factor 2 overview. a UMAP of observed organoids. Color labels represent the manually determined morphology labels. **b** Representative images of solid and cystic organoid phenotypes (cyan = DNA, magenta = actin, yellow = cell permeability; average images were selected and embedded in black background; scale bar: 50µm). **c** GSEA of LGR5 gene expression signature over ranked factor 2 gene expression loadings (ranking from high factor 2 to low factor 2 loading). **d** Distributions of treatment activity score loadings grouped by targets for factor 2. **e** Relationship of selected treatment activity scores (AUROC) with factor 2 score. Example images curated with support from Johannes Betge. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). Johannes Betge retrieved the example images for part b.

To test if factor 2 comprised Wnt signaling and intestinal stem cell identity related gene expression programs, gene set enrichment analyses (GSEA) was performed for cell identity signatures previously described by Merlos-Suarez et al. (Merlos-Suárez et al. 2011). GSEA revealed an enrichment of Lgr5+ stem cell signature-related genes for the factor 2 loadings (FDR=0.002, NES=1.74) (Figure 3.17 c). Activity of Wnt signaling inhibitors and EGFR inhibitors were the strongest average contributors to a positive factor 2 score (t statistic = 3.02, FDR = 0.046 and t statistic = 3.08, FDR = 0.046, respectively), while activity of ERK and MEK inhibitors were associated with a low factor 2 score (Figure 3.17 d), albeit not significantly. As expected from these results, factor 2 high organoid lines showed a stronger morphological response to the Wnt pathway inhibitor PRI-724. (Figure 3.17 e).

Next, I again projected phenotype profiles of treated organoids and approximated how drug treatment shifted

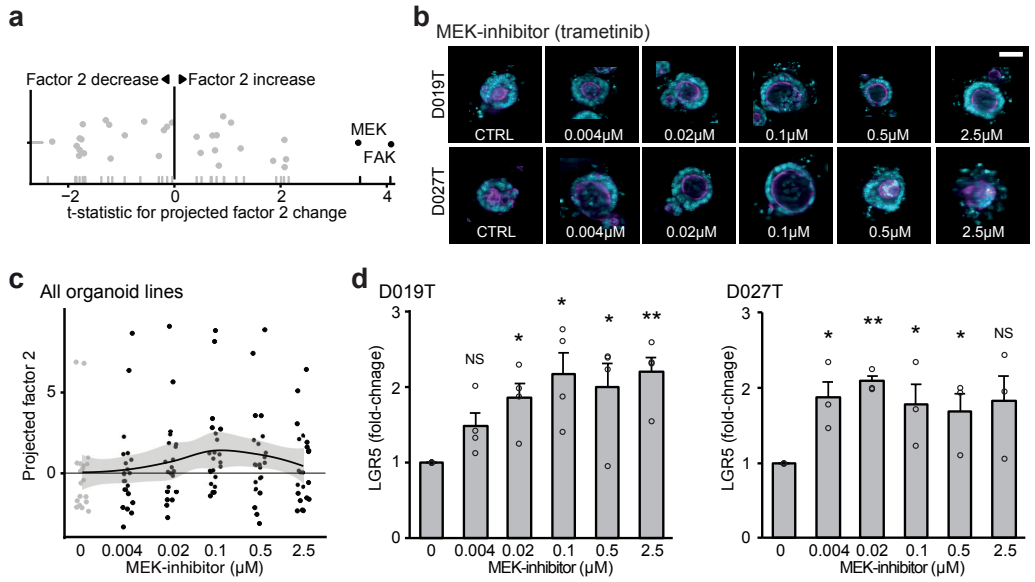

Figure 3.18: Factor 2 extended overview. a Projection of factor 2 scores for drug-induced phenotypes. Highlighted are drug targets leading to a significant change in projected factor scores across all organoid lines (ANOVA). **b** Representative images of organoid phenotypes across 5 increasing concentrations of MEK inhibitor treatment and negative control (0.1% DMSO) (cyan = DNA, magenta = actin, yellow = cell permeability; average images were selected and embedded in black background; scale bar: 50µm). **c** Projected dose-dependent changes in factor 2 scores after treatment with the MEK inhibitor binimetinib. The horizontal black line indicates the median factor 2 values of all binimetinib treatment observations. A Loess-fit with 95% confidence interval (gray background) is provided. **d** Dose-dependent changes in LGR5 transcript abundance after treatment with the MEK inhibitor trametinib, as assessed by RT-qPCR, data from 3 (D027T) and 4 (D019T) independent replicates are presented as mean + s.e.m. *p < 0.05, **p < 0.005, NS = not significant, two-sided Student's t test. p-values: D019T: p = 0.061 (0.004 µM), p = 0.0196 (0.02 µM), p = 0.0187 (0.1 µM), p = 0.024 (0.5 µM), P = 0.0024 (2.5 µM), D027T: p = 0.0051 (0.004 µM), p = 0.00038 (0.02 µM), p = 0.045 (0.1 µM), p = 0.048 (0.5 µM), P = 0.090 (2.5 µM). Example images curated with support from Johannes Betge. Figure adapted from *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). Johannes Betge retrieved the example images for part b. and contributed the validating RT-qPCR results in part d.

organoids along factor 2. I observed MEK and focal adhesion kinase inhibitors significantly shifting all tested organoid lines towards higher factor 2 scores (Figure 3.18 a). This change in factor 2 scores was concentration-dependent for MEK inhibitors and coincided with a visual shift in organoid morphology (Figure 3.18 b and c). Given the observation that factor 2 was enriched for an LGR5+ stem cell signature, the expression of LGR5 transcripts at different concentrations of MEK inhibitor treatment was measured and an analogous dose-dependent increase in transcript abundance was noticed (Figure 3.18 d). These findings were in agreement with the observation that MEK inhibitor activity had a negative contribution to factor 2 3.17 d). While organoids are shifted to a factor 2 high state by MEK inhibition, within the factor 2 high state itself, organoids are relatively insensitive to this treatment. In summary, factor 2 represents an organoid state with cystic architecture, increased expression of LGR5+ stem cell related genes and increased sensitivity to Wnt signaling inhibitors that could be induced by MEK inhibition.

What I cannot create, I do not understand.

Richard Feynman

4

Image-based Profiling of genetically engineered models of early colorectal cancer development

4.1 Motivation

In the previous chapter I demonstrated that (1) interpretable multi-view factors can be learned from patient derived organoid morphological and molecular state, (2) that these factors are associated with differences in small molecule treatment sensitivity, and (3) that non-toxic treatment-induced morphological changes can be interpreted using these factors to guide their mechanistic validation. Building on these observations, I wondered whether this approach could be extended to further understand early colorectal cancer pathogenesis and identify small molecule treatments that interfere with disease-associated states. To this end, a set of four genetically

engineered mouse colon organoid lines that model the initiation of colorectal cancer were generated. The models harbored tumorigenic *Apc*$^{-/-}$ and *Kras*$^{G12D/+}$ genotypes in isolation and combination, leading to four genetically distinct organoid lines. The organoid models were then characterized using transcript expression, proteomics and lipidomics measurements. A high-throughput image-based profiling experiment covering c.a 1700 well-annotated FDA-approved, natural and targeted small molecules was performed. The resulting data were integrated using

multi-omics factorization, and it was evaluated whether this multi-view representation model could be used to identify non-toxic small molecule treatments that acted specifically along previously identified axes of phenotype variation caused by tumorigenic *Apc*$^{-/-}$ and *Kras*$^{G12D/+}$ mutations.

4.2 Generation and characterization of organoid models

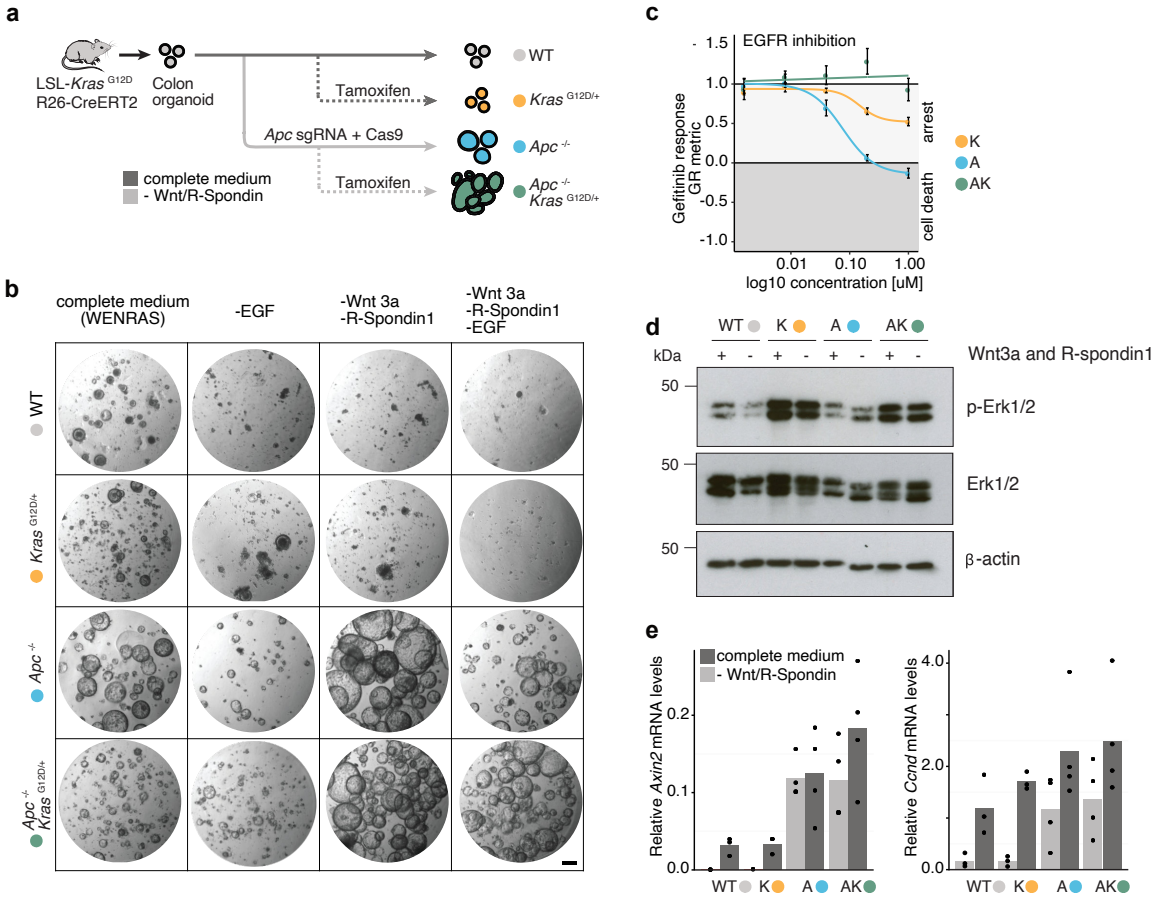

Figure 4.1: Establishing organoid models of colon adenoma, a Overview of organoid model establishment. Mouse colon organoids were isolated from a transgenic donor animal carrying an inactive conditional oncogenic *Kras*^{G12D} allele. Homozygous truncation of *Apc* via CRISPR and activation of the heterozygous *Kras*^{G12D} allele lead to four different genetically defined organoid models. **b** *in-vitro* growth factor dependency of adenoma models. Organoids were cultured in complete or modified medium containing combinations of Wnt3A, R-Spondin1-Fc and Egf for 120h and subsequently imaged. Scale bar = 200 µm. **c** Oncogenic *Kras*^{G12D} increases resistance to EGFR inhibitors. Organoid ATP levels were measured 4 days after gefitinib treatment and adjusted for organoid growth rate. Points represent the mean of n=2 independent experiments. Error bars represent the standard error of the mean. **d** Erk phosphorylation is increased by oncogenic *Kras*^{G12D}. Organoid models were cultured with or without Wnt3A and R-Spondin1-Fc for 72h and analyzed for protein levels. p, phospho. **e** Loss of *Apc* induces transcription of canonical Wnt-signaling target genes. qRT–PCR for Axin2 and Ccnd in the presence or absence of Wnt3A and R-Spondin1-Fc after 120h of culture. Expression levels are normalized to Sdha and Hprt transcript abundance. Bar graphs represent the mean of n=4 independent experiments. Wilcoxon rank sum test.

4.2.1 Generation of organoid colon adenoma models

The emergence of colorectal cancer via the chromosomal instability mechanism is a well-understood sequence of genetic events that start with hyperactivation of canonical Wnt signaling, i.e., through truncating mutations of *APC*, followed by the hyperactivation of ERK-MAPK signaling, i.e., via oncogenic mutations of *KRAS*. I genetically engineered four mouse colon organoid models carrying *Apc* truncating mutations, a mutant *Kras*^{G12D} allele, or both alleles in combination. Thereby, the first set of genetic events within the chromosomal instability mechanism was modeled.

To model the formation of colon adenomas *in-vitro*, I used a transgenic mouse to derive organoid cultures. The transgenic animal carried a conditional tamoxifen inducible *Kras*$^{G12D/+}$ allele (Jackson et al. 2001) (Figure 4.1 a). After isolation, I confirmed that extracted colon organoids did not express an activated form of *Kras*^{G12D} (Figure 4.2a) and defined these organoids as wild type (WT). To model loss-of-function mutations of the tumor suppressor *Apc*, the ortholog of the frequently mutated mutation-cluster-region on the *APC* gene was targeted by CRISPR (Figure 4.1 a).Wnt3A and R-Spondin1 ligands were removed from the organoid culture medium seven days after CRISPR treatment to select for loss-of-function mutations in the Apc gene. The resulting pool of organoids harbored heterogeneous biallelic loss-of-function mutations in *Apc* (Figure 4.2a). Subsequent activation of oncogenic *Kras*$^{G12D/+}$ by treatment with 4-Hydroxytamoxifen led to four distinct organoid adenoma models (Figure 4.1 a and Figure 4.2a-b); wild type (WT), *Apc*$^{-/-}$ (A), *Kras*$^{G12D/+}$ (K), and *Apc*$^{-/-}$ / *Kras*$^{G12D/+}$ (AK).

Similar to genetically modified human colon organoids (Drost et al. 2015; Matano et al. 2015), murine colon organoids showed characteristic niche requirements. Both *Apc* mutant organoid lines grew independent of the Wnt-signaling activating factors Wnt3A and R-Spondin1 (Figure 4.1 b) and at an accelerated proliferation rate. In fact, *Apc* mutant lines showed an increased growth in a Wnt3A and R-Spondin1 free environment when compared to the complete medium. Organoid models with an activated *Kras*^{G12D} allele were less sensitive to removal of Egf from the media. However, as observed before (Drost et al. 2015), the mutant *Kras*^{G12D} allele was insufficient to compensate completely for the loss of Egf from the medium. Nevertheless, *Kras*^{G12D} mutant organoid lines were more resistant to pharmacological inhibition of EGFR signaling (Figure 4.1 c). In conclusion, organoid model genotypes were reflected in characteristic growth factor dependencies in experimental conditions.

Next, I investigated the effects of mutations in *Apc* and *Kras* on both canonical Wnt- and Erk dependent signaling. While the presence of the *Kras*$^{G12D/+}$ allele led to an increase in Erk-phosphorylation across models, *Apc*$^{-/-}$ / *Kras*$^{G12D/+}$ organoids showed no marked additional increase in Erk-phosphorylation when compared to *Kras*$^{G12D/+}$ organoids (Figure 4.1 d). Moreover, *Apc*$^{-/-}$ / *Kras*$^{G12D/+}$ adenoma models showed no significant differences in expression of the Wnt target genes *Axin2* and *Ccnd* when compared to *Apc*$^{-/-}$ single-mutant models ($p > 0.34$ for both medium conditions, Wilcoxon rank sum test, Figure 4.1 e). These results indicate that organoid adenoma models show genotype-dependent activity of characteristic signaling pathways, while there is no

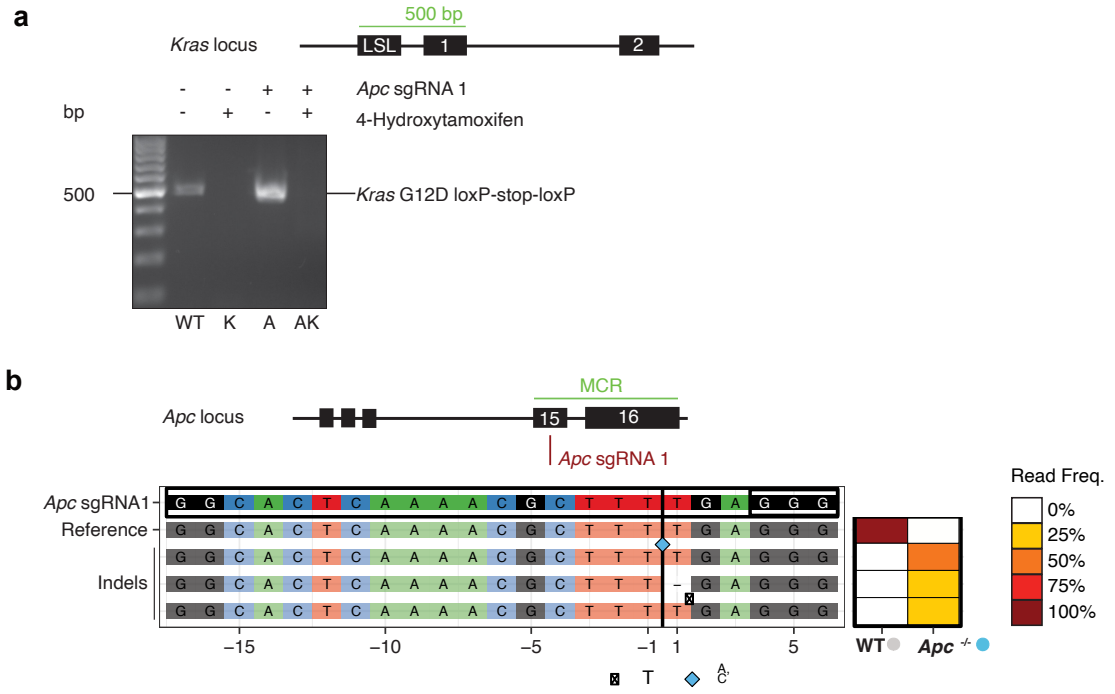

Figure 4.2: Structural validation of organoid colon adenoma models, a Allele-specific PCR products of colon organoid models isolated from a transgenic mouse with a conditional tamoxifen inducible *Kras*$^{G12D/+}$ allele. **b** Amplicon sequencing result of the murine mutation cluster region ortholog for organoids transfected with an *Apc* targeting sgRNA and Cas9 carrying plasmid. The sequencing results show the presence of 3 different insertion/deletions within the pool of sgRNA treated organoid models. wild type sequences are absent within the CRISPR targeted pool, while mutant sequences are absent in the untreated organoid pool.

extensive crosstalk between the *Apc*$^{-/-}$ and *Kras*$^{G12D/+}$ allele in mouse colon organoids that is reflected in the performed experiments to assess canonical Wnt- and Erk dependent signaling.

4.2.2 Molecular characterization of organoid models

To comprehensively characterize the four organoid models, untargeted transcriptome, proteome and lipidome profiling were performed using mRNA microarrays and mass spectrometry, respectively (Figure 4.3). For these measurements, organoids were kept in the same culture condition and duration that were used during the subsequent image-based profiling experiment: After passaging, all organoids were kept in Wnt3A, R-Spondin1 rich medium (WENRAS) to model conditions within the niche and stimulate outgrowth before the medium was changed after 72h to a Wnt3A-free medium (ENR) to model conditions outside the niche for 96h. Transcriptome profiling of organoid models showed an increased expression of the stem-cell marker Lgr5 and negative Wnt-signaling regulators such as *Nkd1*, *Notum*, *Wif1* and *Znrf3* in *Apc* mutant organoid lines (Figure 4.3 b). In contrast, *Apc* wild type organoid lines showed an increased expression of epithelial differentiation markers, such as *Krt20*, *Alpp* and *Abcb1* (P-glycoprotein). Overall, the number of genes with significant expression changes after *Apc* loss was 2.5 times greater compared to isolated $Kras^{G12D}$ activation (FDR = 0.1, $Apc^{-/-}$: 44.5%, $Kras^{G12D/+}$: 18.3% of assessed genes). A related observation was made during the analysis of protein abundance. Again, Wnt signaling regulators (Axin2, Notum) were frequently expressed in *Apc* mutant organoid lines and the number of proteins with significantly altered abundance after *Apc* loss was 2.5 times greater compared to an isolated $Kras^{G12D}$ activation (FDR = 0.1, $Apc^{-/-}$: 260, $Kras^{G12D/+}$: 105 assessed proteins). Principal component analysis of both transcriptome, proteome and lipidome data showed related axes of variation across measurements. In all observed modalities, the first principal component captured differences between *Apc* wild type and *Apc* mutant organoid models, while the second (in case of proteomics measurements the third) principal component captured differences between wild type and $Kras^{G12D/+}$ single-mutant models (Figure 4.3b, 4.3c and 4.3d). In every modality, a high degree of similarity was observed among $Apc^{-/-}$ and $Apc^{-/-}$ / $Kras^{G12D/+}$ organoid lines. While activation of oncogenic $Kras^{G12D}$ in wild type organoids led to global changes in transcript, protein and lipid expression, these changes were not as pronounced in organoids without functional *Apc*. In fact, only the mRNA expression of 91 genes was significantly altered between $Apc^{-/-}$ and $Apc^{-/-}$ / $Kras^{G12D/+}$ organoids (FDR = 0.1).

To explore active biological processes, gene set enrichment analysis on organoid transcript expression data was performed. The strongest changes in expression after loss of *Apc* were linked to an increased proliferation rate (Figure 4.3 e). Gene set enrichment analysis of published intestinal cell-proliferation and stem cell signatures showed an enrichment of both signatures in $Apc^{-/-}$ organoids (Figure 4.3 e) (Merlos-Suárez et al. 2011). In contrast, a signature for differentiating, transit-amplifying cells was depleted. Gene set enrichment analysis of $Apc^{-/-}$ / $Kras^{G12D/+}$ double-mutant organoids showed the same results. Next to these published signatures, I explored the enrichment of curated gene sets from the Reactome database (Griss et al. 2020). Here, both $Apc^{-/-}$ and $Apc^{-/-}$ / $Kras^{G12D/+}$ double-mutant lines showed a positive enrichment of cell cycle and DNA repair related genes when compared to wild type organoids (Figure 4.4a). Unique to the $Kras^{G12D/+}$ organoid line was a decreased expression of citric acid cycle and respiratory chain related genes (Figure 4.4b). This effect, was not

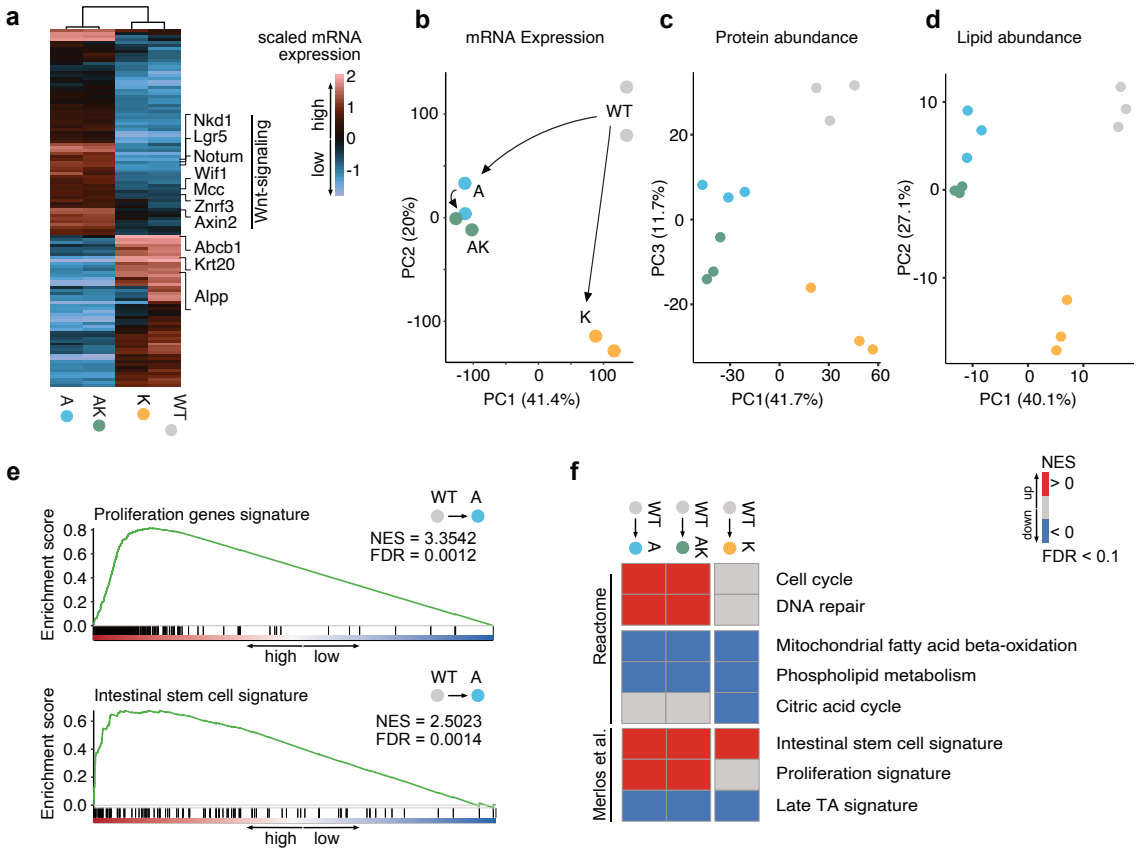

Figure 4.3: Molecular characterization of organoid adenoma models. a Differential gene expression of adenoma models. Shown are scaled expression values for the top 125 differentially expressed genes for every organoid line. Selected genes are highlighted. All organoids were cultured for 3 days in WENRAS before exposure to ENR for 4 days. Cell number was controlled between experiments. Whole organoid lysates were analyzed. **b** Transcript abundance data. Shown are the first two principal components of scaled gene expression data. The proportion of variance of each principal component is listed in parentheses. **c** Protein abundance data. Shown are the first and third principal component of scaled protein expression data. The proportion of variance of each principal component is listed in parentheses. **d** Lipid species abundance. Shown are the first two principal components of scaled lipid abundance data. The proportion of variance of each principal component is listed in parenthesis. **e** Loss of *Apc* leads to increased expression of proliferation and intestinal stem cell associated genes. Shown is a gene set enrichment analysis of differentially expressed genes between *Apc* mutant and WT organoids. Intestinal gene expression signatures were used according to Merlos-Suarez et al. NES, normalized enrichment score. **f** Overview of cellular processes in organoid adenoma models. Shown are selected enriched differential gene expression signatures from Reactome and Merlos-Suarez et al. NES, normalized enrichment score. NES > 0 suggests an enriched/ activated biological process. FDR < 0.1.

observed in *Apc*$^{-/-}$ / *Kras*$^{G12D/+}$ double mutant organoids (Figure 4.3f). In addition, organoid models with a *Kras*$^{G12D/+}$ genotype showed a downregulation of the *Egfr* receptor, in line with a potential negative feedback response to hyperactivated ERK-MAPK signaling (Figure 4.4b). Both *Apc*$^{-/-}$ and *Kras*$^{G12D/+}$ organoid models showed a strong reduction of lipid metabolism and beta-oxidation (Figure 4.4a,b). In summary, (1) loss of *Apc* leads to a global shift in transcript, protein, and lipid composition in colon organoids, including a strong increase in cell proliferation associated transcripts; (2) Activation of isolated oncogenic *Kras*^{G12D} leads to pronounced reduction in citric acid cycle related transcripts while this phenotype was not seen in organoid models with an additional loss of *Apc*; (3) Both *Apc* loss and activation of oncogenic *Kras*^{G12D} lead to a reduction of lipid beta-oxidation related transcripts.

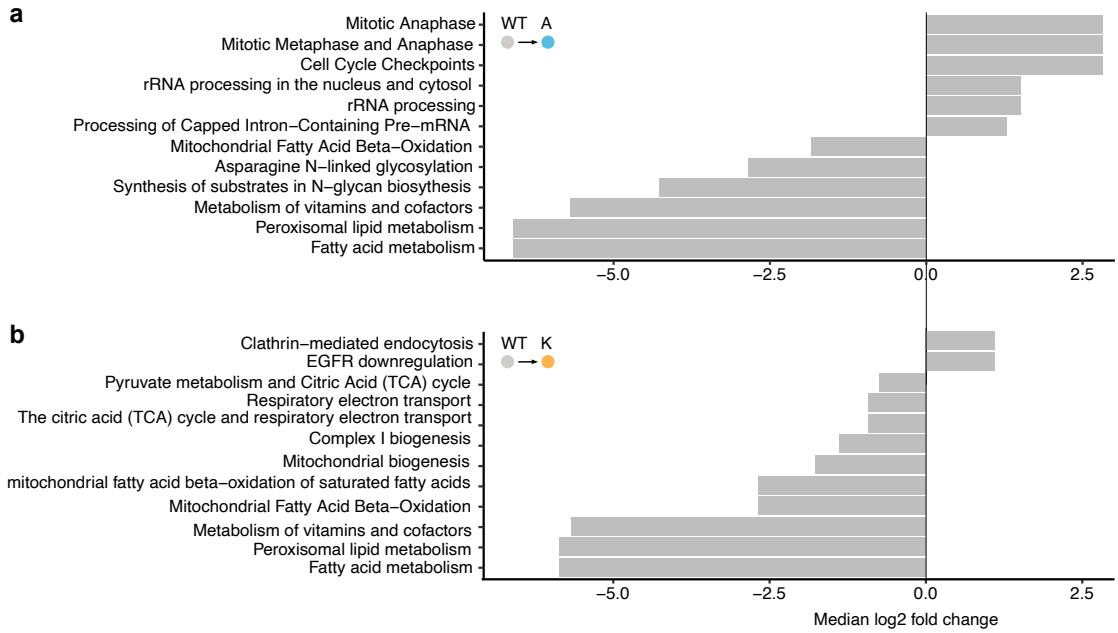

Figure 4.4: Representative up and down-regulated transcriptional processes after loss of *Apc* and activation of oncogenic *Kras*^{G12D}. **a** Representative up and down-regulated transcriptional processes after loss of *Apc*. Expression signatures were sourced from Reactome and average $\log_2$ fold changes for included transcripts are illustrated. FDR < 0.1. **b** Representative up and down-regulated transcriptional processes after activation of oncogenic *Kras*^{G12D}. Expression signatures were sourced from Reactome and average $\log_2$ fold changes for included transcripts are shown. FDR < 0.1.

To further understand the pronounced changes in fatty acid metabolism that were indicated by transcript expression changes in both *Apc*$^{-/-}$ and *Kras*$^{G12D/+}$ organoid models, differences in lipid composition were measured using untargeted lipid extraction and mass spectrometry. In total, more than 350 lipids from 15 species were identified across all samples (Figure 4.5 a). The majority of identified lipids had fatty acid chain lengths of 20 to 40 carbon atoms, while triglycerides (TAG) had an increased length of 40 to 60 carbon atoms (Figure 4.5 b). Major differences between organoid lines were seen especially for storage lipids -triglycerides (TAG) and cholesterol esters (CE) (Figure 4.5 c and d, respectively). Both lipid species were more abundant in the two *Apc*$^{-/-}$ organoid lines (TAG: $t > 4.8$, CE: $t > 3.7$, $p < 0.05$ ANOVA). In single-mutant *Kras*$^{G12D/+}$ organoids, triglycerides were also more abundant compared to wild type organoids ($t = 5.9$, $p < 0.05$ ANOVA), while cholesterol esters were depleted ($t = -3.7$, $p < 0.05$ ANOVA). The increase in the abundance of storage lipids as a result of *Apc* loss of function and oncogenic *Kras*^{G12D} mutation were aligned with the previously described transcriptional changes that indicated a reduced rate of beta-oxidation in these models (Figure 4.4a,b). In summary, both *Apc*$^{-/-}$ and *Kras*$^{G12D/+}$ genotypes cause an accumulation of storage lipids, including triglycerides, which is likely the result of reduced lipid catabolism.

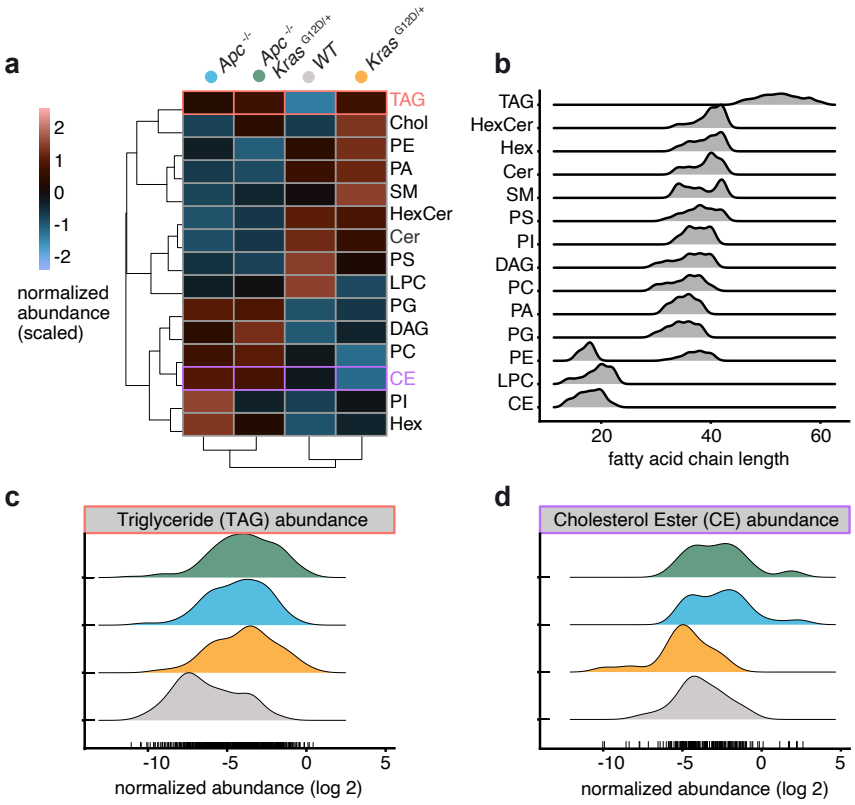

Figure 4.5: Lipid composition changes across organoid adenoma models. a Lipid species abundance of adenoma models. Shown are scaled abundance values for major lipid species for every organoid line. Selected lipid species are highlighted (Lipid Maps Abbreviations: HexCer - Hexosylceramide; PA - Phosphatidate; Cer - Ceramide; SM - Sphingomyelin; LPC - Lysophosphatidylcholine; PS -Phosphatidylserine; PE - Phosphatidylethanolamine; PI - Phosphatidylinositol; PG - Phosphatidylglycerol; DAG - Diacylglycerol; PC - Phosphatidylcholine; TAG - Triacylglycerol; Hex2Cer - Dihexosylceramide; CE - Cholesterol ester, Chol - Cholesterol). All organoids were cultured for 3 days in WENRAS before exposure to ENR for 4 days. Cell number was controlled between experiments. Whole organoid lysates were analyzed. **b** Fatty acid chain lengths by lipid species. Shown are the distributions of fatty acid chain lengths. **c** Distribution of triacylglycerol (TAG) abundance across organoid adenoma models. ANOVA was performed to model average lipid abundance as a function of organoid line across replicates. **d** Distribution of cholesterol ester (CE) abundance across organoid adenoma models. ANOVA was performed to model average lipid abundance as a function of organoid line across replicates.

4.3 Image-based profiling of organoid models

Once models were characterized on a molecular level, the previously developed image-based profiling approach was applied. Organoid models of the four different genotypes were perturbed with a library of ca. 1700 compounds and morphological profiles were systematically observed (Figure 4.6 a). As previously described, classic morphological features were collected for single organoids, normalized, and principal components were calculated, of which 25 components (accounting for >80% of the variance) were retained and used throughout the further analysis.

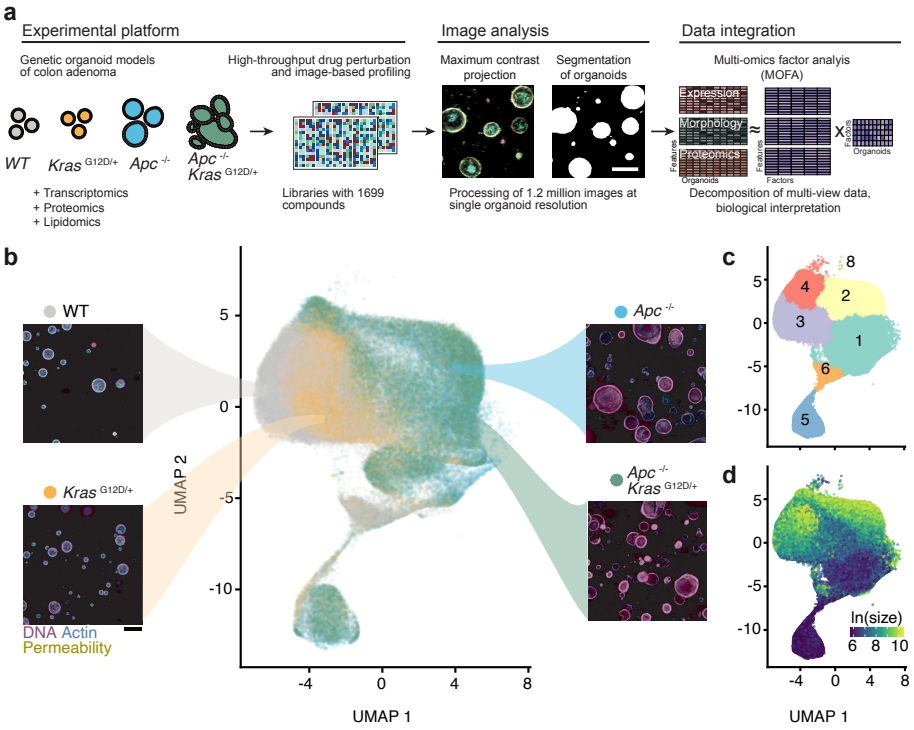

Figure 4.6: Image-based profiling of organoid adenoma models. a Overview of experiments. Organoids were isolated from a transgenic mouse model and genetically edited. Organoids were dissociated and evenly seeded in 384-well plates before perturbation with an experimental small molecule library. After treatment, high-throughput fluorescence microscopy was used to capture the morphology of organoids in 16 selected z-layers and 3 channels. 3D imaging data were projected on a 2D plane using a maximum contrast projection. Here, only pixel areas with the largest contrast among the z-axis were retained. Morphological features were computed based on the projection. Untreated organoid morphology, organoid size and treatment activity scores were integrated with transcript expression, protein abundance, lipid abundance and genotype data in a Multi-Omics Factor Analysis (MOFA) model. Figure created with support from Johannes Betge (graphical presentation). **b** Uniform Manifold Approximation and Projection (UMAP) of organoid-level features for a random 5% sample out of imaged organoids. The identical sample is used for visualizations throughout the figure. Organoid genotype is color-coded and representative images are displayed (magenta = DNA, cyan = actin, cell permeability = yellow, scale-bar: 200 µm). **c** Graph-based clustering of organoids by morphology, with 8 resulting clusters. **d** Organoid size distribution. Color corresponds to the log-scaled organoid area (dark blue: minimum size, yellow: maximum size). Parts of the figure are based on *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022)

4.3.1 Single-Organoid level phenotypes are organized by model genotypes

A UMAP projection of single-organoid morphology showed distinct genotype-dependent morphological states for identified organoids (Figure 4.6 b). Graph based clustering of organoid morphology profiles resulted in 8 clusters (Figure 4.6 c). Organoids within cluster 4 and 3 were enriched for *Apc*$^{+/+}$ organoid models, cluster 2 and 1 were populated by *Apc*$^{-/-}$ models. Analogous to transcript expression, lipidomics and proteomics representation space, the two *Apc* mutant organoid models were less distinct from each other, than organoids with a WT and isolated *Kras*$^{G12D/+}$ genotype (Figure 4.7 b). As observed in patient-derived organoids in the previous chapter, the distributions of DMSO-treated organoids and small molecule perturbed organoids in morphology feature overlapped considerably (Figure 4.7 a).

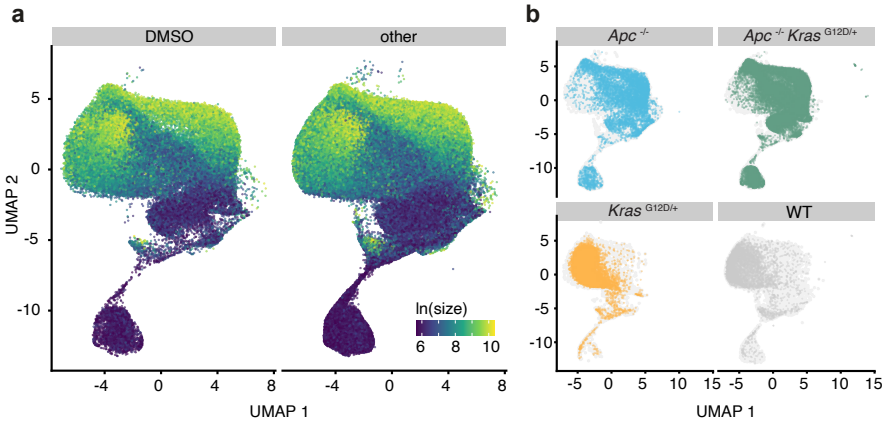

Figure 4.7: Treatment and genotype dependent effects within the organoid morphology distribution. a UMAP representation of DMSO treated (vehicle) and small molecule treated organoids. **b**, UMAP embeddings of four organoid genotypes (baseline state = 0.1% DMSO control-treated organoids), gray background consists of randomly sampled organoids.

When comparing the morphologies of different organoid lines in detail, characteristic differences were identifiable (Figure 4.8 a). DMSO-treated *Apc*$^{+/+}$ organoids showed a strong, regular apical actin cytoskeleton (high average actin intensity) that organized the multicellular formation into a symmetric regular-patterned spherical morphology (low average eccentricity). In contrast, *Apc*$^{-/-}$ organoids showed a relative lack of a regular actin cytoskeleton (low average actin intensity) and an irregular, non-symmetric, non-spherical morphology (high average eccentricity). In summary, organoid models showed genotype-dependent differences in morphology. Analogous to differences in molecular state, a primary source of variation was caused by loss of the tumor suppressor gene *Apc*. Organoids with *Apc*$^{-/-}$ presented with a higher proliferation rate, increased overall DNA staining intensity and loss of the regular spherical apical actin cytoskeleton that was observed in *Apc*$^{+/+}$ organoid models.

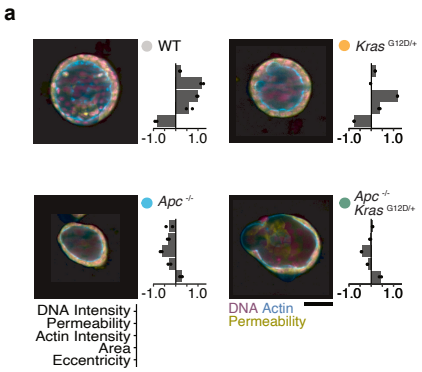

Figure 4.8: Genotype dependent effects on organoid morphology. a Morphological organoid profiles from vehicle-treated adenoma models were aggregated. Shown are representative individual organoids with selected features. Points show the mean phenotype for each independent biological replicate. Representative, interpretable features and their z-scores relative to all single organoid profiles are shown (magenta = DNA, cyan = actin, cell permeability = yellow, scale-bar: 25 µm)

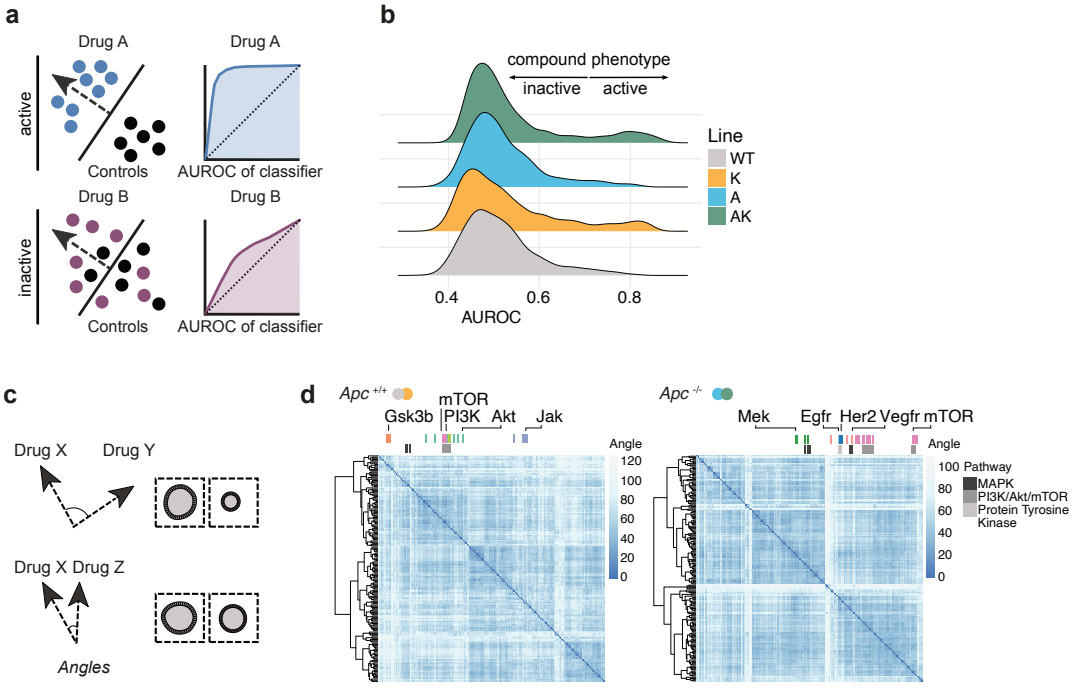

Figure 4.9: Treatment activity scoring. a A logistic regression classifier is trained to distinguish morphology profiles of individual treated and untreated organoids across all available replicates. Afterward, the classifier is applied to a validation set of organoids and the classification performance is estimated using the area under the receiver operating characteristic curve (AUROC) metric. Method implemented by Jan Sauer. **b** Distribution of centered treatment activity scores for all organoid lines, replicates, and perturbations. **c** Identifying related treatment-induced phenotypes. Normal vectors of treatment-specific classifiers were compared by calculating the angular distance (related to cosine similarity, ranging from 0-180 degrees). Small angular distance between vectors corresponds to a high similarity between the treatment-induced organoid phenotypes. Method implemented by Jan Sauer and published in *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). **d** A clustered heatmap of compound induced phenotypes for *Apc* mutant and *Apc* wild type organoids. Highlighted are clusters of compound induced phenotypes with related targets. Normal vectors for *Apc* mutant and *Apc* wild type organoids were concatenated before angular distance calculation. Method implemented by Jan Sauer and published in *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022).

After identifying genotype-dependent morphological differences, the next step was to explore the effects of small molecule treatment on different organoid models. To describe the activity of a treatment, the classification-based approach developed during the study of patient derived cancer organoid phenotypes in the previous chapter was used (Figure 4.9 a). The distribution of AUROC activity scores across organoid lines showed that most compounds did not lead to an identifiable morphological change (Figure 4.9 b). Next to identifying differences around the number of active treatments, I was interested in what small molecules were active in a given organoid genotype and their treatment-induced morphology change. Based on the observation that the primary source of variation for treatment activity was the state of the *Apc* allele, organoid lines were aggregated by their *Apc* allele for further analysis. In line with the approach taken in the previous chapter, normal vectors of the logistic regression classifiers were compared using the cosine distance (Figure 4.9 d). The resulting clustering of treatments showed an enrichment for small molecules with related mechanism of action (Figure 4.9 e and f). For example, EGFR inhibitors were significantly enriched in *Apc*-mutant organoid lines, while GSK3-beta-inhibitors,

which leads to a stimulation of canonical Wnt signaling, were only enriched in *Apc*-wild type organoid models.

4.4 Multi-omics factor analysis identifies factors linking molecular state and morphology

To jointly model the biological and morphological state of the four organoid lines, I performed multi-omics factor analysis (MOFA). Analogous to the process described in the previous chapter, molecular and morphological features from untreated organoids were factorized using k=4 factors (Figure 4.10 a and 4.11 a). The learned model was based on both morphological (e.g., morphology, size, small molecule activity) and molecular (e.g., genotype, proteomics, lipidomics and transcript expression) information. As done previously, to reduce the dimensionality of data modalities with a high number of features, only high variance features from gene expression and proteomics analysis were used.

4.4.1 Learning a multi-view representation with MOFA

The resulting factorization explained the data, ranging from ca. 90% (gene expression) to 50% (morphology) of explained variance (R^2) across the analyzed views. The first three factors captured the majority of variance, >40%, ca. 10%, and <10%, respectively (Figure 4.10 a). The learned model explained most variance within the mRNA expression and genotype data, while measurements within the organoid morphology data had the lowest explained variance (Figure 4.10 b). A similar low explained variance for morphology features was also observed previously in patient derived organoid modeling. When inspecting factor weights within the morphology domain closer, the principal component 1 feature, which accounted for ca. 30% of variance within all observed (treated and untreated) morphology data during the image processing had small relative assigned feature weights across all learned factors (from high to low absolute feature weight: rank 8/25, 11/25, 24/25 and 25/25 for factors 1 through 4, respectively). In contrast, the principal components 2, 3, and 4 had the highest assigned feature weights across all four factors. In summary, MOFA was used to learn k=4 factors across a set of distinct data modalities. In the morphology domain, the learned representation captured only a subspace of the observed morphology, possibly because of high noise levels in morphological data and the fact that only observations from untreated organoids were contained in the support set.

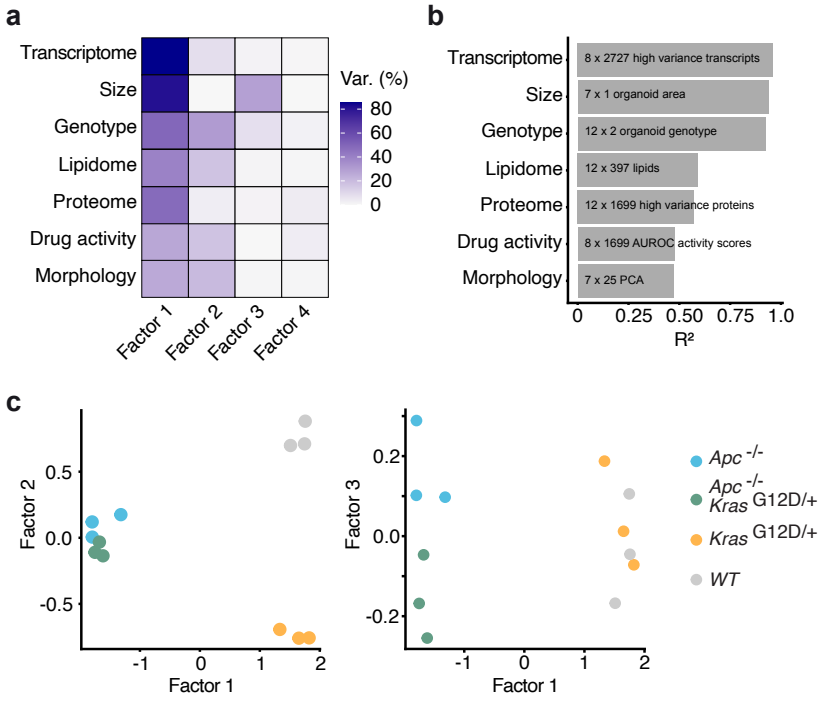

Figure 4.10: Multi-omics factor analysis (MOFA) to identify shared factors linking morphology, size, gene expression, lipidomics, proteomics, genotype and treatment activity. a Percent variance explained by the MOFA model for each factor. Untreated organoid morphology, organoid size and treatment activity scores were integrated with genotype, proteomics, lipidomics and mRNA expression data. **b** Cumulative proportion of total variance explained by each experimental data modality within the MOFA model. **c**, Visualization of samples in factor space, showing factors 1 and 2 as well as factor 1 and 3. Shown are independent replicates for each organoid line. Technical approach and related figure legend also used in *The drug-induced phenotypic landscape of colorectal cancer organoids*(Betge et al. 2022).

Visual inspection of factors as well as exploration of factor weights within the genotype view showed that factor 1 explained differences caused by *Apc* loss of function, while factor 2 explained differences caused by the activation of $Kras^{G12D}$ in an $Apc^{+/+}$ genotype (Figure 4.10 c and 4.11 b). In contrast to factor 2, factor 3 captured differences between Kras+/+ and $Kras^{G12D/+}$ organoids with *Apc* loss of function. While the number of factors is a user-defined hyperparameter within MOFA, the method automatically drops excess factors if they are not considered effective based on an applied automatic relevance determination (ARD) (Argelaguet, Velten, et al. 2018).

Increasing the number of factors above k=4 in this analysis, did not lead to an increased number of interpretable factors. In fact, factor 4 already did not capture differences between organoid genotypes and was not interpretable from a biological perspective by the author (Figure 4.11 b). Of note, the determined number of explanatory

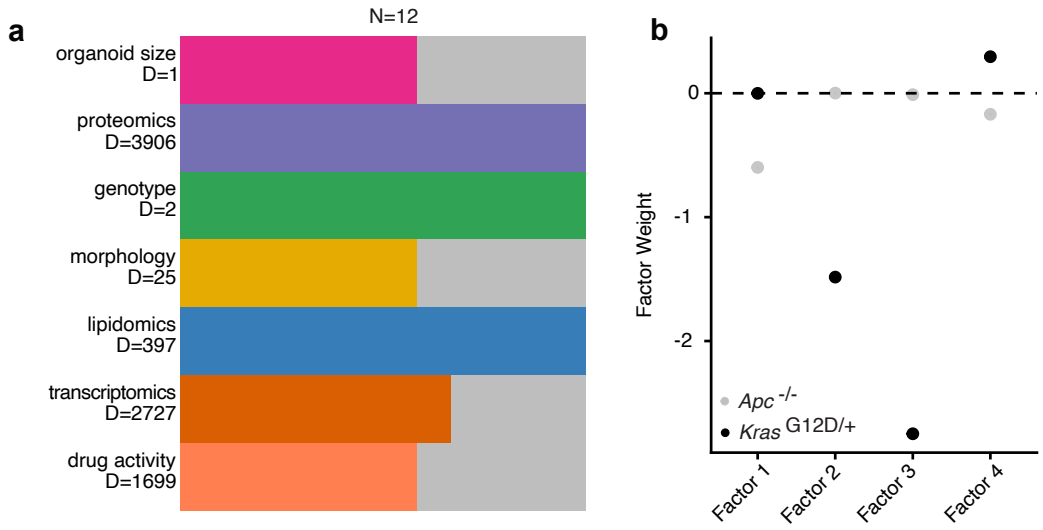

Figure 4.11: Multi-omics factor analysis input data and loadings. a Measurement modalities, dimensionality, and number of measurements. A third replicate of measurements was available for proteomics and lipidomics only. **b** Factor loadings for genotype information.

factors corresponded to the hypothesized intrinsic dimensionality of the data: effects attributed to the $Apc^{-/-}$ allele, the $Kras^{G12D}$ allele, and their interaction. Next, I inspected every factor and associated compound sensitivities separately in more detail.

4.4.2 Factor 1 separates $Apc^{+/+}$ from $Apc^{-/-}$ organoid models

To understand the molecular changes associated with factor 1, factor loadings for mRNA expression data were analyzed using Reactome gene-set enrichment analysis (Figure 4.12 a). Three clusters of biological processes were significantly associated with a negative factor loading, linked to *Apc* loss-of-function: 1) Mitotic Anaphase related processes, including spindle checkpoints; 2) Mitotic S-phase, including DNA replication and 3) DNA repair mechanisms, including homology directed repair.

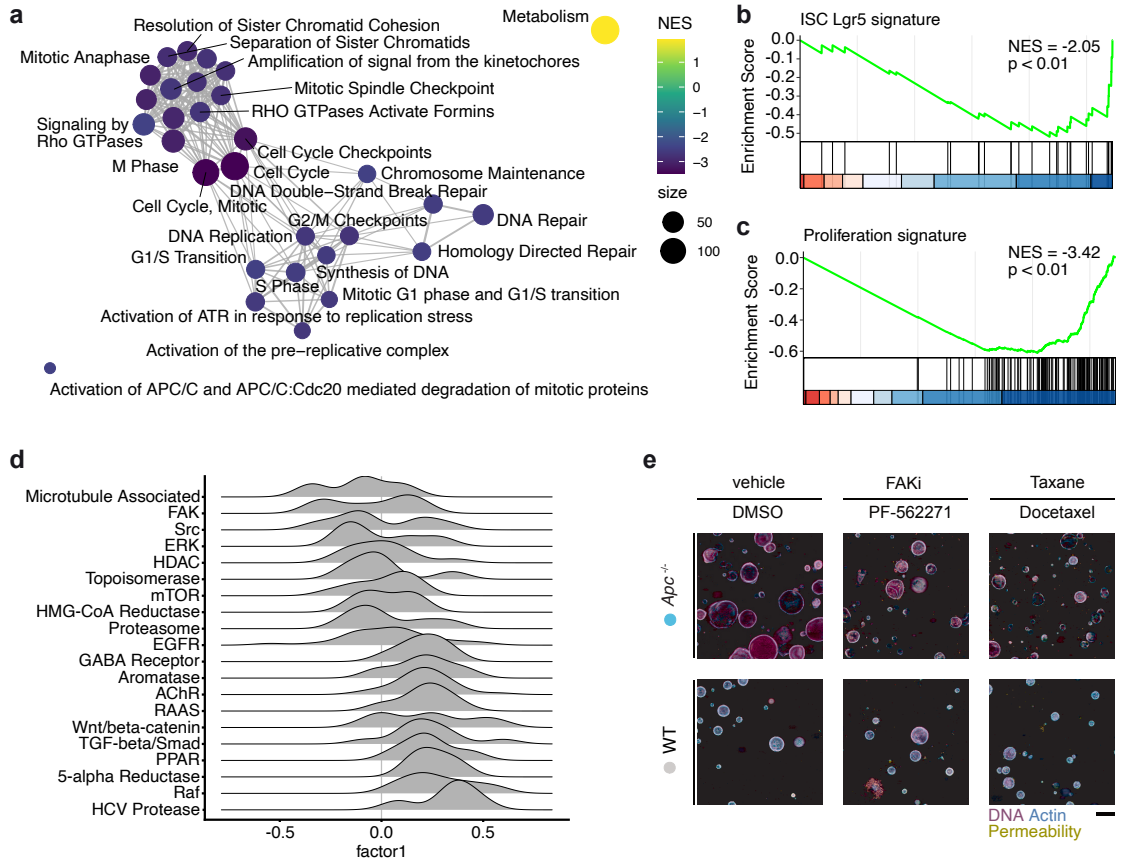

Figure 4.12: Factor 1 overview. a Gene-set enrichment analysis of factor 1 transcript expression weights. An edge connects Reactome pathways with more than 20% overlap. Central enriched processes include mitosis, DNA replication and DNA damage repair. **b and c** Gene set enrichment results of the "Lgr5 intestinal stem cell" and "proliferation" signature by Merlos-Suarez et al (Merlos-Suárez et al. 2011). (NES = normalized enrichment score). **d** Distributions of treatment activity weights by annotated drug target for factor 1. **e** Example images of compound treated organoids with WT or $Apc^{-/-}$ genotype. Representative images are displayed (magenta = DNA, cyan = actin, yellow = cell permeability, scale-bar: 200µm).

In line with the enrichment of processes associated with cell proliferation, factor 1 loadings showed an enrichment of the previously described intestinal proliferation signature (Figure 4.12 c) and an LGR5+ intestinal stem cell identity signature (Figure 4.12 b). These findings are in line with the long-standing evidence that loss of *Apc* leads to a hyperactivation of canonical Wnt signaling, which in turn leads to increased intestinal cell proliferation and Myc-dependent changes towards a stem-like cell state (Sansom et al. 2007; Satoh et al. 2017). When focusing on compound activity, a low factor 1 score was significantly linked to increased activity of microtubules and focal adhesion kinase (FAK) targeting small molecules (Figure 4.12 d). This morphological

sensitivity presented itself primarily as reduced organoid size and number relative to the DMSO vehicle control (Figure 4.12 e). In contrast, the average treatment activity scores of small molecules targeting Wnt signaling were associated with high factor 1 scores (Figure 4.12 d).

Further exploration of the association between the treatment activity score and *Apc* genotype showed that small molecule inhibitors of the protein Porcupine (Porcn), IWP-L6 and LGK-974, were more active in $Apc^{+/+}$ organoids relative to their $Apc^{-/-}$ counterparts (Figure 4.13a). Porcupine is a central mediator of Wnt ligand secretion (Liu et al. 2013). In contrast, this effect was not observable for PRI-724, a small molecule inhibitor targeting the interaction of beta-catenin and CREB-binding-protein in the canonical Wnt signaling pathway (Okazaki et al. 2019) (Figure 4.13a). The observed differences in treatment activity scores among small molecule inhibitors are likely related to their targets' relative location to Apc in the canonical Wnt signaling cascade. While Porcn-dependent Wnt secretion is generally upstream of the Apc-scaffolded destruction complex, the interaction of beta-catenin and the transcriptional coactivator CREB-binding-protein is located downstream of it. As a consequence, inhibition of destruction complex function by loss of *Apc* is expected to render cells less sensitive to perturbations of the Wnt secretion cascade than direct perturbations of transcription factor binding properties (Figure 4.13b).

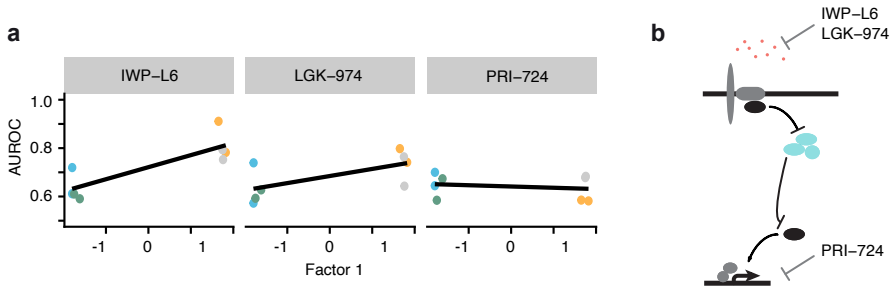

Figure 4.13: Activity of small molecule Wnt signaling inhibitors. a AUROC activity score for three small molecule inhibitors of canonical Wnt signaling. **b** Target proteins for small molecules within the canonical Wnt signaling cascade, with their relative position to the destruction complex (highlighted in blue).

4.4.3 Factor 2 separates *Kras*$^{G12D/+}$ from *Kras* wild type organoid models

While the *Apc*$^{-/-}$ genotype contributed primarily to factor 1 (Figure 4.11 b), the *Kras*^{G12D} allele showed a strong loading for both factor 2 and factor 3. This observation corresponded with the fact that only organoid models without *Apc* loss of function were separated by factor 2 (Figure 4.10 c). Factor 2 described a *Kras*^{G12D} dependent change in cell state in the presence of intact *Apc*.

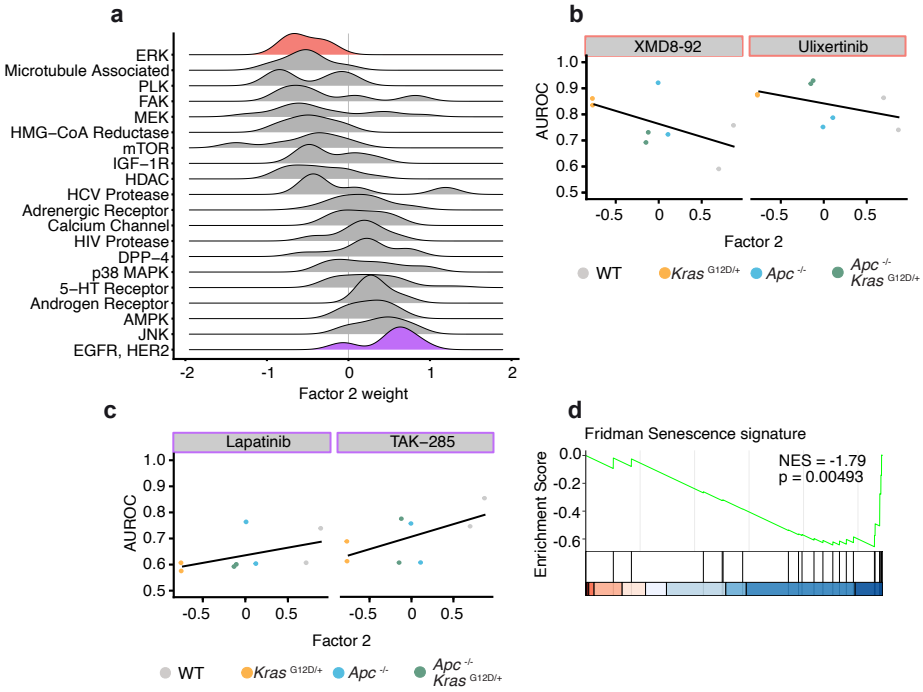

Figure 4.14: Factor 2, isolated *Kras*^{G12D} activity and oncogene induced senescence. **a** Distributions of treatment activity loadings grouped by drug target for factor 2. **b** Relationship of representative ERK inhibitor activity with factor 2 score. Shown are compounds from highlighted groups in panel (a). **c** Relationship of representative EGFR inhibitor activity with factor 2 score. Shown are compounds from highlighted groups in panel (a). **d** Gene set enrichment results of a senescence signature (Fridman et al. 2008) over factor 2 transcript expression weights (NES = normalized enrichment score).

As above, to understand the molecular mechanisms represented by factor 2, features with large absolute loadings were identified. Inhibitors of ERK and MEK were more active in factor 2 low models (*Kras*$^{G12D+/-}$) while EGFR/HER2 inhibitors were more active in factor 2 high organoids (WT, figure 4.14 a and b). This juxtaposition in treatment activity against ERK-MAPK pathway members was reminiscent of the previous observations made for canonical Wnt signaling inhibitors (Figure 4.13). With oncogenic *Kras* localized between the receptor-layer (including Egfr and Her2) and downstream mediating kinases (Erk), hyperactive *Kras* signaling likely leads to a cell state with relative resistance to EGFR inhibitors and increased dependency on Erk signaling. The previously observed transcriptional process of *Egfr*-downregulation as a response to *Kras*$^{G12D+/-}$ is in line with these observations (Figure 4.4b).

Oncogene induced senescence is a cell state marked by an arrest of the cell cycle and expression of pro-inflammatory mediators as a response to an oncogenic perturbation. An activated *Kras*$^{G12D+/-}$ genotype leads

to oncogene induced senescence of colon epithelial cells in vivo (Bennecke et al. 2010). Prompted by previous reports on the effect of an isolated oncogenic *Kras* allele, I identified an enrichment of a senescence related gene expression signature by Fridman et al. within the loadings of factor 2 (Fridman et al. 2008) (Figure 4.14 c). Transcripts linked to cell senescence, including Igfbp3 (factor 2 loading ca. -0.99) and Hmga2 (factor 2 loading ca. -1.05) ranked among the strongest contributors to the factor.

4.4.4 Factor 3 separates *Apc*$^{-/-}$ from *Apc*$^{-/-}$ / *Kras*$^{G12D/+}$ organoid models

While factor 2 captured the effect of oncogenic *Kras* in an *Apc* wild type state, factor 3 scores separated organoid models by *Kras* genotype in the context of an *Apc* loss of function (Figure 4.10 c). On average, factor 3 only accounted for less than 10% of variance across modalities within the MOFA model, supporting the overall similarity of *Apc*$^{-/-}$ and Apc-/- *Kras*$^{G12D/+}$ organoids previously observed separately on the transcriptome, proteome, lipidome and morphology level (Figure 4.3a-d and 4.7 b).

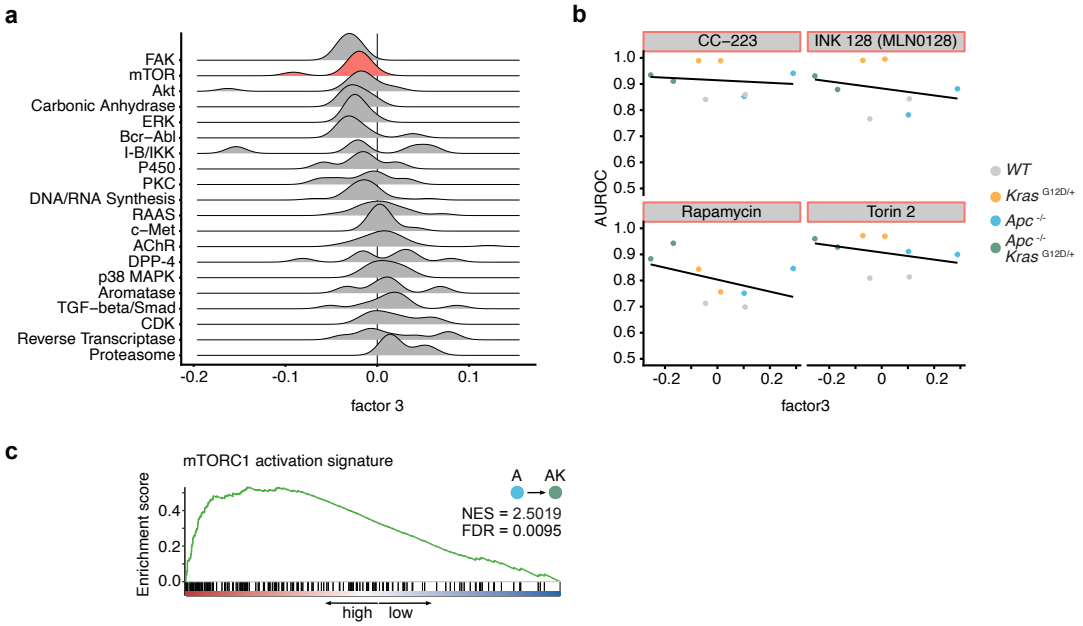

Figure 4.15: Factor 3, *Kras*^{G12D} effects in the context of *Apc* loss of function. **a** Distributions of treatment activity loadings grouped by drug target for factor 3. **b** Relationship of representative mTOR inhibitor activity with factor 3 score. **c** Gene set enrichment results of a Reactome mTORC1 activation signature (NES = normalized enrichment score).

Exploration of factor 3 loadings identified mTOR and FAK inhibitor activity to contribute to a negative factor score. Double-mutant organoid models (*Apc*$^{-/-}$ *Kras*$^{G12D+/-}$), which had a low factor 3 score, showed a greater treatment activity for these small molecule inhibitors than their single-mutant *Apc*$^{-/-}$ counterparts (Figure 4.15 a). This difference in treatment activity was observable for both ATP-competitive (e.g., INK-128, Sapanisertib) and non-ATP-competitive (e.g., Rapamycin) inhibitors (Figure 4.15 b). In line with the increased sensitivity to mTOR inhibitors, differential transcript expression analysis comparing single (Apc -/-, abbreviated A) and double mutant (*Apc*$^{-/-}$ *Kras*$^{G12D+/-}$, abbreviated AK) organoids identified a significant increase in mTORC1 activation according to a Reactome signature (Figure 4.15 c, FDR=0.0095, NES=2.5019). In summary, factor 3 was aligned with the effect of oncogenic *Kras* in *Apc* mutant colon organoids and was characterized by increased transcriptional activity and sensitivity to mTOR signaling.

4.5 Interpreting treatment-induced organoid morphologies with Multi-omics factor analysis to identify small molecules with factor-specific effects

To project observed organoid morphologies (1522 filtered treatments x 25 morphology features) into the previously learned MOFA representation, the pseudoinverse of the morphology factor-weight matrix was approximated, and factor scores were estimated (1522 treatments x 3 factors) (Figure 4.16 a and b). Projection of treated organoid morphologies recovered the previously identified distribution of untreated organoid genotypes (Figure 4.16 c and d). As expected, treatments which showed a strong dispersion away from the centers formed by their respective genotype showed a high AUROC score (Figure 4.16 e and f).

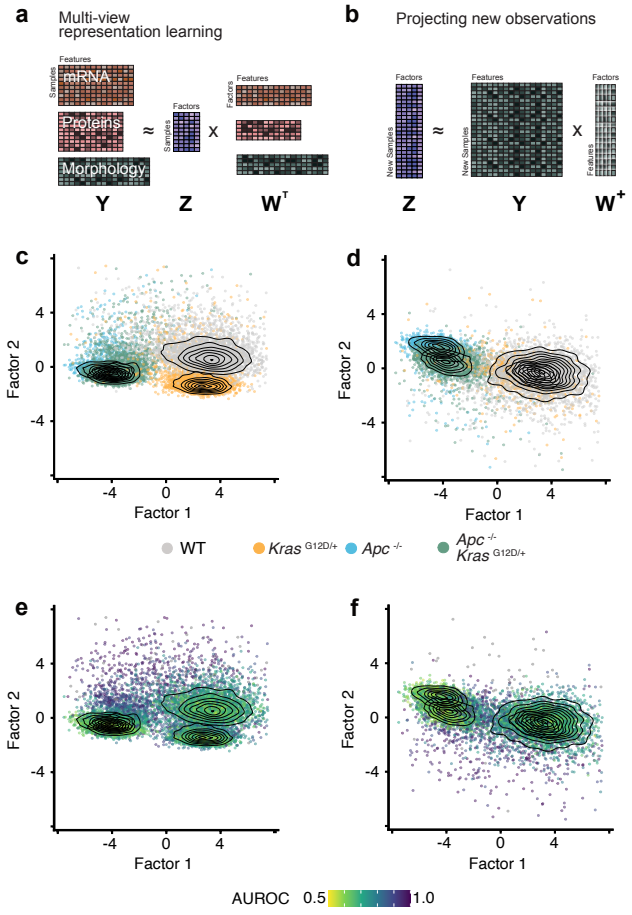

Figure 4.16: Visualization of treated samples projected into factor space. a and b Visual summary of multi-view matrix factorization and projection of new observations **c and d** Visualization of projected samples along factor 1, factor 2 and factor 3, respectively. Samples are colored by genotype. **e and f** Visualization of projected samples along factor 1, factor 2 and factor 3, respectively. Samples are colored by AUROC activity score.

4.5.1 Interpreting treatment-induced organoid morphologies using a learned multi-omics factor model

Next, I was interested in identifying non-toxic treatments that changed organoid morphology along one of the two primary factors (Factors 1 and 2) that were previously identified. Small molecules with the ability to move treated organoids from a transformed state towards a wild type-like state are of particular clinical interest (i.e., causing *Apc*$^{-/-}$ organoids to appear like wild type organoids). To this end, an ANOVA modeling the factor scores as a function of the treatment and organoid line (without an interaction term) was performed for each factor (Figure 4.17 a). Of the 1522 evaluated treatments, 66 (ca. 0.4 %) showed a significant effect on either factor 1 (6 treatments) or factor 2 (60 treatments), 56 (ca. 0.4 %) showed an effect on both factors, and 1400 showed no effect - thereby creating three groups: (1) treatments with a specific effect, (2) unspecific effects, and (3) no observed effect.

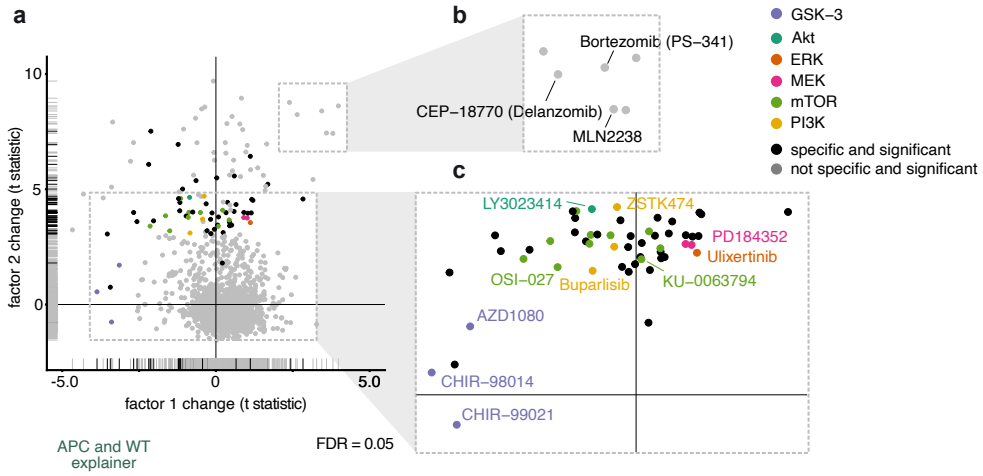

Figure 4.17: Treatment-induced changes on factor 1 and factor 2 scores. a Highlighted are treatments with a significant effect on projected scores for factor 1 and factor 2 across all organoid lines (ANOVA, FDR = 0.05). Treatments with a significant effect on only one of the three evaluated factors are highlighted in black. **b** Example for treatments with unspecific effects. Shown are Proteasome inhibitors. **c** Example for treatment with specific effects on one factor. Treatments are colored by their literature-derived mechanism of action.)

As treatments that cause low-viability phenotypes, which were not modeled during training of the MOFA model, can cause spurious and unspecific effects in the learned factor space, treatments with unspecific effects were inspected first. Among these treatments were multiple compounds which were tested at toxic concentrations, including Proteasome inhibitors (i.e., Bortezomib, Delanzomib, Carfilzomib; Figure 4.17 b), conventional DNA-intercalating agents (i.e., Epirubicin), and compounds with PAINS-like properties (Ellagic Acid).
Among factor-specific treatments, small molecules that displayed a negative effect on factor 1 (shifting from a state associated with an *Apc* wild type allele to an *Apc*$^{-/-}$ associated state) were enriched for GSK3 beta inhibitors (i.e., CHIR-98014, CHIR-99021, LY2090314; Figure 4.17 c). Along factor 2, small molecules targeting

downstream mediators of ERK-MAPK signaling (i.e., PD184352, Ulixertinib) and PI3K-AKT-mTOR signaling (i.e., Buparlisib, KU−0063794) had a positive effect (shifting from a state associated with *Kras*^{G12D} +/- genotype to a *Kras* wild type associated state). Unfortunately, no treatment with positive effect on factor 1, indicating a shift from an *Apc*$^{-/-}$ associated state towards an *Apc* wild type state, was identified.

4.5.2 GSK3 BETA INHIBITORS MOVE $Apc^{+/+}$ ORGANOIDS SPECIFICALLY ALONG FACTOR 1

GSK3 beta is a kinase with a central function within the canonical Wnt signaling destruction complex, (Stamos et al. 2013) (Figure 4.18 a) and inhibition of GSK3 beta has been shown to lead to hyperactivation of canonical Wnt signaling (Stambolic et al. 1996). Since Apc -/- organoids are already in a state of high canonical Wnt signaling, I hypothesized that GSK3 beta inhibition should lead to a larger observed effect on factor 1 in $Apc^{+/+}$ organoids than their $Apc^{-/-}$ counterparts, a case of treatment-genotype interaction.

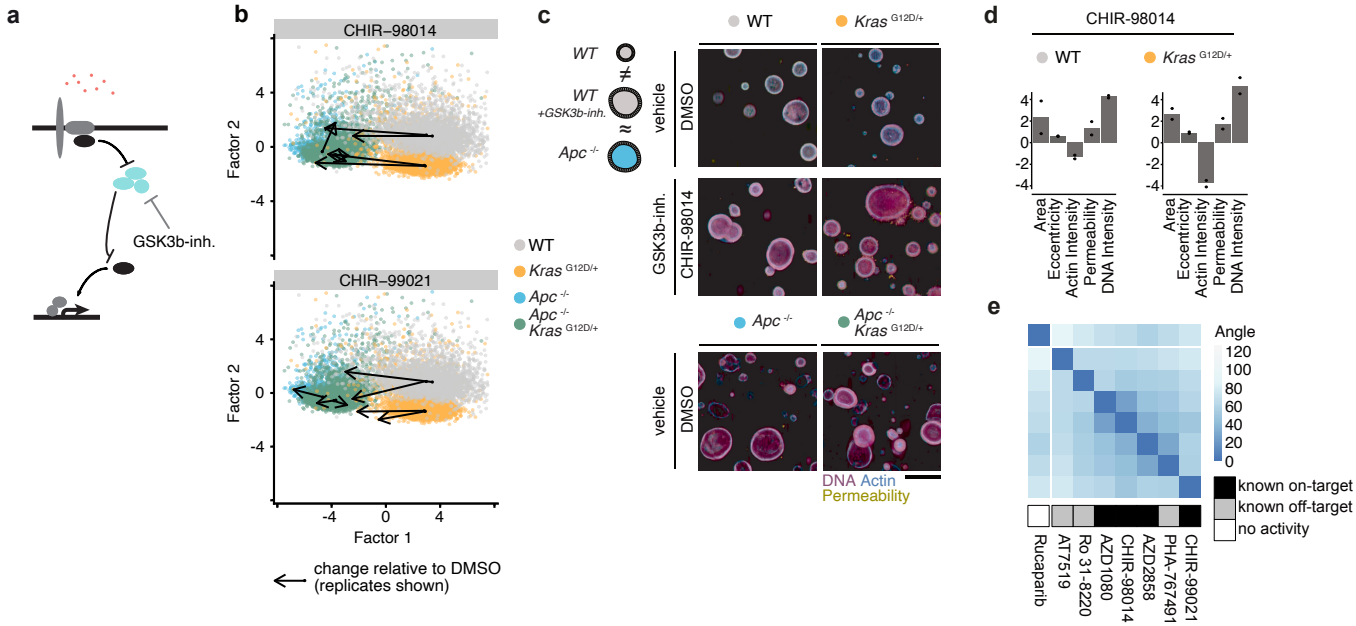

Figure 4.18: GSK3 beta inhibition dependent morphology in colon organoid models. a Wnt signaling pathway diagram. The effect of GSK3 beta inhibitors on the destruction complex is displayed. **b** Projected shift in factor scores after treatment of GSK3 beta inhibitors CHIR98014 and CHIR–99021. Shown are projected treatments for all four organoid genotypes. Drawn vectors correspond to the shift in factor scores attributable to the highlighted small molecule, and the DMSO control. Two vectors are shown for every organoid line, corresponding to the two screening replicates. For $Apc^{-/-}$ (blue) only one replicate was included in the analysis. **c** Small molecule inhibition of GSK3 beta (CHIR98014) leads to phenocopying of $Apc^{-/-}$ genotype organoid models. **d** Shift of morphological features of wild type and $Kras^{G12D}$ +/- organoid models treated with CHIR98014. Shown is an increase in organoid size (Area) and DNA intensity. **e** Excerpt of clustering from figure 4.9 d, labeled with known binding activity of listed small molecules. Rucaparib is not a member of the cluster and shown for comparison.

Indeed, the effect of the GSK3-beta inhibitors CHIR-98014 and CHIR-99021 on factor 1 was stronger for WT and $Kras^{G12D+/-}$ organoid lines (Figure 4.18 b). Treatment with the GSK3 beta inhibitor CHIR-98014 led to treatment-induced phenotypes in WT and $Kras^{G12D+/-}$ organoids that phenocopied the unperturbed morphology of $Apc^{-/-}$ and $Apc^{-/-}$ / $Kras^{G12D+/-}$ organoid models (Figure 4.18 c). On the individual morphology feature level, treatment which CHIR-98014 led to an increase in organoid size and DNA intensity (Figure 4.18 d) in $Apc^{+/+}$ models. This change in morphology was likely due to an increased proliferation rate of mutant cells, leading to rising organoid size and a higher density of nuclei per analyzed object. Guided by the identification of an interpretable GSK3 beta inhibition induced phenotype, I analyzed small molecules that clustered with known inhibitors of this kinase based on the similarity of their drug effect vectors (Figure 4.18 e). In line with the observations in this experiment, all small molecules clustering with well-described inhibitors of GSK3 beta had

previously described off-target binding activity against this kinase within the LINCS KINOMEScan database (Subramanian et al. 2017).

In summary, projecting treatment-induced morphologies into a factor space that was previously learnt from annotated organoid samples helped interpret small molecule effects and recovered previously known annotated on-target and off-target activity against GSK3-beta. In the future, I expect such an approach to assist in the therapeutic discovery of small molecules targeted at early functional genetic events during colorectal cancer pathogenesis.

5

Discussion

5.1 Disclosure

Parts of the discussion, in particular the section *"Two multi-omics factors defining patient-derived colorectal cancer organoid phenotype"*, have been adapted from a joint first-author publication, *The drug-induced phenotypic landscape of colorectal cancer organoids* (Betge et al. 2022). The adapted text was originally written by myself.

5.2 Image-based profiling of organoids

5.2.1 Using organoids as high validity *in-vitro* models for image-based profiling

Organoids are representative *in-vitro* models for diverse human tissues and can be used for image-based profiling. While the prospective use of cancer organoids as a diagnostic is currently limited by high sample dropout and long turnaround time, a high overall predictive validity for multiple therapeutic regimens has been reported (S. N. Ooft et al. 2021). The use of organoid models in early-stage drug discovery, however, is not limited by the same constraints existing in a diagnostic context. For therapeutic discovery, previous studies have successfully used organoids to perform medium-scale small molecule treatment assays. The most commonly used method is

screening with an ATP-based cell viability readouts (Wetering et al. 2015). Additionally, imaging studies with organoids have been used to characterize developmental processes such as the self-organization of intestinal cells (Lukonin et al. 2020; Boehnke et al. 2016) or the morphological response to individual drugs (Badder et al. 2020; Serra et al. 2019). While image-based profiling of *in-vitro* models has become an important method in functional genomics and therapeutic discovery (Carpenter 2007), performing such high-throughput experiments in organoid models has been a technological challenge given the 3D multicellular growth pattern of organoids as well as their required culture conditions. In this thesis, sparse 3D imaging as well as adjustment to the seeding and staining protocol enabled small molecule image-based profiling of patient-derived and genetically engineered colorectal cancer organoid models.

5.2.2 Towards multi-view profiling

A challenge in image-based profiling is the low interpretability of the morphological representation that is generated during the image analysis step. Classic quantitative features of cellular morphology are rarely descriptive, and even when dimensionality reduction methods or self-supervised learning techniques are used, the learned representations are rarely directly interpretable within the context of cellular mechanisms. To increase the interpretability of observations in a given image-based profiling experiment, supporting multi-view data, for example transcriptomics, proteomics or metabolomics data can be collected to ”annotate” images. A simple and idealized approach to perform such annotation would be to collect the complete multi-view information for all tested treatment conditions (i.e., imaging, transcript abundance, protein abundance, lipid abundance). However, both from a physical (some measurement methods destroy the sample) and cost perspective, this approach is inaccessible.

In this thesis, an alternative model-based approach was chosen to increase the interpretability of morphology information: First, a well-annotated multi-view representation of untreated organoids was learned by modeling morphological features as well as molecular features (i.e., transcript expression). In a second step, single-view treatment-induced morphologies were projected into the learned representation. The resulting factor scores were used to identify treatments with an interpretable effect. Examples of such identified effects include (1) the increase in organoid size and IGF1R-signaling as a response to mTOR inhibition (factor 1, patient-derived organoids); (2) the spherical reorganization and increase in LGR5 transcript expression as a response to MEK inhibition (factor 2, patient-derived organoids); as well as (3) the increase in organoid size, DNA staining intensity and canonical Wnt signaling as a response to GSK3-beta inhibition (factor 1, mouse organoids). While this approach has utility in guiding the interpretation of otherwise challenging-to-interpret morphology information, a series of important limitations exist that are outlined below.

A central limitation of the described multi-view representation learning approach are **out of distribution observations** during the projection of treated organoids. Observations during factor-learning were drawn from an

untreated distribution of organoid states, while projected observations were drawn from a distinct distribution of treated organoids. As a result, an observed treatment within the query set that is far outside the distribution learned from the support set is likely to be misinterpreted. For example, if observation from *Apc* mutant organoid lines had not been included in the support set that the multi-view representation was learned from, the effect of GSK_3-beta inhibitors on *Apc* wild type organoids would not have been interpretable. These treatment-induced morphologies would have been grouped with the set of non-specific but active treatments, such as Proteasome inhibitors. In this thesis, potential out-of-distribution treatments were identified through a high AUROC score and non-specific treatment effects along multiple factors. In the future, more principled methods to detect out-of-distribution samples, such as obtaining the likelihood or reconstruction error, could be used. Given the limitation of out of distribution observations, it is important to include conditions in the support set that are well annotated **and** representative of the effects that are supposed to be explored, as the representation learned from this data will be used to interpret all remaining observations within the query set. Put differently, a support set needs to be constructed with the biological question in mind because all other observations will be interpreted using a model that was learned from it.

A second limitation associated with the used approach is a **lack of causality** in the interpretation of treatment effects on factors. For example, CDK inhibitors led to a decrease in factor 1 scores among patient-derived colorectal cancer organoids, with a markedly smaller organoid size in treated organoids. This decrease in factor 1 scores does, however, not inform about the causal structure of the mechanism that is leading to a shift in factor 1. For example, while all the following causal diagrams are as likely in the light of the factor score shift, they are distinct in terms of their biological plausibility: It is unlikely that CDK is a regulator of IGF1R signaling (CDK $\rightarrow$ IGF1R $\rightarrow$ organoid size), while it is more likely that CDK is a mediating variable of IGF1R signaling (IGF1R $\rightarrow$ CDK $\rightarrow$ organoid size) or an independent effector (IGF1R $\rightarrow$ organoid size, CDK $\rightarrow$ organoid size). In summary, treatment effects on factors aid in the generation of hypotheses but need to be used as a starting point for further validating experiments.

A third limitation of this approach is **data drift**. This limitation is most likely relevant in the context of larger profiling campaigns: Over the course of multiple experimental batches, the distribution of morphology states that is observed for both untreated and treated organoids is shifting. A model that has been trained at day 1 on a distribution of untreated organoid phenotypes and is used to interpret phenotypes on day 30 of treated organoids might show unexpected behavior because of a shift of experimental conditions, i.e., the staining protocol. The conditions in the experiments performed during this thesis were not fully representative of such scenarios, as untreated and treated organoid morphologies were collected in parallel and all morphological data were pre-processed together (i.e., features scaling, PCA transformation). As a result, the ability of multi-view group factor analysis to project ”unseen” treatment induced morphologies into interpretable factors was assessed under idealized conditions with considerable data leakage, which needs to be accounted for if using such an approach in larger discovery or diagnostic projects.

Despite these limitations, the combination of multi-view representation learning and image-based profiling of organoids introduced in this thesis has the potential to be built upon and improved. A set of possible directions exist, such as (1) further increasing the predictive validity of in-vitro models by imaging organoid models within co-culture systems (Cattaneo et al. 2020); (2) further increasing the robustness of multi-view profiling by introducing selective RNA-Sequencing of formalin fixed cells after a microscopy run is completed; and (3) improving the image analysis and modeling process by using self-supervised learning methods instead of classical morphological features (Perakis et al. 2021).

5.3 Multi-view profiling of patient-derived colorectal cancer organoids identifies factors of cancer organoid architecture

A set of 11 patient-derived colorectal cancer organoids were profiled to understand the underlying factors determining organoid morphology, as well as characteristic differences in small molecule sensitivity.

5.3.1 Two multi-omics factors defining patient-derived colorectal cancer organoid phenotype

The first identified axis of phenotype variation was an IGF1R signaling program associated with increased organoid size, EGFR inhibitor resistance, which can be induced by mTOR inhibition. Insulin-like growth factors are central and conserved regulators promoting cell size, organ size and organism growth (Puche et al. 2012; Sun et al. 2006). The IGF1 receptor (IGF1R) signaling cascade is activated in around 20% of colorectal cancer patients and leads to downstream mitotic stimuli via ERK-MAPK signaling and mTOR (Zhong et al. 2017). In patient-derived cancer organoids, we observed that average organoid size was positively correlated with elevated IGF1R signaling activity, as determined by transcript expression.

Increased IGF1R signaling activity also presented with a characteristic pharmacological sensitivity profile. In accordance with previous observations(Isella et al. 2017a), colorectal cancer organoids in a high IGF1R signaling state were less responsive to EGFR inhibitors and more responsive to IGF1R and MEK blockade, demonstrating the central role of IGF1R mediated mitogen activated protein kinase (IGFR1-MAPK) signaling. In fact, reciprocal resistance between IGF1R and EGFR signaling inhibitors has been described in multiple cancer types (Hua et al. 2020). Moreover, organoids could be moved into a state of increased IGF1R-MAPK signaling by inhibition of mTOR, a downstream mediator of IGF1R activity. In line with this observation, a reactive induction of IGF1R signaling has been previously described as a resistance mechanism to small molecule mTOR inhibitors in cancer (S. V. Sharma et al. 2010; Yoon et al. 2017).

Its role in controlling cell size and the high number of interactions with other signaling mechanisms, such as EGFR and mTOR, have made IGF1R signaling an attractive target for therapeutic discovery. So far, however, neither mono- nor combination therapies containing IGF1R receptor inhibitors have shown clinical utility (Beckwith et al. 2015; Jentzsch et al. 2023). A speculative explanation for this failure in clinical trials might be related to its signaling network centrality. While the IGF1R receptor and its downstream effectors are central mediators of growth-stimulating processes, they do not themselves constitute a necessary condition for tumor growth (Beckwith et al. 2015), which is, in contrast, often the case for genetic events involved in disease development, such as *APC* loss-of-function (Dow et al. 2015) or *KRAS* gain-of-function.

While the potential of IGF1R as a target for therapeutic discovery is low in light of past evidence, direct implications for the method of organoid culture - in particular the role of supplementing IGF-1 to increase organoid growth *in-vitro* - might be drawn from the results of this study. Next to the observation that organoids with strong IGF1R signaling were of greater average size, the emerging role of IGF1R signaling in organoid

culture was recently emphasized by the observation that addition of the IGF-1 ligand, relative to EGF, increased culture efficiency of organoids from healthy human intestinal tissue (Fujii, Matano, et al. 2018). Given the association of IGF1R signaling with organoid size within this thesis as well as the observation from organoid isolation studies (Fujii, Matano, et al. 2018), I hypothesize that addition of IGF-1 ligand to colorectal cancer organoid culture media could (1) further increase culture establishment efficiency and (2) reduce additional genetic bottleneck effects that might bias isolated organoid cultures towards increased IGF1R signaling.

The second axis of phenotype variation was an LGR5+ program associated with cystic organoid architecture and Wnt signaling inhibitor sensitivity, which can be induced by inhibition of MEK. Organoid models with a high factor score showed a monolayer organization with a characteristic actin cytoskeleton and an enrichment of an LGR5 intestinal stem cell signature along the factor weights. These LGR5+ organoids were more sensitive to Wnt inhibitor (Pri-724) treatment and showed a relative resistance to Erk and Mek small molecule inhibitors. Surprisingly, not only were organoids in this stem-like state more morphologically resistant to Mek inhibition, treatment with Mek inhibitors shifted organoids towards this LGR5+ state. Put differently, a subset of organoids which are LGR5+ are insensitive to MEK inhibition and Mek inhibition drives (surviving) organoids towards this LGR5+ state.

The ability of ERK-MAPK inhibition (here EGFR and MEK) to move intestinal organoids into a quiescent LGR5+ state was described by (Basak et al. 2017) and led to a ”clutch and gas pedal” model of Wnt and ERK-MAPK signaling in which Wnt signaling is necessary and sufficient to maintain the LGR5+ stemlike state, while ERK-MAPK signaling is a primary signal to cause (stem cell) proliferation (Basak et al. 2017). Additional work by (Tianzuo Zhan et al. 2019) recently extended this model by demonstrating that inhibition of MEK in colorectal cancer cell lines and organoid models leads to an increase in canonical Wnt signaling through inhibition of EGR1-dependent transcription of AXIN1, a necessary member of the destruction complex. According to this extended model, Wnt signaling is necessary and sufficient to maintain the LGR5+ cell state, and ERK-MAPK signaling is not only the primary proliferative signal, but its inhibition can even further shift the cellular state towards the LGR5+ identity.

5.3.2 Are patient-derived organoids representative of known molecular colorectal cancer states?

The multi-view factors that were identified in patient-derived organoids during this thesis capture differences in IGF1R-, EGFR-, and canonical Wnt signaling (high Factors 1, low Factor 1, and high Factor 2, respectively). A natural question to ask is whether the identified molecular axes of variation are aligned with previously identified molecular axes or variation (and the states they define) of colorectal cancer tissue samples. Molecular states, also referred to as subtypes, of colorectal cancer have been proposed by a series of consortia and are often centered around a single data modality. Most common subtype classifications are based on bulk transcript expression data (Guinney et al. 2015) while other classifications, for example based on single-cell transcript expression (Joanito et al. 2022), have been developed. The Consensus Molecular Subtypes (CMS) classification (Guinney et al. 2015) is among the most widely used subtype classifications and was developed by a consortium of researchers that had previously published multiple independent classifications. The CMS classification groups colorectal cancer tissue samples that have undergone bulk transcript expression measurements into four subtypes that are summarized briefly below:

- CMS 1: hypermutated, frequently BRAF-mutated, MSI and CIMP samples with corresponding signs of inflammation
- CMS 2: high Wnt signaling activity, chromosomal instability (CIN) common
- CMS 3: *KRAS*-mutated MSS samples, signs of strong metabolic deregulation
- CMS 4: stromal infiltration and high TGF-beta activation, worse overall survival

Since its publication (Guinney et al. 2015), however, the classification has not found use in clinical care. Two prospective studies in which the CMS classifier was applied and the treatment choices of Cetuximab vs. Bevacizumab (both in combination with FOLFIRI/FOLFOX) were evaluated, came to -at first sight- conflicting conclusions (Stintzing et al. 2019; Lenz et al. 2019; Aderka et al. 2019; Sveen et al. 2019). While one group concluded that the CMS classification had ”no direct impact on clinical decision-making” (Stintzing et al. 2019) the other group concluded that their ”findings highlight the possible clinical utility of CMS” (Lenz et al. 2019). While a range of specific explanations for this particular discordance have been proposed (i.e., cohort differences in UICC stage, pretreatment history, and chosen chemotherapy backbone), a set of general observations that highlight current limitations of clinical transcriptome measurement and classification that go beyond the use of the CMS classifier itself might be more instructive for this discussion:

- Clinical transcriptome measurements and classification of multiple biopsies from the same patient can give discordant results, for example when comparing primary tumor and metastasis (Eide et al. 2021)
- Clinical transcriptome measurements and classification are still not a standardized procedure, thereby reducing their reliability (Sveen et al. 2019). Different transcript expression measurement methods and

software libraries are used —even between clinical trials that aim to use the same classification (Stintzing et al. 2019; Lenz et al. 2019).

- Clinical transcriptome measurements capture and classify the whole tumor sample, including the tumor microenvironment surrounding the malignant cells. In such contexts, it is generally challenging to draw conclusions for the malignant cells in isolation.

Especially because the CMS classification was established on whole tissue samples, and not isolated malignant cells, care should be taken when comparing subtype classifications between in-vitro models and clinical samples. Within the collection of patient-derived organoids profiled within this study, the majority of organoids were classified as CMS2 (11/13). Of note, no CMS1 or CMS4 organoid lines were established. This distribution of assigned subtypes was in line with reports from other colorectal cancer organoid studies (Wetering et al. 2015; Schütte et al. 2017) and showed an under-representation of CMS1 and CMS4. Motivated to identify a possible alignment with published classification systems, I noticed a partial alignment of modeled states with the colorectal cancer intrinsic subtype (CRIS) model (Isella et al. 2017b). The CRIS model was identified based on transcriptomic data from patient-derived xenografts (PDX), an alternative method to propagate patient-derived cancer cells *in-vitro* with high efficiency. By removing genes that are not expressed in epithelial cells, the classification suppresses signal from the micro-environment and thereby reduces previously described instabilities when applying it to heterogeneous clinical samples (Dunne et al. 2017).
Abbreviated, the five states comprising the CRIS model are:

- CRIS A: BRAF-mutated MSI, and KRAS-mutated MSS samples
- CRIS B: TGF-beta pathway activation, enriched for MSI samples
- CRIS C: EGFR activation, and sensitive to inhibition
- CRIS D: IGF signaling activation and EGFR inhibitor resistance
- CRIS E: high Wnt signaling activity with a ”Paneth cell-like” transcriptome signature

A subset of CRIS classes was identified within the presented organoid profiling data. The gene sets defining CRIS C and D were found to be separated along factor 1, with low factor 1 scores ($F1-$; $F2-$) associated with CRIS C and high factor 1 scores ($F1+$; $F2-$) associated with CRIS D. There was no significant enrichment identified for the other states outlined within the CRIS model, however, a potential association between CRIS E and $F2+$ might motivate further investigation.

While differences between tissue sample classification and organoid models are to be expected, the complete absence of CMS1, CRIS A or CRIS B organoid models is worthy of further discussion. All three subtypes show an enrichment of MSI positive colorectal cancer samples (Guinney et al. 2015; Isella et al. 2017b). Throughout this project, no MSI positive organoids were isolated successfully. Both MSI positive colorectal cancer cell lines

(Imkeller et al. 2022) and PDX models (Isella et al. 2017b) have been reported to show lower levels of intrinsic canonical Wnt signaling and an enrichment for RNF43 mutations, a Wnt signaling related functional event that -in contrast to truncating mutations of *APC*- does not render cells independent of Wnt ligands (Wetering et al. 2015). At this point, it is important to state that the quality of Wnt ligands in organoid medium is a common source of poor colon organoid culture efficiency (Kim Boonekamp, direct correspondence). Based on this observation, it is conceivable that the organoid methodology applied at the time of this study exerted a systematic bias against the development of MSI positive colon cancer organoids -possibly due to methodological challenges related to maintaining organoid culture conditions with sufficient Wnt ligand activity. Further standardizing and optimizing reagents related to colon cancer organoid culture will accelerate the development of fully representative *in-vitro* model collections for translational colorectal cancer research.

In summary, image-based profiling of colon cancer organoids revealed characteristic differences in IGF1R-, EGFR-, and canonical Wnt signaling (high Factors 1, low Factor 1, and high Factor 2, respectively) and their corresponding organoid morphologies. The resulting identified organoid states showed a partial agreement with previously described molecular subtypes of colorectal cancer. At the same time, imbalances in the prevalence of certain subtypes of both CMS and CRIS classifications among organoid models reveals directions for further methodological improvements: both in the development of robust colorectal cancer state descriptions and in the development of low-bias organoid culture protocols.

5.4 Multi-view profiling of step-wise *in-vitro* models of colorectal cancer pathogenesis

After multi-view factors which partially determined organoid morphology in patient-derived organoids were identified, the question arose to what extent such a factorization could help understand the multi-omics changes as well as treatment sensitivity differences among engineered models of early colon cancer pathogenesis.

5.4.1 Using mouse colon organoid models to study early colorectal cancer pathogenesis

For this project, murine colon organoids were chosen as a model. Murine tissue is a popular source of organoids (Sato, Vries, et al. 2009) and murine organoids show a generally more robust growth *in-vitro* which is an important factor to enable profiling experiments with ca. 1700 small molecule treatment conditions. An additional factor for the use of murine organoids was the availability of existing genetic mouse models, such as the *LSL-Kras^{G12D} CreERT2* conditional allele used in this project (Jackson et al. 2001). While the colon, in contrast to the small intestine, is the region in which carcinomas are more frequent in humans, the inverse is true for popular genetic mouse models (Luo et al. 2009). Despite the mouse small intestine being more frequently transformed in genetic models of colorectal cancer, and the ease of mouse small intestinal organoid culture, colon organoid models were chosen as a model. The primary reason for this choice was to more closely represent human colon tissue.

While enabling the observation of a series of known and also not yet directly reported molecular changes (i.e., the accumulation of storage lipids, outlined below), the engineered colon organoid models and the conditions under which they were interrogated had a set of limitations. One limitation exists regarding the comparison of *Apc$^{-/-}$* and double-mutant *Apc$^{-/-}$* / *Kras$^{G12D/+}$* organoid models. Comparing the two genotypes showed a high degree of molecular similarity. While this high similarity was observed consistently across different modalities, it was potentially pronounced given the presence of Egf ligands in the organoid culture medium. In the presence of a high concentration of Egf ligands, the effect of *Kras$^{G12D/+}$* in a cell with high *Apc$^{-/-}$*-dependent stemness may be less pronounced compared to *in-vivo* models in which tumors form outside of high Egf conditions.

5.4.2 Effects of *Apc* truncation and *Kras*$^{G12D/+}$ on colon organoid models

Multi-view profiling of organoids with an *Apc* loss of function recovered a set of known molecular mechanisms of adenoma formation. Introduction of insertions/ deletions into the *Apc* gene via CRISPR led to increase in proliferation rate in standard medium, canonical Wnt signaling and growth in Wnt-ligand free medium. The hyper-activation of canonical Wnt signaling also manifested itself in the characteristic sensitivity differences to inhibitors of Wnt secretion and downstream transcription, as well as in the phenocopying of the *Apc*$^{-/-}$ related morphology (factor 1) by inhibition of the destruction complex member GSK3-beta with small molecules.

Additional *Apc* loss of function-specific effects included an increased sensitivity towards FAK-inhibitors as well as microtubule-interfering compounds, such as taxanes. While multiple explanations for the mechanism leading to these vulnerabilities exist (Ashton et al. 2010), they point towards additional roles of Apc outside the destruction complex, involving the regulation of focal adhesion- (Ashton et al. 2010; Matsumoto et al. 2010) as well as microtubule-dynamics (Stolz et al. 2015) within the cell.

Analogous profiling of organoids with an isolated *Kras*$^{G12D/+}$ allele recovered characteristic small molecule sensitivity differences. Organoids with a *Kras*$^{G12D/+}$ genotype were more sensitive towards ERK-MAPK cascade mediators downstream of *Kras* (i.e., ERK and MEK subfamily inhibitors) and more resistant to inhibitors of upstream receptors (i.e., EGFR inhibitors). This resistance to EGFR inhibitors coincided with a down regulation of the Egfr receptor, as suggested by transcript expression profiling. Previously, oncogene induced senescence has been described in colon epithelial cells with isolated *Kras*$^{G12D/+}$ (Bennecke et al. 2010) *in-vivo*. In line with these observations, an oncogene-induced senescence signature was identified within factor 2.

New observations were made related to the metabolic effects of *Apc*$^{-/-}$ and *Kras*$^{G12D/+}$ in colon organoids. Both genotypes led to a reduced expression of beta-oxidation associated transcripts and an accumulation of storage lipids, such as triacylglycerides (TAG). Of note, *Apc*$^{-/-}$ organoids also showed an accumulation of cholesterol esters (CE), while this lipid species was depleted relative to the wild type state in *Kras*$^{G12D/+}$ organoids. Both TAGs and CEs are components of lipid droplets, phospholipid-monolayer enclosed droplets of neutral lipids within the cytoplasm that have been observed in various cancer types, including colorectal cancer (Cruz, Barreto, et al. 2020; Accioly et al. 2008). While the function of lipid droplets is not clear, their anabolic accumulation is likely mediated via mTOR and Sterol regulatory element binding proteins (SREBP) (Cruz, Barreto, et al. 2020) and has been observed as the result of ERK-MAPK activation via oncogenic *Hras*^{V12} in mouse embryonic fibroblasts (Cruz, Carrossini, et al. 2019) as well as in primary colon cancer cells with high canonical Wnt signaling activity (Tirinato et al. 2015).

Besides pronounced changed in lipid metabolism, *Kras*$^{G12D/+}$ engineered organoid models also showed changes in their cellular respiration. Introduction of a *Kras*$^{G12D/+}$ genotype led to significant reduction in oxidative phosphorylation based on transcript abundance data, in line with the expected Warburg effect (DeBerardinis et al. 2020). In contrast, *Apc*$^{-/-}$ organoids did not present themselves with such a change in their transcript abundance,

in line with recent work by Imkeller et al. in which an analysis across in-vitro colorectal cancer models identified a higher rate of oxidative phosphorylation and mitochondrial dependency in *Apc*$^{-/-}$ / high Wnt-signaling models compared to *Apc*$^{+/+}$ / low Wnt-signaling models (Imkeller et al. 2022).

In summary, engineered colon organoid models were able to recapitulate multiple previously described molecular changes associated with colorectal cancer pathogenesis *in-vivo*. In addition, the accumulation of lipid droplets in both *Apc*$^{-/-}$ and *Kras*$^{G12D/+}$ models was identified, pointing towards a potential additional mechanism through which mTOR signaling, besides global translation rate control (Smit et al. 2020), is a point of interaction between these two common and correlated genetic events during colorectal cancer formation.

5.5 Perspective

In-vitro models, such as organoids, are an essential component of the therapeutic discovery process. Therapeutic discovery is a sequential decision-making process in which, for example, a candidate small molecule can not, for ethical and practical reasons, be tested directly in a clinical context to observe its treatment effect. Instead, therapeutic candidates have to be evaluated using one or more models to approximate the treatment effect that could be observed in large clinical trials. In practice, multiple *in-silico*, *in-vitro* and *in-vivo* models are combined and used sequentially to guide cost-effective decision-making during a discovery program. The higher the predictive validity and the lower the cost of using a model, the more value does it provide to the therapeutic discovery process. While determining the validity of an *in-vitro* model is an empirical process, a set of axioms to prioritize models have been formulated by Vincent et al. and are presented below in a modified format (Vincent et al. 2015):

1. The model must have a clear link to the disease of interest (i.e., matched tissue of origin, representative culture conditions) and, if engineered, model the disease state based on the best understanding of the disease pathophysiology.

2. The treatment should only represent the planned clinical intervention. Additional stimuli, such as cytokines or damage-inducing chemicals, that are required to model the disease associated phenotype should be avoided.

3. The phenotype observed during the experiment should represent a coarse-grained function (cytotoxicity, muscle contraction) that is related to the desired clinical endpoint, rather than a molecular biomarker (transcript expression).

Designing or ”fitting” *in-vitro* biological models to a disease along these axioms might be a cost intensive process, but already minor changes in predictive validity of an experiment can offset these costs in a discovery program. In fact, economic research into the overall cost of therapeutic discovery by Scannel et al. (Scannell et al. 2016) concludes with the statement:

> ”The rate of creation of valid screening and disease models may be the major constraint on R&D productivity.”

Hopefully, the work presented in this thesis on early- and late-stage organoid models of colorectal cancer can contribute to further increasing the predictive validity of image-based therapeutic discovery programs targeting this disease.

Summary

Colorectal cancer belongs to the most common and lethal forms of cancer. Organoids are multicellular, three-dimensional in-vitro models which enable the culture of both healthy and transformed adult tissue, including the culture of healthy and cancerous colon epithelium. In culture, organoids show highly heterogeneous morphologies. The molecular factors underlying these differences in organoid architecture are, however, not well understood.

The aims of this thesis were (1) to understand the factors determining the diversity of organoid morphologies, (2) to identify differences in small molecule treatment sensitivity linked to these factors, and (3) to explore whether non-toxic small molecules existed that could move organoids from one morphology to another morphology along the previously identified factors.

First, organoid models were generated that represented the colorectal carcinoma and adenoma stages. For the former, patient-derived organoids were derived from endoscopic biopsies of colorectal carcinoma. For the latter, organoids were isolated from a genetic *LSL-Kras*^{G12D} *CreERT2* mouse model and, via CRISPR/Cas9, *Apc* insertion/deletions were introduced which yielded four organoid models with distinct genotypes.

Organoid models were then subjected to image-based profiling, a microscopy-based high-throughput assay, and more than 2,000 small molecule treatment conditions were evaluated in total. The resulting morphological information was analyzed jointly with multi-omics data using group factor analysis. The identified factors captured variation of molecular, morphological, and small molecule treatment sensitivity differences.

Two primary factors of variation were identified in patient-derived colorectal cancer organoids: The first factor organized organoids according to their size and the activity of the IGF1R signaling pathway. Organoids with a large diameter and high activity of IGF1R signaling were sensitive towards small molecule inhibitors of the IGF1 receptor and resistant towards EGFR inhibition.

The second factor organized organoids according to their Lgr5+ stemness. Organoids with a high Lgr5+ stemness signature showed a spherical and symmetric multicellular organization. These organoids showed a sensitivity towards inhibitors of Wnt signaling and a relative resistance towards inhibitors of ERK-MAPK signaling.

Using these identified factors, known, non-lethal, small molecule treatment effects were identifiable: For example, small molecule inhibitors of mTOR showed a treatment-induced morphology that corresponded to a positive shift along factor 1. In fact, treatment of organoids with mTOR inhibitors during validating experiments led to a reactive overexpression of the IGF1R signaling effector Irs-1. A related effect was observed in MEK inhibitors regarding factor 2. Treatment of organoids with MEK inhibitors led to an increased expression of *LGR5* transcripts.

A similar approach was used in genetically engineered mouse organoids. Here, three prominent factors were identified: The first factor separated *Apc* wild type from *Apc* mutated organoid models. Loss of *Apc* led to an increase in transcript expression of proliferation- and DNA-repair markers. These organoids also showed an increased sensitivity towards perturbation of the cytoskeleton, for example through taxanes or FAK inhibitors, as well as a loss of the spherical and symmetrical multicellular structure. Organoids with an $Apc^{-/-}$ genotype grew independently of the otherwise essential Wnt ligands in the culture medium. Regarding changes in the cellular metabolism, loss of *Apc* led to a reduced expression of beta-oxidation-related transcripts and to an accumulation of triglycerides and cholesterol-ester storage lipids.

The second factor separated $Kras^{G12D/+}$ organoid models from wild type organoids. Organoids with $Kras^{G12D/+}$ genotype showed signs of oncogene-induced senescence, reduced expression of beta-oxidation related transcripts, an accumulation of triglycerides as well as a marked reduction in the expression of oxidative phosphorylation related transcripts. Organoids showed an increased sensitivity towards ERK- and MEK-inhibitors and a low relative sensitivity towards EGFR inhibitors.

The third factor separated $Apc^{-/-}$ from $Apc^{-/-}$ / $Kras^{G12D/+}$ organoids. Double mutant organoid models showed an increased sensitivity towards mTOR inhibitors and an increased expression of mTORC1-activation related transcripts. As demonstrated previously, learned factors were useful to identify non-lethal small molecule treatment effects that moved organoids along the identified factors. For example, inhibitors of GSK3-beta, which are known activators of canonical Wnt signaling, had a marked effect on factor 1.

The results presented within this thesis elucidate primary factors of morphological and molecular heterogeneity in organoid models of colorectal cancer and its precursor states. The results also demonstrate the potential of image-based profiling of organoids within translational medical research.

Zusammenfassung

Das kolorektale Karzinom gehört zu den häufigsten und tödlichsten Krebserkrankungen. Organoide sind multizelluläre, dreidimensionale *in-vitro* Modelle, welche die effiziente Kultur von gesundem und malignen adulten Gewebe ermöglichen, unter anderem auch die Kultur von Dickdarmepithel und kolorektalen Karzinomen. *In-vitro* zeigen Organoide eine ausgeprägte morphologische Heterogenität. Die molekularen Faktoren, welche diese Unterschiede in der Organoidarchitektur bedingen, sind jedoch nicht gut verstanden.

Das Ziel dieser Promotion war es (1) molekulare Faktoren, welche die Organoidarchitektur determinieren, zu identifizieren, (2) Faktor-abhängige Unterschiede in der Medikamentensensibilität zu beschreiben, und (3) Medikamente zu charakterisieren, welche die Organoidarchitektur entlang der identifizierten Faktoren verändern können.

Zuerst wurden Organoid Modelle angelegt, welche das kolorektale Karzinom-Stadium sowie die frühe Pathogenese im Adenom-Stadium repräsentieren. Für Modelle des Karzinom-Stadiums wurden endoskopisch isolierte Biopsien von kolorektalen Karzinomen als Grundlage genutzt. Für Modelle des Adenom-Stadiums wurden Organoide aus dem Dickdarmepithel einer *LSL-Kras*^{G12D} *CreERT2* transgenen Maus gewonnen und mittels CRISPR *Apc* Insertion/Deletions-Mutationen eingefügt.

Daraufhin wurden Organoide mittels bildbasierter Hochdurchsatzmedikamententestung untersucht und insgesamt 2000 Behandlungsbedingungen getestet. Die gesammelten morphologischen Informationen wurden daraufhin zusammen mit unterstützenden multi-omics Daten mittels einer Gruppen-Faktor-Analyse analysiert und Faktoren, welche molekulare und morphologische Unterschiede als auch Unterschiede in der Medikamentensensitivität erklärten, identifiziert.

In patientenstämmigen kolorektalen Organoiden wurden zwei prominente Faktoren identifiziert: Der erste Faktor organisierte Organoide nach ihrer Größe und der Aktivität des IGF-Signalweges. Organoide mit einem größeren Diameter und hoher Aktivität des IGF-Signalweges waren sensitiv gegenüber der Hemmung des IGF1 Rezeptors und resistent gegenüber EGFR-Inhibitoren.

Der zweite Faktor organisierte Organoide nach der Ausprägung der LGR5+ Stammzellen-Identität. Organoide mit einer ausgeprägten LGR5+ Stammzellen-Identität zeichneten sich durch eine sphärische und symmetrische multizelluläre Organisation aus. Diese Organoide zeigten eine hohe relative Sensitivität gegenüber Inhibitoren des Wnt Signalweges und eine relative Resistenz gegenüber Inhibitoren der ERK-MAPK Signalweges.

Mithilfe der gelernten Faktoren konnten auch bekannte nicht letale Medikamenteneffekte systematisch identifiziert werden: Inhibitoren des mTOR-Komplexes zeigten etwa in den Experimenten eine behandlungsinduzierte Morphologie, welcher auf einen positiven Effekt auf Faktor 1 hindeutete. Tatsächlich führte Behandlung von Organoiden mit einem mTOR-Inhibitor in validierenden Experimenten zu einer reaktiven Expression des IGF-Signalwegkomponenten Irs-1. Ein ähnlicher Effekt zeigte sich bei MEK Inhibitoren in Bezug auf Faktor 2. In validierenden Experimenten führte die Behandlung zu einer erhöhten Expression von *LGR5* Transkripten.

In genetisch modifizierten Maus Organoiden wurde ein ähnlicher experimenteller Ansatz verfolgt. Hier wurden drei prominente Faktoren identifiziert: Der erste Faktor unterteilte *Apc* Wildtyp von $Apc^{-/-}$ Organoiden. Verlust der *Apc* Funktion führte zu erhöhter Expression von Proliferations- und DNA Reparatur-Markern. Es führte ebenso zu erhöhter Sensibilität gegenüber Perturbationen des Cytoskeletts, zum Beispiel durch Taxane und FAK-Inhibitoren, sowie morphologisch zu einem Verlust der sphärischen und symmetrischen multizellulären Organisation. Organoide mit $Apc^{-/-}$ Genotyp zeigten eine Unabhängigkeit von üblicherweise essenziellen Wnt Liganden im Kulturmedium. Mit Bezug auf den zellularen Metabolismus, führte der Verlust von *Apc* zu einer Reduktion von Transkripten der beta-Oxidation, sowie einer Akkumulation von Triacylglycerine und Cholesterolester Speicherlipiden.

Der zweite Faktor unterteilte $Kras^{G12D/+}$ Organoide von Wildtyp Organoiden. Organoide mit $Kras^{G12D/+}$ Genotyp zeigten Zeichen einer Onkogen-induzierten Seneszenz, reduzierte Expression von Transkripten der beta-Oxidation, eine Akkumulation von Speicherlipiden, sowie eine deutliche Reduktion von Transkripten der oxidativen Phosphorylierung. Es bestand eine hohe relative Sensibilität gegenüber ERK- und MEK-Inhibitoren, und eine geringe relative Sensibilität gegenüber Inhibitoren des EGF Rezeptors.

Der dritte Faktor trennte $Apc^{-/-}$ Organoide von $Apc^{-/-}$ / $Kras^{G12D/+}$ Organoiden. Modelle mit $Apc^{-/-}$ / $Kras^{G12D/+}$ Genotyp zeigten eine erhöhte Sensitivität gegenüber mTOR-Inhibitoren sowie eine erhöhte Expression einer mTORC1-Aktivierungssignatur. Wie zuvor konnten die gelernten Faktoren genutzt werden, um nicht letale Medikamenteneffekte systematisch zu identifizieren. Hierbei zeigten Inhibitoren der GSK3-Beta-Kinase einen deutlichen Effekt auf Faktor 1.

Die Ergebnisse dieser Arbeit beleuchten primäre molekulare Faktoren der Organoidarchitektur und ihre molekulare Bedeutung für das kolorektale Karzinom.

Bibliography

Accioly, Maria T., Patricia Pacheco, Clarissa M. Maya-Monteiro, Nina Carrossini, Bruno K. Robbs, Silvia S. Oliveira, Cristiane Kaufmann, José A. Morgado-Diaz, Patricia T. Bozza, and João P. B. Viola (2008). Lipid bodies are reservoirs of cyclooxygenase-2 and sites of prostaglandin-E2 synthesis in colon cancer cells. *Cancer Research* 68, pp. 1732–1740. DOI: 10.1158/0008-5472.CAN-07-1999.

Aderka, Dan, Sebastian Stintzing, and Volker Heinemann (2019). Explaining the unexplainable: discrepancies in results from the CALGB/SWOG 80405 and FIRE-3 studies. *The Lancet Oncology* 20, e274–e283. DOI: 10.1016/S1470-2045(19)30172-X.

André, Thierry, Kai-Keen Shiu, Tae Won Kim, Benny Vittrup Jensen, Lars Henrik Jensen, Cornelis Punt, Denis Smith, Rocio Garcia-Carbonero, Manuel Benavides, Peter Gibbs, Christelle de la Fouchardiere, Fernando Rivera, Elena Elez, Johanna Bendell, Dung T. Le, Takayuki Yoshino, Eric Van Cutsem, Ping Yang, Mohammed Z. H. Farooqui, Patricia Marinello, Luis A. Diaz, and KEYNOTE-177 Investigators (2020). Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. *The New England Journal of Medicine* 383, pp. 2207–2218. DOI: 10.1056/NEJMoa2017699.

Argelaguet, Ricard, Damien Arnol, Danila Bredikhin, Yonatan Deloro, Britta Velten, John C. Marioni, and Oliver Stegle (2020). MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biology* 21, p. 111. DOI: 10.1186/s13059-020-02015-1.

Argelaguet, Ricard, Britta Velten, Damien Arnol, Sascha Dietrich, Thorsten Zenz, John C Marioni, Florian Buettner, Wolfgang Huber, and Oliver Stegle (2018). Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Molecular Systems Biology* 14, e8124. DOI: 10.15252/msb.20178124.

Ashton, Gabrielle H., Jennifer P. Morton, Kevin Myant, Toby J. Phesse, Rachel A. Ridgway, Victoria Marsh, Julie A. Wilkins, Dimitris Athineos, Vanesa Muncan, Richard Kemp, Kristi Neufeld, Hans Clevers, Valerie Brunton, Douglas J. Winton, Xiaoyan Wang, Rosalie C. Sears, Alan R. Clarke, Margaret C. Frame, and Owen J. Sansom (2010). Focal adhesion kinase is required for intestinal regeneration and tumorigenesis downstream of Wnt/c-Myc signaling. *Developmental Cell* 19, pp. 259–269. DOI: 10.1016/j.devcel.2010.07.015.

Badder, Luned M., Andrew J. Hollins, Bram Herpers, Kuan Yan, Kenneth B. Ewan, Mairian Thomas, Jennifer R. Shone, Delyth A. Badder, Marc Naven, Kevin E. Ashelford, Rachel Hargest, Alan R. Clarke, Christina Esdar, Hans-Peter Buchstaller, J. Mark Treherne, Sylvia Boj, Bahar Ramezanpour, Dirk Wienke, Leo S. Price, Paul H. Shaw, and Trevor C. Dale (2020). 3D imaging of colorectal cancer organoids identifies responses to Tankyrase inhibitors. *PloS One* 15, e0235319. DOI: 10.1371/journal.pone.0235319.

Basak, Onur, Joep Beumer, Kay Wiebrands, Hiroshi Seno, Alexander van Oudenaarden, and Hans Clevers (2017). Induced Quiescence of Lgr5+ Stem Cells in Intestinal Organoids Enables Differentiation of Hormone-Producing Enteroendocrine Cells. *Cell Stem Cell* 20, 177–190.e4. DOI: 10.1016/j.stem.2016.11.001.

Beckwith, Heather and Douglas Yee (2015). Minireview: Were the IGF Signaling Inhibitors All Bad? *Molecular Endocrinology* 29, pp. 1549–1557. DOI: 10.1210/me.2015-1157.

Bennecke, Moritz, Lydia Kriegl, Monther Bajbouj, Kristin Retzlaff, Sylvie Robine, Andreas Jung, Melek C. Arkan, Thomas Kirchner, and Florian R. Greten (2010). Ink4a/Arf and oncogene-induced senescence prevent tumor progression during alternative colorectal tumorigenesis. *Cancer Cell* 18, pp. 135–146. DOI: 10.1016/j.ccr.2010.06.013.

Betge, Johannes, Niklas Rindtorff, Jan Sauer, Benedikt Rauscher, Clara Dingert, Haristi Gaitantzi, Frank Herweck, Kauthar Srour-Mhanna, Thilo Miersch, Erica Valentini, Kim E. Boonekamp, Veronika Hauber, Tobias Gutting, Larissa Frank, Sebastian Belle, Timo Gaiser, Inga Buchholz, Ralf Jesenofsky, Nicolai Härtel, Tianzuo Zhan, Bernd Fischer, Katja Breitkopf-Heinlein, Elke Burgermeister, Matthias P. Ebert, and Michael Boutros (2022). The drug-induced phenotypic landscape of colorectal cancer organoids. *Nature Communications* 13, p. 3135. DOI: 10.1038/s41467-022-30722-9.

Billmann, Maximilian, Thomas Horn, Bernd Fischer, Thomas Sandmann, Wolfgang Huber, and Michael Boutros (2016). A genetic interaction map of cell cycle regulators. *Molecular Biology of the Cell* 27, pp. 1397–1407. DOI: 10.1091/mbc.E15-07-0467.

Boehnke, Karsten, Philip W. Iversen, Dirk Schumacher, María José Lallena, Rubén Haro, Joaquín Amat, Johannes Haybaeck, Sandra Liebs, Martin Lange, Reinhold Schäfer, Christian R. A. Regenbrecht, Christoph Reinhard,

and Juan A. Velasco (2016). Assay Establishment and Validation of a High-Throughput Screening Platform for Three-Dimensional Patient-Derived Colon Cancer Organoid Cultures. *SLAS Discovery* 21, pp. 931–941. DOI: 10.1177/1087057116650965.

Borowsky, Jennifer, Troy Dumenil, Mark Bettington, Sally-Ann Pearson, Catherine Bond, Lochlan Fennell, Cheng Liu, Diane McKeone, Christophe Rosty, Ian Brown, Neal Walker, Barbara Leggett, and Vicki Whitehall (2018). The role of APC in WNT pathway activation in serrated neoplasia. *Modern Pathology: An Official Journal of the United States and Canadian Academy of Pathology, Inc* 31, pp. 495–504. DOI: 10.1038/modpathol.2017.150.

Boutin, Adam T., Wen-Ting Liao, Melody Wang, Soyoon Sarah Hwang, Tatiana V. Karpinets, Hannah Cheung, Gerald C. Chu, Shan Jiang, Jian Hu, Kyle Chang, Eduardo Vilar, Xingzhi Song, Jianhua Zhang, Scott Kopetz, Andrew Futreal, Y. Alan Wang, Lawrence N. Kwong, and Ronald A. DePinho (2017). Oncogenic *Kras* drives invasion and maintains metastases in colorectal cancer. *Genes & Development* 31, pp. 370–382. DOI: 10.1101/gad.293449.116.

Caicedo, Juan C., Sam Cooper, Florian Heigwer, Scott Warchal, Peng Qiu, Csaba Molnar, Aliaksei S. Vasilevich, Joseph D. Barry, Harmanjit Singh Bansal, Oren Kraus, Mathias Wawer, Lassi Paavolainen, Markus D. Herrmann, Mohammad Rohban, Jane Hung, Holger Hennig, John Concannon, Ian Smith, Paul A. Clemons, Shantanu Singh, Paul Rees, Peter Horvath, Roger G. Linington, and Anne E. Carpenter (2017). Data-analysis strategies for image-based cell profiling. *Nature Methods* 14, pp. 849–863. DOI: 10.1038/nmeth.4397.

Caicedo, Juan C, Shantanu Singh, and Anne E Carpenter (2016). Applications in image-based profiling of perturbations. *Current Opinion in Biotechnology* 39, pp. 134–142. DOI: 10.1016/j.copbio.2016.04.003.

Carpenter, Anne E. (2007). Image-based chemical screening. *Nature Chemical Biology* 3, pp. 461–465. DOI: 10.1038/nchembio.2007.15.

Cattaneo, Chiara M, Krijn K Dijkstra, Lorenzo F Fanchi, Sander Kelderman, Sovann Kaing, Nienke van Rooij, Stieneke van den Brink, Ton N Schumacher, and Emile E Voest (2020). Tumor organoid – T cell co-culture systems. *Nature protocols* 15, pp. 15–39. DOI: 10.1038/s41596-019-0232-9.

Chandrasekaran, Srinivas Niranj, Hugo Ceulemans, Justin D. Boyd, and Anne E. Carpenter (2021). Image-based profiling for drug discovery: due for a machine-learning upgrade? *Nature Reviews Drug Discovery* 20, pp. 145–159. DOI: 10.1038/s41573-020-00117-w.

Cho, Kathleen R. and Bert Vogelstein (1992). Genetic alterations in the adenoma–carcinoma sequence. *Cancer* 70, pp. 1727–1731. DOI: 10.1002/1097-0142(19920915)70:4+<1727::AID-CNCR2820701613>3.0.CO;2-P.

Clevers, Hans (2011). The cancer stem cell: premises, promises and challenges. *Nature Medicine* 17, pp. 313–319. DOI: 10.1038/nm.2304.

— (2013). The intestinal crypt, a prototype stem cell compartment. *Cell* 154, pp. 274–284. DOI: 10.1016/j.cell.2013.07.004.

Cruz, André L. S., Ester de A. Barreto, Narayana P. B. Fazolini, João P. B. Viola, and Patricia T. Bozza (2020). Lipid droplets: platforms with multiple functions in cancer hallmarks. *Cell Death & Disease* 11, pp. 1–16. DOI: 10.1038/s41419-020-2297-3.

Cruz, André L. S., Nina Carrossini, Leonardo K. Teixeira, Luis F. Ribeiro-Pinto, Patricia T. Bozza, and João P. B. Viola (2019). Cell Cycle Progression Regulates Biogenesis and Cellular Localization of Lipid Droplets. *Molecular and Cellular Biology* 39, e00374–18. DOI: 10.1128/MCB.00374-18.

DeBerardinis, Ralph J. and Navdeep S. Chandel (2020). We need to talk about the Warburg effect. *Nature Metabolism* 2, pp. 127–129. DOI: 10.1038/s42255-020-0172-2.

Dixit, Atray, Oren Parnas, Biyu Li, Jenny Chen, Charles P. Fulco, Livnat Jerby-Arnon, Nemanja D. Marjanovic, Danielle Dionne, Tyler Burks, Raktima Raychowdhury, Britt Adamson, Thomas M. Norman, Eric S. Lander, Jonathan S. Weissman, Nir Friedman, and Aviv Regev (2016). Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* 167, 1853–1866.e17. DOI: 10.1016/j.cell.2016.11.038.

Dow, Lukas E., Kevin P. O'Rourke, Janelle Simon, Darjus F. Tschaharganeh, Johan H. van Es, Hans Clevers, and Scott W. Lowe (2015). Apc Restoration Promotes Cellular Differentiation and Reestablishes Crypt Homeostasis in Colorectal Cancer. *Cell* 161, pp. 1539–1552. DOI: 10.1016/j.cell.2015.05.033.

Driehuis, Else, Arne van Hoeck, Kat Moore, Sigrid Kolders, Hayley E. Francies, M. Can Gulersonmez, Edwin C. A. Stigter, Boudewijn Burgering, Veerle Geurts, Ana Gracanin, Gergana Bounova, Folkert H. Morsink, Robert Vries, Sylvia Boj, Johan van Es, G. Johan A. Offerhaus, Onno Kranenburg, Mathew J. Garnett, Lodewyk Wessels, Edwin Cuppen, Lodewijk A. A. Brosens, and Hans Clevers (2019). Pancreatic cancer organoids recapitulate disease and allow personalized drug screening. *Proceedings of the National Academy of Sciences* 116, pp. 26580–26590. DOI: 10.1073/pnas.1911273116.

Drost, Jarno, Richard H. van Jaarsveld, Bas Ponsioen, Cheryl Zimberlin, Ruben van Boxtel, Arjan Buijs, Norman Sachs, René M. Overmeer, G. Johan Offerhaus, Harry Begthel, Jeroen Korving, Marc van de Wetering, Gerald Schwank, Meike Logtenberg, Edwin Cuppen, Hugo J. Snippert, Jan Paul Medema, Geert J. P. L. Kops, and

Hans Clevers (2015). Sequential cancer mutations in cultured human intestinal stem cells. *Nature* 521, pp. 43–47. DOI: 10.1038/nature14415.

Dunne, Philip D., Matthew Alderdice, Paul G. O'Reilly, Aideen C. Roddy, Amy M. B. McCorry, Susan Richman, Tim Maughan, Simon S. McDade, Patrick G. Johnston, Daniel B. Longley, Elaine Kay, Darragh G. McArt, and Mark Lawler (2017). Cancer-cell intrinsic gene expression signatures overcome intratumoural heterogeneity bias in colorectal cancer patient classification. *Nature Communications* 8, p. 15657. DOI: 10.1038/ncomms15657.

Eide, Peter W., Seyed H. Moosavi, Ina A. Eilertsen, Tuva H. Brunsell, Jonas Langerud, Kaja C. G. Berg, Bård I. Røsok, Bjørn A. Bjørnbeth, Arild Nesbakken, Ragnhild A. Lothe, and Anita Sveen (2021). Metastatic heterogeneity of the consensus molecular subtypes of colorectal cancer. *NPJ Genomic Medicine* 6, p. 59. DOI: 10.1038/s41525-021-00223-7.

Fang, Jing Yuan and Bruce C. Richardson (2005). The MAPK signalling pathways and colorectal cancer. *The Lancet. Oncology* 6, pp. 322–327. DOI: 10.1016/S1470-2045(05)70168-6.

Fearon, Eric R. (2011). Molecular genetics of colorectal cancer. *Annual Review of Pathology* 6, pp. 479–507. DOI: 10.1146/annurev-pathol-011110-130235.

Fridman, A. L. and M. A. Tainsky (2008). Critical pathways in cellular senescence and immortalization revealed by gene expression profiling. *Oncogene* 27, pp. 5975–5987. DOI: 10.1038/onc.2008.213.

Fujii, Masayuki, Mami Matano, Kohta Toshimitsu, Ai Takano, Yohei Mikami, Shingo Nishikori, Shinya Sugimoto, and Toshiro Sato (2018). Human Intestinal Organoids Maintain Self-Renewal Capacity and Cellular Diversity in Niche-Inspired Culture Condition. *Cell Stem Cell* 23, 787–793.e6. DOI: 10.1016/j.stem.2018.11.016.

Fujii, Masayuki, Mariko Shimokawa, Shoichi Date, Ai Takano, Mami Matano, Kosaku Nanki, Yuki Ohta, Kohta Toshimitsu, Yoshihiro Nakazato, Kenta Kawasaki, Toshio Uraoka, Toshiaki Watanabe, Takanori Kanai, and Toshiro Sato (2016). A Colorectal Tumor Organoid Library Demonstrates Progressive Loss of Niche Factor Requirements during Tumorigenesis. *Cell Stem Cell* 18, pp. 827–838. DOI: 10.1016/j.stem.2016.04.003.

Ganesh, Karuna, Chao Wu, Kevin P. O'Rourke, Bryan C. Szeglin, Youyun Zheng, Charles-Etienne Gabriel Sauvé, Mohammad Adileh, Isaac Wasserman, Michael R. Marco, Amanda S. Kim, Maha Shady, Francisco Sanchez-Vega, Wouter R. Karthaus, Helen H. Won, Seo-Hyun Choi, Raphael Pelossof, Afsar Barlas, Peter Ntiamoah, Emmanouil Pappou, Arthur Elghouayel, James S. Strong, Chin-Tung Chen, Jennifer W. Harris, Martin R. Weiser, Garrett M. Nash, Jose G. Guillem, Iris H. Wei, Richard N. Kolesnick, Harini Veeraraghavan, Eduardo J. Ortiz, Iva Petkovska, Andrea Cercek, Katia O. Manova-Todorova, Leonard B. Saltz, Jessica A. Lavery, Ronald P. DeMatteo, Joan Massagué, Philip B. Paty, Rona Yaeger, Xi Chen, Sujata Patil, Hans Clevers, Michael F. Berger, Scott W. Lowe, Jinru Shia, Paul B. Romesser, Lukas E. Dow, Julio Garcia-Aguilar, Charles L. Sawyers, and J. Joshua Smith (2019). A rectal cancer organoid platform to study individual responses to chemoradiation. *Nature Medicine* 25, pp. 1607–1614. DOI: 10.1038/s41591-019-0584-2.

Gautier, Laurent, Leslie Cope, Benjamin M. Bolstad, and Rafael A. Irizarry (2004). affy–analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics (Oxford, England)* 20, pp. 307–315. DOI: 10.1093/bioinformatics/btg405.

Gehart, Helmuth and Hans Clevers (2019). Tales from the crypt: new insights into intestinal stem cells. *Nature reviews. Gastroenterology & hepatology* 16, pp. 19–34. DOI: 10.1038/s41575-018-0081-y.

Gibson, Christopher C., Weiquan Zhu, Chadwick T. Davis, Jay A. Bowman-Kirigin, Aubrey C. Chan, Jing Ling, Ashley E. Walker, Luca Goitre, Simona Delle Monache, Saverio Francesco Retta, Yan-Ting E. Shiu, Allie H. Grossmann, Kirk R. Thomas, Anthony J. Donato, Lisa A. Lesniewski, Kevin J. Whitehead, and Dean Y. Li (2015). Strategy for identifying repurposed drugs for the treatment of cerebral cavernous malformation. *Circulation* 131, pp. 289–299. DOI: 10.1161/CIRCULATIONAHA.114.010403.

Griss, Johannes, Guilherme Viteri, Konstantinos Sidiropoulos, Vy Nguyen, Antonio Fabregat, and Henning Hermjakob (2020). ReactomeGSA - Efficient Multi-Omics Comparative Pathway Analysis. *Molecular & cellular proteomics: MCP* 19, pp. 2115–2125. DOI: 10.1074/mcp.TIR120.002155.

Guinney, Justin, Rodrigo Dienstmann, Xin Wang, Aurélien de Reyniès, Andreas Schlicker, Charlotte Soneson, Laetitia Marisa, Paul Roepman, Gift Nyamundanda, Paolo Angelino, Brian M. Bot, Jeffrey S. Morris, Iris M. Simon, Sarah Gerster, Evelyn Fessler, Felipe De Sousa E Melo, Edoardo Missiaglia, Hena Ramay, David Barras, Krisztian Homicsko, Dipen Maru, Ganiraju C. Manyam, Bradley Broom, Valerie Boige, Beatriz Perez-Villamil, Ted Laderas, Ramon Salazar, Joe W. Gray, Douglas Hanahan, Josep Tabernero, Rene Bernards, Stephen H. Friend, Pierre Laurent-Puig, Jan Paul Medema, Anguraj Sadanandam, Lodewyk Wessels, Mauro Delorenzi, Scott Kopetz, Louis Vermeulen, and Sabine Tejpar (2015). The consensus molecular subtypes of colorectal cancer. *Nature Medicine* 21, pp. 1350–1356. DOI: 10.1038/nm.3967.

Hafner, Marc, Mario Niepel, Mirra Chung, and Peter K. Sorger (2016). Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. *Nature Methods* 13, pp. 521–527. DOI: 10.1038/nmeth.3853.

He, Xi C., Jiwang Zhang, Wei-Gang Tong, Ossama Tawfik, Jason Ross, David H. Scoville, Qiang Tian, Xin Zeng, Xi He, Leanne M. Wiedemann, Yuji Mishina, and Linheng Li (2004). BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt–β-catenin signaling. *Nature Genetics* 36, pp. 1117–1121. DOI: 10.1038/ng1430.

Heigwer, Florian, Grainne Kerr, and Michael Boutros (2014). E-CRISP: fast CRISPR target site identification. *Nature Methods* 11, pp. 122–123. DOI: 10.1038/nmeth.2812.

Hospedales, Timothy, Antreas Antoniou, Paul Micaelli, and Amos Storkey (2020). Meta-Learning in Neural Networks: A Survey.

Hua, Hui, Qingbin Kong, Jie Yin, Jin Zhang, and Yangfu Jiang (2020). Insulin-like growth factor receptor signaling in tumorigenesis and drug resistance: a challenge for cancer therapy. *Journal of Hematology & Oncology* 13, p. 64. DOI: 10.1186/s13045-020-00904-3.

Huber, Wolfgang, Vincent J. Carey, Robert Gentleman, Simon Anders, Marc Carlson, Benilton S. Carvalho, Hector Corrada Bravo, Sean Davis, Laurent Gatto, Thomas Girke, Raphael Gottardo, Florian Hahne, Kasper D. Hansen, Rafael A. Irizarry, Michael Lawrence, Michael I. Love, James MacDonald, Valerie Obenchain, Andrzej K. Oleś, Hervé Pagès, Alejandro Reyes, Paul Shannon, Gordon K. Smyth, Dan Tenenbaum, Levi Waldron, and Martin Morgan (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods* 12, pp. 115–121. DOI: 10.1038/nmeth.3252.

Huber, Wolfgang, Anja von Heydebreck, Holger Sültmann, Annemarie Poustka, and Martin Vingron (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics (Oxford, England)* 18 Suppl 1, S96–104. DOI: 10.1093/bioinformatics/18.suppl_1.s96.

Huch, Meritxell, Craig Dorrell, Sylvia F. Boj, Johan H. van Es, Vivian S. W. Li, Marc van de Wetering, Toshiro Sato, Karien Hamer, Nobuo Sasaki, Milton J. Finegold, Annelise Haft, Robert G. Vries, Markus Grompe, and Hans Clevers (2013). In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. *Nature* 494, pp. 247–250. DOI: 10.1038/nature11826.

Imkeller, Katharina, Giulia Ambrosi, Nancy Klemm, Ainara Claveras Cabezudo, Luisa Henkel, Wolfgang Huber, and Michael Boutros (2022). Metabolic balance in colorectal cancer is maintained by optimal Wnt signaling levels. *Molecular Systems Biology* 18, e10874. DOI: 10.15252/msb.202110874.

Irizarry, Rafael A., Bridget Hobbs, Francois Collin, Yasmin D. Beazer-Barclay, Kristen J. Antonellis, Uwe Scherf, and Terence P. Speed (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics (Oxford, England)* 4, pp. 249–264. DOI: 10.1093/biostatistics/4.2.249.

Isella, Claudio, Francesco Brundu, Sara E. Bellomo, Francesco Galimi, Eugenia Zanella, Roberta Porporato, Consalvo Petti, Alessandro Fiori, Francesca Orzan, Rebecca Senetta, Carla Boccaccio, Elisa Ficarra, Luigi Marchionni, Livio Trusolino, Enzo Medico, and Andrea Bertotti (2017a). Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. *Nature Communications* 8, p. 15107. DOI: 10.1038/ncomms15107.

— (2017b). Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. *Nature Communications* 8, p. 15107. DOI: 10.1038/ncomms15107.

Jackson, E. L., N. Willis, K. Mercer, R. T. Bronson, D. Crowley, R. Montoya, T. Jacks, and D. A. Tuveson (2001). Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes & Development* 15, pp. 3243–3248. DOI: 10.1101/gad.943001.

Jentzsch, Valerie, Leeza Osipenko, Jack W. Scannell, and John A. Hickman (2023). Costs and Causes of Oncology Drug Attrition With the Example of Insulin-Like Growth Factor-1 Receptor Inhibitors. *JAMA Network Open* 6, e2324977. DOI: 10.1001/jamanetworkopen.2023.24977.

Joanito, Ignasius, Pratyaksha Wirapati, Nancy Zhao, Zahid Nawaz, Grace Yeo, Fiona Lee, Christine L. P. Eng, Dominique Camat Macalinao, Merve Kahraman, Harini Srinivasan, Vairavan Lakshmanan, Sara Verbandt, Petros Tsantoulis, Nicole Gunn, Prasanna Nori Venkatesh, Zhong Wee Poh, Rahul Nahar, Hsueh Ling Janice Oh, Jia Min Loo, Shumei Chia, Lih Feng Cheow, Elsie Cheruba, Michael Thomas Wong, Lindsay Kua, Clarinda Chua, Andy Nguyen, Justin Golovan, Anna Gan, Wan-Jun Lim, Yu Amanda Guo, Choon Kong Yap, Brenda Tay, Yourae Hong, Dawn Qingqing Chong, Aik-Yong Chok, Woong-Yang Park, Shuting Han, Mei Huan Chang, Isaac Seow-En, Cherylin Fu, Ronnie Mathew, Ee-Lin Toh, Lewis Z. Hong, Anders Jacobsen Skanderup, Ramanuj DasGupta, Chin-Ann Johnny Ong, Kiat Hon Lim, Emile K. W. Tan, Si-Lin Koo, Wei Qiang Leow, Sabine Tejpar, Shyam Prabhakar, and Iain Beehuat Tan (2022). Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and refines the consensus molecular classification of colorectal cancer. *Nature Genetics* 54, pp. 963–975. DOI: 10.1038/s41588-022-01100-4.

Kamburov, Atanas, Konstantin Pentchev, Hanna Galicka, Christoph Wierling, Hans Lehrach, and Ralf Herwig (2011). ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Research* 39, pp. D712–717. DOI: 10.1093/nar/gkq1156.

Klami, Arto, Seppo Virtanen, Eemeli Leppäaho, and Samuel Kaski (2014). Group Factor Analysis.

Kopetz, Scott, Axel Grothey, Rona Yaeger, Eric Van Cutsem, Jayesh Desai, Takayuki Yoshino, Harpreet Wasan, Fortunato Ciardiello, Fotios Loupakis, Yong Sang Hong, Neeltje Steeghs, Tormod K. Guren, Hendrik-Tobias Arkenau, Pilar Garcia-Alfonso, Per Pfeiffer, Sergey Orlov, Sara Lonardi, Elena Elez, Tae-Won Kim, Jan H. M. Schellens, Christina Guo, Asha Krishnan, Jeroen Dekervel, Van Morris, Aitana Calvo Ferrandiz, L. S. Tarpgaard, Michael Braun, Ashwin Gollerkeri, Christopher Keir, Kati Maharry, Michael Pickard, Janna Christy-Bittel, Lisa Anderson, Victor Sandor, and Josep Tabernero (2019). Encorafenib, Binimetinib, and Cetuximab in BRAF V600E-Mutated Colorectal Cancer. *The New England Journal of Medicine* 381, pp. 1632–1643. DOI: 10.1056/NEJMoa1908075.

Koren, Yehuda, Robert Bell, and Chris Volinsky (2009). Matrix Factorization Techniques for Recommender Systems. *Computer* 42, pp. 30–37. DOI: 10.1109/MC.2009.263.

Laufer, Christina, Bernd Fischer, Maximilian Billmann, Wolfgang Huber, and Michael Boutros (2013). Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. *Nature Methods* 10, pp. 427–431. DOI: 10.1038/nmeth.2436.

Leek, Jeffrey T. and John D. Storey (2007). Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. *PLOS Genetics* 3, e161. DOI: 10.1371/journal.pgen.0030161.

Lenz, Heinz-Josef, Fang-Shu Ou, Alan P. Venook, Howard S. Hochster, Donna Niedzwiecki, Richard M. Goldberg, Robert J. Mayer, Monica M. Bertagnolli, Charles D. Blanke, Tyler Zemla, Xueping Qu, Pratyaksha Wirapati, Sabine Tejpar, Federico Innocenti, and Omar Kabbarah (2019). Impact of Consensus Molecular Subtype on Survival in Patients With Metastatic Colorectal Cancer: Results From CALGB/SWOG 80405 (Alliance). *Journal of Clinical Oncology* 37, pp. 1876–1885. DOI: 10.1200/JCO.18.02258.

Li, Yingming, Ming Yang, and Zhongfei Zhang (2019). A Survey of Multi-View Representation Learning. *IEEE Transactions on Knowledge and Data Engineering* 31, pp. 1863–1883. DOI: 10.1109/TKDE.2018.2872063.

Liu, Jun, Shifeng Pan, Mindy H. Hsieh, Nicholas Ng, Fangxian Sun, Tao Wang, Shailaja Kasibhatla, Alwin G. Schuller, Allen G. Li, Dai Cheng, Jie Li, Celin Tompkins, AnneMarie Pferdekamper, Auzon Steffy, Jane Cheng, Colleen Kowal, Van Phung, Guirong Guo, Yan Wang, Martin P. Graham, Shannon Flynn, J. Chad Brenner, Chun Li, M. Cristina Villarroel, Peter G. Schultz, Xu Wu, Peter McNamara, William R. Sellers, Lilli Petruzzelli, Anthony L. Boral, H. Martin Seidel, Margaret E. McLaughlin, Jianwei Che, Thomas E. Carey, Gary Vanasse, and Jennifer L. Harris (2013). Targeting Wnt-driven cancer through the inhibition of Porcupine by LGK974. *Proceedings of the National Academy of Sciences of the United States of America* 110, pp. 20224–20229. DOI: 10.1073/pnas.1314239110.

Lukonin, Ilya, Denise Serra, Ludivine Challet Meylan, Katrin Volkmann, Janine Baaten, Rui Zhao, Shelly Meeusen, Karyn Colman, Francisca Maurer, Michael B. Stadler, Jeremy Jenkins, and Prisca Liberali (2020). Phenotypic landscape of intestinal organoid regeneration. *Nature* 586, pp. 275–280. DOI: 10.1038/s41586-020-2776-9.

Luo, Feijun, David G Brooks, Hongtao Ye, Rifat Hamoudi, George Poulogiannis, Charles E Patek, Douglas J Winton, and Mark J Arends (2009). Mutated K-rasAsp12 promotes tumourigenesis in ApcMin mice more in the large than the small intestines, with synergistic effects between K-ras and Wnt pathways. *International Journal of Experimental Pathology* 90, pp. 558–574. DOI: 10.1111/j.1365-2613.2009.00667.x.

Markowitz, Sanford D. and Monica M. Bertagnolli (2009). Molecular origins of cancer: Molecular basis of colorectal cancer. *The New England Journal of Medicine* 361, pp. 2449–2460. DOI: 10.1056/NEJMra0804588.

Matano, Mami, Shoichi Date, Mariko Shimokawa, Ai Takano, Masayuki Fujii, Yuki Ohta, Toshiaki Watanabe, Takanori Kanai, and Toshiro Sato (2015). Modeling colorectal cancer using CRISPR-Cas9–mediated engineering of human intestinal organoids. *Nature Medicine* 21, pp. 256–262. DOI: 10.1038/nm.3802.

Matsumoto, Shinji, Katsumi Fumoto, Tetsuji Okamoto, Kozo Kaibuchi, and Akira Kikuchi (2010). Binding of APC and dishevelled mediates Wnt5a-regulated focal adhesion dynamics in migrating cells. *The EMBO Journal* 29, pp. 1192–1204. DOI: 10.1038/emboj.2010.26.

Merker, Sebastian R., Jürgen Weitz, and Daniel E. Stange (2016). Gastrointestinal organoids: How they gut it out. *Developmental Biology* 420, pp. 239–250. DOI: 10.1016/j.ydbio.2016.08.010.

Merlos-Suárez, Anna, Francisco M. Barriga, Peter Jung, Mar Iglesias, María Virtudes Céspedes, David Rossell, Marta Sevillano, Xavier Hernando-Momblona, Victoria da Silva-Diz, Purificación Muñoz, Hans Clevers, Elena Sancho, Ramón Mangues, and Eduard Batlle (2011). The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* 8, pp. 511–524. DOI: 10.1016/j.stem.2011.02.020.

Mina, Marco, Franck Raynaud, Daniele Tavernari, Elena Battistello, Stephanie Sungalee, Sadegh Saghafinia, Titouan Laessle, Francisco Sanchez-Vega, Nikolaus Schultz, Elisa Oricchio, and Giovanni Ciriello (2017). Conditional Selection of Genomic Alterations Dictates Cancer Evolution and Oncogenic Dependencies. *Cancer Cell* 32, 155–168.e6. DOI: 10.1016/j.ccell.2017.06.010.

Moser, A. R., C. Luongo, K. A. Gould, M. K. McNeley, A. R. Shoemaker, and W. F. Dove (1995). ApcMin: a mouse model for intestinal and mammary tumorigenesis. *European Journal of Cancer (Oxford, England: 1990)* 31A, pp. 1061–1064. DOI: 10.1016/0959-8049(95)00181-h.

Moshkov, Nikita, Tim Becker, Kevin Yang, Peter Horvath, Vlado Dancik, Bridget K. Wagner, Paul A. Clemons, Shantanu Singh, Anne E. Carpenter, and Juan C. Caicedo (2023). Predicting compound activity from phenotypic profiles and chemical structures. *Nature Communications* 14, p. 1967. DOI: 10.1038/s41467-023-37570-1.

Nishihara, Reiko, Kana Wu, Paul Lochhead, Teppei Morikawa, Xiaoyun Liao, Zhi Rong Qian, Kentaro Inamura, Sun A. Kim, Aya Kuchiba, Mai Yamauchi, Yu Imamura, Walter C. Willett, Bernard A. Rosner, Charles S. Fuchs, Edward Giovannucci, Shuji Ogino, and Andrew T. Chan (2013). Long-term colorectal-cancer incidence and mortality after lower endoscopy. *The New England Journal of Medicine* 369, pp. 1095–1105. DOI: 10.1056/NEJMoa1301969.

Nüsslein-Volhard, Christiane and Eric Wieschaus (1980). Mutations affecting segment number and polarity in Drosophila. *Nature* 287, pp. 795–801. DOI: 10.1038/287795a0.

Ogino, Shuji, Katsuhiko Nosho, Gregory J. Kirkner, Takako Kawasaki, Jeffrey A. Meyerhardt, Massimo Loda, Edward L. Giovannucci, and Charles S. Fuchs (2009). CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut* 58, pp. 90–96. DOI: 10.1136/gut.2008.155473.

Okazaki, Hiroyasu, Seidai Sato, Kazuya Koyama, Shun Morizumi, Shuichi Abe, Momoyo Azuma, Yajuan Chen, Hisatsugu Goto, Yoshinori Aono, Hirohisa Ogawa, Kozo Kagawa, Haruka Nishimura, Hiroshi Kawano, Yuko Toyoda, Hisanori Uehara, Hiroyuki Kouji, and Yasuhiko Nishioka (2019). The novel inhibitor PRI-724 for Wnt/β-catenin/CBP signaling ameliorates bleomycin-induced pulmonary fibrosis in mice. *Experimental Lung Research* 45, pp. 188–199. DOI: 10.1080/01902148.2019.1638466.

Ooft, S. N., F. Weeber, L. Schipper, K. K. Dijkstra, C. M. McLean, S. Kaing, J. van de Haar, W. Prevoo, E. van Werkhoven, P. Snaebjornsson, L. R. Hoes, M. Chalabi, D. van der Velden, M. van Leerdam, H. Boot, C. Grootscholten, A. D. R. Huitema, H. J. Bloemendal, E. Cuppen, and E. E. Voest (2021). Prospective experimental treatment of colorectal cancer patients based on organoid drug responses. *ESMO Open* 6. DOI: 10.1016/j.esmoop.2021.100103.

Ooft, Salo N., Fleur Weeber, Krijn K. Dijkstra, Chelsea M. McLean, Sovann Kaing, Erik van Werkhoven, Luuk Schipper, Louisa Hoes, Daniel J. Vis, Joris van de Haar, Warner Prevoo, Petur Snaebjornsson, Daphne van der Velden, Michelle Klein, Myriam Chalabi, Henk Boot, Monique van Leerdam, Haiko J. Bloemendal, Laurens V. Beerepoot, Lodewyk Wessels, Edwin Cuppen, Hans Clevers, and Emile E. Voest (2019). Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Science Translational Medicine* 11, eaay2574. DOI: 10.1126/scitranslmed.aay2574.

O'Reilly, Kathryn E., Fredi Rojo, Qing-Bai She, David Solit, Gordon B. Mills, Debra Smith, Heidi Lane, Francesco Hofmann, Daniel J. Hicklin, Dale L. Ludwig, Jose Baselga, and Neal Rosen (2006). mTOR inhibition induces upstream receptor tyrosine kinase signaling and activates Akt. *Cancer Research* 66, pp. 1500–1508. DOI: 10.1158/0008-5472.CAN-05-2925.

Pancione, Massimo, Andrea Remo, and Vittorio Colantuoni (2012). Genetic and Epigenetic Events Generate Multiple Pathways in Colorectal Cancer Progression. *Pathology Research International* 2012, e509348. DOI: 10.1155/2012/509348.

Parsons, Marie J., Tuomas Tammela, and Lukas E. Dow (2021). WNT as a Driver and Dependency in Cancer. *Cancer Discovery* 11, pp. 2413–2429. DOI: 10.1158/2159-8290.CD-21-0190.

Pau, Grégoire, Florian Fuchs, Oleg Sklyar, Michael Boutros, and Wolfgang Huber (2010). EBImage–an R package for image processing with applications to cellular phenotypes. *Bioinformatics (Oxford, England)* 26, pp. 979–981. DOI: 10.1093/bioinformatics/btq046.

Pauli, Chantal, Benjamin D. Hopkins, Davide Prandi, Reid Shaw, Tarcisio Fedrizzi, Andrea Sboner, Verena Sailer, Michael Augello, Loredana Puca, Rachele Rosati, Terra J. McNary, Yelena Churakova, Cynthia Cheung, Joanna Triscott, David Pisapia, Rema Rao, Juan Miguel Mosquera, Brian Robinson, Bishoy M. Faltas, Brooke E. Emerling, Vijayakrishna K. Gadi, Brady Bernard, Olivier Elemento, Himisha Beltran, Francesca Demichelis, Christopher J. Kemp, Carla Grandori, Lewis C. Cantley, and Mark A. Rubin (2017). Personalized *In Vitro* and *In Vivo* Cancer Models to Guide Precision Medicine. *Cancer Discovery* 7, pp. 462–477. DOI: 10.1158/2159-8290.CD-16-1154.

Perakis, Alexis, Ali Gorji, Samriddhi Jain, Krishna Chaitanya, Simone Rizza, and Ender Konukoglu (2021). "Contrastive Learning of Single-Cell Phenotypic Representations for Treatment Classification". In: *Machine Learning in Medical Imaging: 12th International Workshop, MLMI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings*. Berlin, Heidelberg: Springer-Verlag, pp. 565–575. DOI: 10.1007/978-3-030-87589-3_58.

Polakis, Paul (2007). The many ways of Wnt in cancer. *Current Opinion in Genetics & Development* 17, pp. 45–51. DOI: 10.1016/j.gde.2006.12.007.

Puche, Juan E. and Inma Castilla-Cortázar (2012). Human conditions of insulin-like growth factor-I (IGF-I) deficiency. *Journal of Translational Medicine* 10, p. 224. DOI: 10.1186/1479-5876-10-224.

Ran, F. Ann, Patrick D. Hsu, Jason Wright, Vineeta Agarwala, David A. Scott, and Feng Zhang (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* 8, pp. 2281–2308. DOI: 10.1038/nprot.2013.143.

Ritchie, Matthew E., Belinda Phipson, Di Wu, Yifang Hu, Charity W. Law, Wei Shi, and Gordon K. Smyth (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* 43, e47. DOI: 10.1093/nar/gkv007.

Sachs, Norman, Joep de Ligt, Oded Kopper, Ewa Gogola, Gergana Bounova, Fleur Weeber, Anjali Vanita Balgobind, Karin Wind, Ana Gracanin, Harry Begthel, Jeroen Korving, Ruben van Boxtel, Alexandra Alves Duarte, Daphne Lelieveld, Arne van Hoeck, Robert Frans Ernst, Francis Blokzijl, Isaac Johannes Nijman, Marlous Hoogstraat, Marieke van de Ven, David Anthony Egan, Vittoria Zinzalla, Jurgen Moll, Sylvia Fernandez Boj, Emile Eugene Voest, Lodewyk Wessels, Paul Joannes van Diest, Sven Rottenberg, Robert Gerhardus Jacob Vries, Edwin Cuppen, and Hans Clevers (2018). A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. *Cell* 172, 373–386.e10. DOI: 10.1016/j.cell.2017.11.010.

Sakai, Eri, Mizuho Nakayama, Hiroko Oshima, Yuta Kouyama, Atsushi Niida, Satoshi Fujii, Atsushi Ochiai, Keiichi I. Nakayama, Koshi Mimori, Yutaka Suzuki, Chang Pyo Hong, Chan-Young Ock, Seong-Jin Kim, and Masanobu Oshima (2018). Combined Mutation of Apc, Kras, and Tgfbr2 Effectively Drives Metastasis of Intestinal Cancer. *Cancer Research* 78, pp. 1334–1346. DOI: 10.1158/0008-5472.CAN-17-3303.

Sansom, Owen J., Valerie S. Meniel, Vanesa Muncan, Toby J. Phesse, Julie A. Wilkins, Karen R. Reed, J. Keith Vass, Dimitris Athineos, Hans Clevers, and Alan R. Clarke (2007). Myc deletion rescues Apc deficiency in the small intestine. *Nature* 446, pp. 676–679. DOI: 10.1038/nature05674.

Sasaki, Nobuo, Norman Sachs, Kay Wiebrands, Saskia I. J. Ellenbroek, Arianna Fumagalli, Anna Lyubimova, Harry Begthel, Maaike van den Born, Johan H. van Es, Wouter R. Karthaus, Vivian S. W. Li, Carmen López-Iglesias, Peter J. Peters, Jacco van Rheenen, Alexander van Oudenaarden, and Hans Clevers (2016). Reg4+ deep crypt secretory cells function as epithelial niche for Lgr5+ stem cells in colon. *Proceedings of the National Academy of Sciences of the United States of America* 113, E5399–5407. DOI: 10.1073/pnas.1607327113.

Sato, Toshiro and Hans Clevers (2013). Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. *Science (New York, N.Y.)* 340, pp. 1190–1194. DOI: 10.1126/science.1234852.

Sato, Toshiro, Daniel E. Stange, Marc Ferrante, Robert G. J. Vries, Johan H. Van Es, Stieneke Van den Brink, Winan J. Van Houdt, Apollo Pronk, Joost Van Gorp, Peter D. Siersema, and Hans Clevers (2011). Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* 141, pp. 1762–1772. DOI: 10.1053/j.gastro.2011.07.050.

Sato, Toshiro, Robert G. Vries, Hugo J. Snippert, Marc van de Wetering, Nick Barker, Daniel E. Stange, Johan H. van Es, Arie Abo, Pekka Kujala, Peter J. Peters, and Hans Clevers (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* 459, pp. 262–265. DOI: 10.1038/nature07935.

Satoh, Kiyotoshi, Shinichi Yachida, Masahiro Sugimoto, Minoru Oshima, Toshitaka Nakagawa, Shintaro Akamoto, Sho Tabata, Kaori Saitoh, Keiko Kato, Saya Sato, Kaori Igarashi, Yumi Aizawa, Rie Kajino-Sakamoto, Yasushi Kojima, Teruaki Fujishita, Ayame Enomoto, Akiyoshi Hirayama, Takamasa Ishikawa, Makoto Mark Taketo, Yoshio Kushida, Reiji Haba, Keiichi Okano, Masaru Tomita, Yasuyuki Suzuki, Shinji Fukuda, Masahiro Aoki, and Tomoyoshi Soga (2017). Global metabolic reprogramming of colorectal cancer occurs at adenoma stage and is induced by MYC. *Proceedings of the National Academy of Sciences of the United States of America* 114, E7697–E7706. DOI: 10.1073/pnas.1710366114.

Scannell, Jack W. and Jim Bosley (2016). When Quality Beats Quantity: Decision Theory, Drug Discovery, and the Reproducibility Crisis. *PLOS ONE* 11, e0147215. DOI: 10.1371/journal.pone.0147215.

Schütte, Moritz, Thomas Risch, Nilofar Abdavi-Azar, Karsten Boehnke, Dirk Schumacher, Marlen Keil, Reha Yildiriman, Christine Jandrasits, Tatiana Borodina, Vyacheslav Amstislavskiy, Catherine L. Worth, Caroline Schweiger, Sandra Liebs, Martin Lange, Hans-Jörg Warnatz, Lee M. Butcher, James E. Barrett, Marc Sultan, Christoph Wierling, Nicole Golob-Schwarzl, Sigurd Lax, Stefan Uranitsch, Michael Becker, Yvonne Welte, Joseph Lewis Regan, Maxine Silvestrov, Inge Kehler, Alberto Fusi, Thomas Kessler, Ralf Herwig, Ulf Landegren, Dirk Wienke, Mats Nilsson, Juan A. Velasco, Pilar Garin-Chesa, Christoph Reinhard, Stephan Beck, Reinhold Schäfer, Christian R. A. Regenbrecht, David Henderson, Bodo Lange, Johannes Haybaeck, Ulrich Keilholz, Jens Hoffmann, Hans Lehrach, and Marie-Laure Yaspo (2017). Molecular dissection of colorectal cancer in pre-clinical models identifies biomarkers predicting sensitivity to EGFR inhibitors. *Nature Communications* 8, p. 14262. DOI: 10.1038/ncomms14262.

Serra, Denise, Urs Mayr, Andrea Boni, Ilya Lukonin, Markus Rempfler, Ludivine Challet Meylan, Michael B. Stadler, Petr Strnad, Panagiotis Papasaikas, Dario Vischi, Annick Waldt, Guglielmo Roma, and Prisca Liberali (2019). Self-organization and symmetry breaking in intestinal organoid development. *Nature* 569, pp. 66–72. DOI: 10.1038/s41586-019-1146-y.

Sharma, R. P. and V. L. Chopra (1976). Effect of the Wingless (wg1) mutation on wing and haltere development in Drosophila melanogaster. *Developmental Biology* 48, pp. 461–465. DOI: 10.1016/0012-1606(76)90108-1.

Sharma, Sreenath V., Diana Y. Lee, Bihua Li, Margaret P. Quinlan, Fumiyuki Takahashi, Shyamala Maheswaran, Ultan McDermott, Nancy Azizian, Lee Zou, Michael A. Fischbach, Kwok-Kin Wong, Kathleyn Brandstetter, Ben Wittner, Sridhar Ramaswamy, Marie Classon, and Jeff Settleman (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* 141, pp. 69–80. DOI: 10.1016/j.cell.2010.02.027.

Smirnov, Petr, Zhaleh Safikhani, Nehme El-Hachem, Dong Wang, Adrian She, Catharina Olsen, Mark Freeman, Heather Selby, Deena M. A. Gendoo, Patrick Grossmann, Andrew H. Beck, Hugo J. W. L. Aerts, Mathieu Lupien, Anna Goldenberg, and Benjamin Haibe-Kains (2016). PharmacoGx: an R package for analysis of large pharmacogenomic datasets. *Bioinformatics (Oxford, England)* 32, pp. 1244–1246. DOI: 10.1093/bioinformatics/btv723.

Smit, Wouter Laurentius, Claudia Nanette Spaan, Ruben Johannes de Boer, Prashanthi Ramesh, Tânia Martins Garcia, Bartolomeus Joannes Meijer, Jacqueline Ludovicus Maria Vermeulen, Marco Lezzerini, Alyson Winfried MacInnes, Jan Koster, Jan Paul Medema, Gijs Robert van den Brink, Vanesa Muncan, and Jarom Heijmans (2020). Driver mutations of the adenoma-carcinoma sequence govern the intestinal epithelial global translational capacity. *Proceedings of the National Academy of Sciences of the United States of America* 117, pp. 25560–25570. DOI: 10.1073/pnas.1912772117.

Snippert, Hugo J., Laurens G. van der Flier, Toshiro Sato, Johan H. van Es, Maaike van den Born, Carla Kroon-Veenboer, Nick Barker, Allon M. Klein, Jacco van Rheenen, Benjamin D. Simons, and Hans Clevers (2010). Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* 143, pp. 134–144. DOI: 10.1016/j.cell.2010.09.016.

Stambolic, V., L. Ruel, and J. R. Woodgett (1996). Lithium inhibits glycogen synthase kinase-3 activity and mimics wingless signalling in intact cells. *Current biology: CB* 6, pp. 1664–1668. DOI: 10.1016/s0960-9822(02)70790-2.

Stamos, Jennifer L. and William I. Weis (2013). The β-Catenin Destruction Complex. *Cold Spring Harbor Perspectives in Biology* 5. DOI: 10.1101/cshperspect.a007898.

Stein-O'Brien, Genevieve L., Raman Arora, Aedin C. Culhane, Alexander V. Favorov, Lana X. Garmire, Casey S. Greene, Loyal A. Goff, Yifeng Li, Aloune Ngom, Michael F. Ochs, Yanxun Xu, and Elana J. Fertig (2018). Enter the Matrix: Factorization Uncovers Knowledge from Omics. *Trends in Genetics* 34, pp. 790–805. DOI: 10.1016/j.tig.2018.07.003.

Stintzing, S., P. Wirapati, H.-J. Lenz, D. Neureiter, L. Fischer von Weikersthal, T. Decker, A. Kiani, F. Kaiser, S. Al-Batran, T. Heintges, C. Lerchenmüller, C. Kahl, G. Seipelt, F. Kullmann, M. Moehler, W. Scheithauer, S. Held, D. P. Modest, A. Jung, T. Kirchner, D. Aderka, S. Tejpar, and V. Heinemann (2019). 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* 30, pp. 1796–1803. DOI: 10.1093/annonc/mdz387.

Stolz, Ailine, Kim Neufeld, Norman Ertych, and Holger Bastians (2015). Wnt-mediated protein stabilization ensures proper mitotic microtubule assembly and chromosome segregation. *EMBO reports* 16, pp. 490–499. DOI: 10.15252/embr.201439410.

Subramanian, Aravind, Rajiv Narayan, Steven M. Corsello, David D. Peck, Ted E. Natoli, Xiaodong Lu, Joshua Gould, John F. Davis, Andrew A. Tubelli, Jacob K. Asiedu, David L. Lahr, Jodi E. Hirschman, Zihan Liu, Melanie Donahue, Bina Julian, Mariya Khan, David Wadden, Ian C. Smith, Daniel Lam, Arthur Liberzon, Courtney Toder, Mukta Bagul, Marek Orzechowski, Oana M. Enache, Federica Piccioni, Sarah A. Johnson, Nicholas J. Lyons, Alice H. Berger, Alykhan F. Shamji, Angela N. Brooks, Anita Vrcic, Corey Flynn, Jacqueline Rosains, David Y. Takeda, Roger Hu, Desiree Davison, Justin Lamb, Kristin Ardlie, Larson Hogstrom, Peyton Greenside, Nathanael S. Gray, Paul A. Clemons, Serena Silver, Xiaoyun Wu, Wen-Ning Zhao, Willis Read-Button, Xiaohua Wu, Stephen J. Haggarty, Lucienne V. Ronco, Jesse S. Boehm, Stuart L. Schreiber, John G. Doench, Joshua A. Bittker, David E. Root, Bang Wong, and Todd R. Golub (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* 171, 1437–1452.e17. DOI: 10.1016/j.cell.2017.10.049.

Sun, Hongzhi, Xiao Tu, and Renato Baserga (2006). A mechanism for cell size regulation by the insulin and insulin-like growth factor-I receptors. *Cancer Research* 66, pp. 11106–11109. DOI: 10.1158/0008-5472.CAN-06-2641.

Sung, Hyuna, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians* 71, pp. 209–249. DOI: 10.3322/caac.21660.

Sveen, A., C. Cremolini, and R. Dienstmann (2019). Predictive modeling in colorectal cancer: time to move beyond consensus molecular subtypes. *Annals of Oncology* 30, pp. 1682–1685. DOI: 10.1093/annonc/mdz412.

Tetteh, Paul W., Onur Basak, Henner F. Farin, Kay Wiebrands, Kai Kretzschmar, Harry Begthel, Maaike van den Born, Jeroen Korving, Frederic de Sauvage, Johan H. van Es, Alexander van Oudenaarden, and Hans

Clevers (2016). Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters. *Cell Stem Cell* 18, pp. 203–213. DOI: 10.1016/j.stem.2016.01.001.

Tirinato, Luca, Carlo Liberale, Simone Di Franco, Patrizio Candeloro, Antonina Benfante, Rosanna La Rocca, Lisette Potze, Roberto Marotta, Roberta Ruffilli, Vijayakumar P Rajamanickam, Mario Malerba, Francesco De Angelis, Andrea Falqui, Ennio Carbone, Matilde Todaro, Jan Paul Medema, Giorgio Stassi, and Enzo Di Fabrizio (2015). Lipid Droplets: A New Player in Colorectal Cancer Stem Cells Unveiled by Spectroscopic Imaging. *Stem Cells (Dayton, Ohio)* 33, pp. 35–44. DOI: 10.1002/stem.1837.

Van Cutsem, E., A. Cervantes, R. Adam, A. Sobrero, J. H. Van Krieken, D. Aderka, E. Aranda Aguilar, A. Bardelli, A. Benson, G. Bodoky, F. Ciardiello, A. D'Hoore, E. Diaz-Rubio, J.-Y. Douillard, M. Ducreux, A. Falcone, A. Grothey, T. Gruenberger, K. Haustermans, V. Heinemann, P. Hoff, C.-H. Köhne, R. Labianca, P. Laurent-Puig, B. Ma, T. Maughan, K. Muro, N. Normanno, P. Österlund, W. J. G. Oyen, D. Papamichael, G. Pentheroudakis, P. Pfeiffer, T. J. Price, C. Punt, J. Ricke, A. Roth, R. Salazar, W. Scheithauer, H. J. Schmoll, J. Tabernero, J. Taïeb, S. Tejpar, H. Wasan, T. Yoshino, A. Zaanan, and D. Arnold (2016). ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Annals of Oncology: Official Journal of the European Society for Medical Oncology* 27, pp. 1386–1422. DOI: 10.1093/annonc/mdw235.

Vincent, Fabien, Paula Loria, Marko Pregel, Robert Stanton, Linda Kitching, Karl Nocka, Regis Doyonnas, Claire Steppan, Adam Gilbert, Thomas Schroeter, and Marie-Claire Peakman (2015). Developing predictive assays: the phenotypic screening "rule of 3". *Science Translational Medicine* 7, 293ps15. DOI: 10.1126/scitranslmed.aab1201.

Virtanen, Seppo, Arto Klami, Suleiman Khan, and Samuel Kaski (2012). "Bayesian Group Factor Analysis". en. In: *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*. ISSN: 1938-7228. PMLR, pp. 1269–1277.

Vlachogiannis, Georgios, Somaieh Hedayat, Alexandra Vatsiou, Yann Jamin, Javier Fernández-Mateos, Khurum Khan, Andrea Lampis, Katherine Eason, Ian Huntingford, Rosemary Burke, Mihaela Rata, Dow-Mu Koh, Nina Tunariu, David Collins, Sanna Hulkki-Wilson, Chanthirika Ragulan, Inmaculada Spiteri, Sing Yu Moorcraft, Ian Chau, Sheela Rao, David Watkins, Nicos Fotiadis, Maria Bali, Mahnaz Darvish-Damavandi, Hazel Lote, Zakaria Eltahir, Elizabeth C. Smyth, Ruwaida Begum, Paul A. Clarke, Jens C. Hahne, Mitchell Dowsett, Johann de Bono, Paul Workman, Anguraj Sadanandam, Matteo Fassan, Owen J. Sansom, Suzanne Eccles, Naureen Starling, Chiara Braconi, Andrea Sottoriva, Simon P. Robinson, David Cunningham, and Nicola Valeri (2018). Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science (New York, N.Y.)* 359, pp. 920–926. DOI: 10.1126/science.aao2774.

Vogelstein, B., E. R. Fearon, S. R. Hamilton, S. E. Kern, A. C. Preisinger, M. Leppert, Y. Nakamura, R. White, A. M. Smits, and J. L. Bos (1988). Genetic alterations during colorectal-tumor development. *The New England Journal of Medicine* 319, pp. 525–532. DOI: 10.1056/NEJM198809013190901.

Wensink, G. Emerens, Sjoerd G. Elias, Jasper Mullenders, Miriam Koopman, Sylvia F. Boj, Onno W. Kranenburg, and Jeanine M. L. Roodhart (2021). Patient-derived organoids as a predictive biomarker for treatment response in cancer patients. *npj Precision Oncology* 5, pp. 1–13. DOI: 10.1038/s41698-021-00168-1.

Wetering, Marc van de, Hayley E. Francies, Joshua M. Francis, Gergana Bounova, Francesco Iorio, Apollo Pronk, Winan van Houdt, Joost van Gorp, Amaro Taylor-Weiner, Lennart Kester, Anne McLaren-Douglas, Joyce Blokker, Sridevi Jaksani, Sina Bartfeld, Richard Volckman, Peter van Sluis, Vivian S. W. Li, Sara Seepo, Chandra Sekhar Pedamallu, Kristian Cibulskis, Scott L. Carter, Aaron McKenna, Michael S. Lawrence, Lee Lichtenstein, Chip Stewart, Jan Koster, Rogier Versteeg, Alexander van Oudenaarden, Julio Saez-Rodriguez, Robert G. J. Vries, Gad Getz, Lodewyk Wessels, Michael R. Stratton, Ultan McDermott, Matthew Meyerson, Mathew J. Garnett, and Hans Clevers (2015). Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* 161, pp. 933–945. DOI: 10.1016/j.cell.2015.03.053.

Yao, Ye, Xiaoya Xu, Lifeng Yang, Ji Zhu, Juefeng Wan, Lijun Shen, Fan Xia, Guoxiang Fu, Yun Deng, Mengxue Pan, Qiang Guo, Xiaoxue Gao, Yuanchuang Li, Xinxin Rao, Yi Zhou, Liping Liang, Yaqi Wang, Jing Zhang, Hui Zhang, Guichao Li, Lixing Zhang, Junjie Peng, Sanjun Cai, Chen Hu, Jianjun Gao, Hans Clevers, Zhen Zhang, and Guoqiang Hua (2020). Patient-Derived Organoids Predict Chemoradiation Responses of Locally Advanced Rectal Cancer. *Cell Stem Cell* 26, 17–26.e6. DOI: 10.1016/j.stem.2019.10.010.

Yoon, Sang-Oh, Sejeong Shin, Florian A. Karreth, Gwen R. Buel, Mark P. Jedrychowski, David R. Plas, Shoukat Dedhar, Steven P. Gygi, Philippe P. Roux, Noah Dephoure, and John Blenis (2017). Focal Adhesion- and IGF1R-Dependent Survival and Migratory Pathways Mediate Tumor Resistance to mTORC1/2 Inhibition. *Molecular Cell* 67, 512–527.e4. DOI: 10.1016/j.molcel.2017.06.033.

Yu, Guangchuang, Li-Gen Wang, Yanyan Han, and Qing-Yu He (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS : a Journal of Integrative Biology* 16, pp. 284–287. DOI: 10.1089/omi.2011.0118.

Zhan, T., N. Rindtorff, and M. Boutros (2017). Wnt signaling in cancer. *Oncogene* 36, pp. 1461–1473. DOI: 10.1038/onc.2016.304.

Zhan, Tianzuo, Giulia Ambrosi, Anna Maxi Wandmacher, Benedikt Rauscher, Johannes Betge, Niklas Rindtorff, Ragna S. Häussler, Isabel Hinsenkamp, Leonhard Bamberg, Bernd Hessling, Karin Müller-Decker, Gerrit Erdmann, Elke Burgermeister, Matthias P. Ebert, and Michael Boutros (2019). MEK inhibitors activate Wnt signalling and induce stem cell plasticity in colorectal cancer. *Nature Communications* 10, p. 2197. DOI: 10.1038/s41467-019-09898-0.

Zhang, Jun and Shane T. Mueller (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika* 70, pp. 203–212. DOI: 10.1007/s11336-003-1119-8.

Zhang, Lixing, Mohammad Adileh, Maria Laura Martin, Stefan Klingler, Julie White, Xiaojing Ma, Louise R. Howe, Anthony M. C. Brown, and Richard Kolesnick (2017). Establishing estrogen-responsive mouse mammary organoids from single Lgr5+ cells. *Cellular Signalling* 29, pp. 41–51. DOI: 10.1016/j.cellsig.2016.08.001.

Zhang, Wei and Hui Tu Liu (2002). MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Research* 12, pp. 9–18. DOI: 10.1038/sj.cr.7290105.

Zhang, Xiaofei, Arne H. Smits, Gabrielle BA van Tilburg, Huib Ovaa, Wolfgang Huber, and Michiel Vermeulen (2018). Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nature Protocols* 13, pp. 530–550. DOI: 10.1038/nprot.2017.147.

Zhong, H., C. Fazenbaker, C. Chen, S. Breen, J. Huang, X. Yao, P. Ren, Y. Yao, R. Herbst, and R. E. Hollingsworth (2017). Overproduction of IGF-2 drives a subset of colorectal cancer cells, which specifically respond to an anti-IGF therapeutic antibody and combination therapies. *Oncogene* 36, pp. 797–806. DOI: 10.1038/onc.2016.248.

Eigenanteil

This thesis comprises two parts, focused on patient-derived organoids and engineered mouse colon organoids.

The parts of this thesis focused on patient-derived organoids were conducted in collaboration with Johannes Betge (Physician Scientist, 2nd medical clinic, Mannheim) and Jan Sauer (PhD Student, Computational Genome Biology Group, DKFZ, Heidelberg) with support by Benedikt Rauscher (PhD Student, Signaling and functional Genomics Group, DKFZ, Heidelberg).

Johannes Betge led the multi-center procurement and oversaw the isolation, amplicon sequencing and transcript expression measurements of patient-derived colorectal cancer organoids. He also oversaw the RT-qPCR (Figure 3.18 d) and Western Blot (Figure 3.15 g) based validation of MEK and mTOR inhibitors in patient-derived organoids, respectively. High-throughput experiments were done with technical support from Johannes Betge.

Jan Sauer implemented the image analysis process, comprising the compression of raw images, the maximum contrast projection method, the segmentation of organoids, the extraction of features as well as the estimation of treatment effect vectors through logistic regression.

Benedikt Rauscher advised and implemented a version of the transcript expression analysis.

Erica Valentini performed the analysis of amplicon sequencing data.

I conducted all experiments outlined in the main part of this dissertation, unless otherwise specified. Data analysis and visualization was performed by me, unless otherwise specified.

The parts of this thesis focused on engineered mouse colon organoids were conducted independently of the collaboration outlined above. The *LSL-Kras^{G12D} CreERT2* mouse was a gift from Claudia Scholl.

Parts of this dissertation have already been published in the following article:

Johannes Betge*, **Niklas Rindtorff***, Jan Sauer*, Benedikt Rauscher*, Clara Dingert, Haristi Gaitantzi, Frank Herweck, Kauthar Srour-Mhanna, Thilo Miersch, Erica Valentini, Veronika Hauber, Tobias Gutting, Larissa Frank, Sebastian Belle, Timo Gaiser, Inga Buchholz, Ralf Jesenofsky, Nicolai Härtel, Tianzuo Zhan, Bernd Fischer, Katja Breitkopf-Heinlein, Elke Burgermeister, Matthias P. Ebert, Michael Boutros, The drug-induced phenotypic landscape of colorectal cancer organoids. Nature Communications; 3135 (2022). https://doi.org/10.1038/s41467-022-30722-9 ***shared first authorship**

The publication corresponds to all experimental results presented in the first chapter of the thesis results with small modifications including the addition of (Figure 3.5). The publication also covers smaller parts of the introduction, as well as the discussion, related to patient-derived organoids. My personal contribution to the publication comprised the performance of all experiments (most of which were completed together with Johannes Betge) except for the amplicon sequencing of organoid samples (conducted by the laboratory engineer T. Miersch under supervision of Johannes Betge), as well as the validating RT-qPCRs, and Western Blots (conducted by the laboratory engineer K. Kaiser under supervision of Johannes Betge). Laboratory engineers, A. Falzone, and A. Kerner supported with the establishment and processing of patient-derived organoids. The underlying data analysis and visualizations for the main figures except for Figure 3, Figure 5i, and Figure 6g were created by me. I contributed to the complete writing process and the majority of the results and introduction section were written by me, as well as the discussion on the limits of using multi-omics factor analysis.

Throughout the period in which this dissertation was created, a set of additional publications was contributed to:

- **Niklas Rindtorff**, MingYu Lu, Nisarg Patel, Huahua Zheng, Alexander D'Amour, A Biologically Plausible Benchmark for Contextual Bandit Algorithms in Precision Oncology Using in vitro Data. Arxiv (2019) https://doi.org/10.48550/arXiv.1911.04389
- Tianzuo Zhan*, **Niklas Rindtorff***, Michael Boutros, Wnt signaling in cancer. Oncogene; 1461–1473 (2017). https://doi.org/10.1038/onc.2016.304
- Tianzuo Zhan*, **Niklas Rindtorff***, Johannes Betge, Matthias P. Ebert, Michael Boutros, CRISPR/Cas9 for cancer research and therapy. Seminars in Cancer Biology; 106-119 (2019). https://doi.org/10.1016/j.semcancer.2018.04.001
- Leonhard Valentin Bamberg, Florian Heigwer, Anna Maxi Wandmacher, Ambika Singh, Johannes Betge, **Niklas Rindtorff**, Johannes Werner, Julia Josten, Olga Valerievna Skabkina, Isabel Hinsenkamp, Gerrit Erdmann, Christoph Röcken, Matthias P Ebert, Elke Burgermeister, Tianzuo Zhan, Michael Boutros, Targeting euchromatic histone lysine methyltransferases sensitizes colorectal cancer to histone deacetylase inhibitors. Int. J. Cancer.; 151:1586–1601 (2022). https://doi.org/10.1002/ijc.34155
- Tianzuo Zhan, Giulia Ambrosi, Anna Maxi Wandmacher, Benedikt Rauscher, Johannes Betge, **Niklas Rindtorff**, Ragna S. Häussler, Isabel Hinsenkamp, Leonhard Bamberg, Bernd Hessling, Karin Müller-Decker, Gerrit Erdmann, Elke Burgermeister, Matthias P. Ebert, Michael Boutros, MEK inhibitors activate Wnt signalling and induce stem cell plasticity in colorectal cancer. Nature Communications; 2197 (2019). https://doi.org/10.1038/s41467-019-09898-0

*__shared first authorship__

A manuscript presenting the results generated during the image-based profiling of engineered mouse colon organoid is in preparation at the time of this thesis submission.

Curriculum Vitae

Persönliche Daten

Name: Niklas Timon Rindtorff
Geburtsdatum: 11.04.1995
Geburtsort: Tübingen

Schulischer und universitärer Werdegang

06/2021	Erteilung der ärztlichen Approbation, Ende des Studiums an der Ruprecht-Karls-Universität Heidelberg
06/2021	3. Abschnitt der Ärztlichen Prüfung (1,0)
04/2021	2. Abschnitt der Ärztlichen Prüfung (Verschoben durch Verordnung zur Abweichung von der Approbationsordnung) (2,0)
04/2020	Praktisches Jahr in den Abteilungen Viszeralchirurgie, klinische Pharmakologie (beide Universitätsklinikum Heidelberg) und Innere Medizin (Vincentius Krankenhaus Heidelberg)
06/2019	Master of Biomedical Informatics, Harvard Medical School (4.0 GPA)
08/2018 - 08/2019	Einjähriges Fulbright Stipendium, Harvard Medical School, Boston, MA, USA
Seit 09/2016	Medizinische Doktorarbeit in der Arbeitsgruppe Signalwege und funktionelle Genomik, DKFZ Heidelberg
07/2015	1. Abschnitt der Ärztlichen Prüfung (1,0)
09/2013	Beginn Studium der Humanmedizin, Ruprecht-Karls-Universität Heidelberg
09/2011 - 06/2013	Lise-Meitner-Gymnasium Böblingen, Abitur (1,0)
08/2010 - 06/2011	Auslandsjahr an der Newfane Senior Highschool, Newfane, NY, USA
09/2006 - 08/2010	Schönbuch Gymnasium Holzgerlingen
09/2001 - 08/2005	Grundschule Weil im Schönbuch

Acknowledgements

I would like to thank Prof. Dr. Michael Boutros for providing the opportunity to work on this ambitious thesis project. I am grateful for the regular feedback and all the support received during my work.

I appreciated the support from Dr. Tianzuo Zhan who supervised me early during my work in the laboratory. Without his support and encouragement to independently pursue the establishment of organoid models, this project would not have developed into its current form.

I am truly grateful to Dr. Johannes Betge who was a wonderful colleague and scientific partner. He joined the laboratory at a time when the isolation and image-based profiling was still at a prototype stage. His support enabled the establishment and profiling of patient derived organoid models. With him, I found a great colleague to pursue this project. We spent many very long evenings in the cell culture laboratory isolating and expanding organoid models. Without his support, the presented work would not have been possible.

I would like to thank Dr. Jan Sauer for his work on image-analysis of 3D organoid cultures. My thanks also go to Dr. Benedikt Rauscher for thoughtful advice throughout this project and an initial analysis of the patient derived organoid transcript expression data.

I am especially thankful to Dr. Kim Boonekamp, who supported me during the completion of this thesis as well as in the discussion of mouse organoid profiling results with very thoughtful comments. Without her scientific and personal support, this thesis would not have been shaped into this current form.

My sincere appreciation goes to the engineering team of the Signaling and Functional Genomics group, especially Thilo Miersch whose expertise in high-throughput liquid handling was instrumental to making the image-based profiling method work for organoid models at scale.

Lastly, I extend my heartfelt thanks to all members of the Signaling and Functional Genomics group. The laboratory has been an ideal place for my scientific training and first independent steps. It was a pleasure to work with you.

I am dedicating this thesis to my family.

Eidestaatliche Versicherung

1. Bei der eingereichten Dissertation zu dem Thema "Image-based profiling of colorectal cancer organoid models" handelt es sich um meine eigenständig erbrachte Leistung.
2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.
3. Die Arbeit oder Teile davon habe ich bisher nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.
4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.
5. Die Bedeutung der eidessstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidessstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Berlin, den 18.12.2023 Niklas Rindtorff