Dissertation

submitted to the
Combined Faculty of Natural Sciences and Mathematics of
the Ruperto Carola University Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
B.Tech Ashna Alladin
born in Mumbai, India

Oral Examination : 24th October, 2019

Studying the origins of primary tumours and residual disease in breast cancer

Referees :

Prof. Dr. Darren Gilmour

Prof. Dr. Ana Martin-Villalba

To Areeb,
now and always

ACKNOWLEDGEMENTS

Many people and various places have made this thesis and the research behind it possible, and I would like to express my deepest gratitude to each and every one of them. **EMBL**, for being an inspiring environment, filled with dedicated and enthusiastic scientists, many of whom have helped me during my PhD. I am grateful for being accepted to pursue my PhD at EMBL and would like to thank the members of the EIPP Graduate Office (**Helke Hillebrand, Monika Lachner, Matija Grgurinovic, Carolina Carolina Garcia Sabate and Meriam Bezohra**). My deepest gratitude is also extended to **Reshma Surendran** from Human Resources and other members of the administrative departments at EMBL.

I owe my heartiest gratitude to **Martin Jechlinger** for giving me a place in his research group in the Cell Biology and Biophysics Unit. I was deeply inspired by his research projects and ideas and am especially grateful for his role in my exposure to innumerable techniques and my overall development as a scientist. My views on scientific method, translational research and group dynamics have been shaped by Martin over the course of the PhD and will guide me for the rest of my life.

As far as guidance goes, I want to take this opportunity to express my gratitude to members of my Thesis Advisory Committee: **Darren Gilmour**, **Ana Martin-Villaba** and **Kiran Raosaheb Patil**. I have benefited greatly from their scientific opinions and project discussions during my committee meetings over the years. I would also like to extend my gratitude to **Micheal Knop** for agreeing to be part of the Defence Committee.

No group of people have been more instrumental in helping me achieve my research goals than the past and present members of the **Jechlinger Group**. **Ksenija Radic** for being an incredible scientist, a wonderful human being and the best trainer and collaborator I have ever had the joy of working with. My deepest gratitude to you Ksenija for helping me every step along the way, from conducting experiments to careful perusal of my thesis. I will always value **Alex Shechter**'s and your contribution to my life and work in Heidelberg. **Marta Garcia Montero** for teaching me how to grow my first organoid and to sacrifice my first mouse, and for being there for whatever favour, personal and professional, that I needed over the last few years. I would also like to thank **Sylwia Gawrzak** for her guidance with experiments and valuable advice on all matters of the PhD; **Savannah Jade Jackson** for her work with the mice and for pensive morning discussions that set the mood for many of my PhD days; **Matthew Stephen Boucher** for his scientific input and epic humour; **Lucas Chaible** for being a vital collaborator and a dear friend during his term and of course for his supreme annual BBQs; **Miriam Valera-Alberni** and **Federico Villa** for all the science and laughter I shared with them during their short stays.

Many of my work hours as a PhD student have been spent at the **ALMF** core facility at EMBL and so I would like to thank **Sabine Reither** for all her expertise and help in guiding me along the light sheet microscopy road; **Stefan Terjung** for his help with all things microscopy; **Tobias Rasse** for his enthusiasm and scientific input. I would also like to thank the members of the Luxendo family (**Monika Loeschinger, Malte Wachsmuth, Soeren Kaufmann**) for their dedication towards facilitating all my imaging experiments. My greatest thanks to the Centre for Bioimage Analysis (**CBA**) and **Christian Tischer** for his brilliance in developing image analysis tools and dedication in helping me decipher umpteen terabytes of data to further my project.

I would like to express my gratitude to project collaborators throughout EMBL including **Katharina Zirngibl** for her work with the RNASeq data and **Jean-Karim Heriche** for his computational analysis and scientific input. Numerous other colleagues and core facilities have contributed my research and growth as a scientist – Laboratory Animal Resources (**Klaus Schmitt** and all the animal technicians); Genomics Core Facility (**Vladimir Benes, Jelena Pistolic**); Chemical Biology Core Facility (**Kerstin Putzker**); Centre for Biomolecular Network Analysis (**Matt Rogon and Bernd Klaus**) and Flow Cytometry Core Facility (**Malte Paulsen**). I would also like to thank **Rocio Sotillo** and her lab at the DKFZ for providing invaluable support, especially **Charles Yuanyuan Chen** for being patient and helpful while I persevered with trying to optimize immunohistochemistry stains. I would also like to thank **Robert Prevede**l and **Jan Ellenberg** for their time and suggestions on how to improve my manuscript.

No account of my gratitude would be complete without the inclusion of my friends and family who have supported me through all the collateral damage that happens when you pursue science whole-heartedly and with passion. **Sanjana Singh** has been the best roommate, bridesmaid and chai date to me for the last four years and I would like to thank her for many brilliant scientific discussions and of course, her dardnaak humour. This PhD would be a lot less fun if she hadn't walked into the lab and broken into a song innumerable times. Many thanks also to **Sebastian Brosig** for translating my thesis summary and for cheering me up in tissue culture when I really needed it. I would also like to thank all my friends and colleagues in the PhD community at EMBL for their help, advice and pursuit of good times.

I am grateful to my alma mater VIT University in Vellore, India for allowing me the chance to be inspired by excellent faculty members and supported by the best of friends (**Sakshi Dixit Ankita Budhraja, Ashwin Koneru**). My special thanks also to **Chetna Jai**n for being the best thing about my hometown and roots.

Finally, I would like to thank my husband **Areeb Jawed** for all he has done for me

for the better part of the last decade. Thank you Areeb, for travelling and exploring the world with me, teaching me to relax and appreciate the little things and greatly improving my standards of living and eating. You are a man of brilliance, love and laughter and I am grateful for you every day.

More than anyone, I owe my eternal gratitude to my **family**. My **Dad**, for his inadvertent lessons on life and the importance of being content. My father made me an independent entity in pursuit of brilliance and I will forever be grateful. My **Mom** for believing in me, motivating me and fuelling my ambition every day. Nothing and no one can ever compare to the light she brings into my life. My sisters **Azra and Zaara** for their incredible hearts and unwavering loyalty. **Shermin**, for all the love, laughter and good living. My grandparents and uncles and aunts and cousins for all the good memories that have made me who I am and given me the courage to pursue my dreams.

It really does take a village.

SUMMARY

Breast cancer is the leading cause of death in women worldwide and these deaths are mostly attributed to metastasis and tumour recurrence following initially successful therapy. Metastasis refers to the development of invasive disease, wherein malignant cells dissociate from primary tumours, infiltrating other organs and tissues to give rise to secondary outgrowths. Previously, metastasis was thought to be initiated in advanced tumours, but breast cancer cellsh with metastatic potential have now been shown to disseminate very early from the primary site via largely unknown mechanisms. These early interactions of tumour cells with their cellular micro-environment and normal neighbours also results in early tumour cell heterogeneity and must therefore be elucidated such that we can prevent metastatic spread in the patient situation and better treat the resulting heterogenous tumours. However, studying tumour initiation is not possible in patients because it happens on a cellular level not detectable by current technology. Tumour recurrence is another major cause of breast cancer related death and is believed to be caused by residual disease cells that survive initial therapy. These are a reservoir of refractory cells that can lay dormant for many years (sometimes decades) before resulting in relapse tumours. They are also difficult to obtain from human patients, since they are very few and cannot be detected easily, and thus their molecular mechanisms have not been fully explored.

In addition to the unavailability of human tissue, mouse models of breast cancer also fall short in helping us study early cancer initiation, because they allow oncogenic expression in all cells of the tissue instead of initiating cancer like in the human situation—one neoplastic transformed cell proliferating unchecked in a normal epithelium. To address this issue, we used primary organoids from an inducible mouse model of breast cancer and lentivirally transduced single cells within these organoids to express oncogenes. We further optimized parameters for long term imaging using light sheet microscopy and developed big data analysis pipelines that lead us to discern that single transformed cells had a lower chance at establishing tumorigenic foci, when compared to clusters of cells. Thus, we postulate a proximity-controlled signalling that is imperative to tumour initiation within epithelial tissues using the first ever in vitro stochastic breast tumorigenesis model system. This new stochastic tumorigenesis system can be further used to identify the molecular interactions in the early breast cancer cells.

Our group has already revealed distinct characteristics, such as dysregulated lipid metabolism, of the residual disease correlate obtained from an inducible mouse model. As survival mechanisms invoked by residual cells remain largely unknown, we analysed the dynamic transcriptome of regressing tumours at important timepoints during the establishment of residual disease. Key molecular players upregulated during regression – like c-Jun and BCL6 – were identified and the inflammatory arm of the Nf-kB cascade was found to be dysregulated among others. Further validation of these molecular targets as potentially synthetic lethal interactors remains to be performed so that they can be used to limit the residual disease reservoir and eventually tumour recurrence.

ZUSAMMENFASSUNG

Brustkrebs ist die häufigste Todesursache bei Frauen weltweit. Die auftretenden Todesfälle sind dabei hauptsächlich auf Metastasen und Brustkrebs im Rezidiv nach anfänglich erfolgreicher Therapie zurückzuführen. Unter Metastasen versteht man Tumorzellen die den Zellverbund des Primärtumors verlassen haben, und sich in weit entfernten Organen sowie Geweben des Körpers neu anzusiedeln. Lange wurde angenommen, dass besonders fortgeschrittene Tumore Metastasen bilden. Heute weiß man, dass sich Brustkrebszellen bereits frühzeitig und noch auf ungeklärte Weise vom Primärtumor lösen können. Die somit schon frühen Interaktionen von Krebszellen mit ihrer Umgebung sowie gesunden Nachbarzellen führen zu heterogenen Tumoren. Diese gilt es eingehend zu erforschen um Metastasen und daraus resultierende heterogene Tumore in Brustkrebspatienten besser behandeln zu können. Die frühzeitige Erkennung der Brustkrebsbildung ist in Patienten derzeit unmöglich, da sich anfangs nur wenige Tumorzellen bilden, welche zudem mit heutigen Techniken nicht zu identifizieren sind. Eine weitere Ursache für den Tod von Brustkrebspatienten ist das erneute ausbrechen bereits erfolgreich behandelter Tumore, sogenannte Rezidive. Rezidive sind auf wenige widerstandsfähige Tumorzellen zurückzuführen, welche während einer ersten Behandlung nicht vollständig abgestorben sind. Diese Zellen können in inaktivem Zustand über Jahre im Körper des Patienten überdauern und auch noch nach Jahrzenten einen weiteren Brustkrebs hervorrufen. Bisher sind solch inaktive Tumorzellen unmöglich aus bereits behandelten Patienten zu isolieren, da sie in geringer Zahl vorliegen und mit heutigen Techniken nicht zu identifiziert sind. Aus diesem Grund sind molekulare Grundlagen für Rezidive unzureichend erforscht.

Um Rezidive besser zu erforschen sind neben der Unverfügbarkeit menschlicher Tumorproben auch Brustkrebsmodelle aus Mäusen wenig Hilfreich. In Mausmodellen lassen sich Tumore derzeit nur in allen Zellen eines Gewebes auslösen. Mein Ziel hingegen ist, Brustkrebs durch nur eine Krebszelle in einem sonst gesunden Gewebe hervorzurufen, ähnlich wie es bei Brustkrebspatienten der Fall ist. Daher verwende ich primäre Organoide welche aus induzierbaren Brustkrebs Mausmodellen gewonnen wurden. Dabei wird eine einzelne Zelle dieses Organoids durch einen Lentivirus infiziert, wodurch entsprechende Krebs Gene aktiviert werden. Zudem habe ich die nötigen Parameter zur Langzeitbildgebung mittels Light Sheet Mikroskopie“ optimiert. Die zusätzliche Entwicklung von big data analysis pipelines“ erlaubte mir herauszufinden, dass einzelne Krebszellen weniger wahrscheinlich zu Tumorherden beitragen als ganze Gruppen von Krebszellen. Durch die erstmalige Verwendung eines stochastischen in-vitro Brustkrebsmodells konnte ich zeigen, dass eine Umgebungskontrollierte Signalübertragung unabdingbar ist um Brustkrebs hervorzurufen. Dieses neuartig stochastische Brustkrebsmodell kann zudem zur Identifizierung von makromolekularen Interaktionen in noch frühen Brustkrebszellen verwendet werden.

Durch induzierbare Brustkrebsmodelle in der Maus konnte unser Labor bereits charakteristische Eigenschaften von Rezidiven identifizieren. Eine Haupteigenschaft war dabei ein dysregulierter Lipid Metabolismus. Da Überlebensmechanismen von inaktiven aber

jederzeit rückfälligen Tumorzellen unbekannt sind, haben wir zu verschiedenen Zeitpunkten das Transkriptom jener Brustkrebszellen analysiert, die auf dem Weg der Inaktivierung waren. Dabei hat sich herausgestellt, dass Schlüsselproteine wie c-Jun und BCL6 während der Tumorregression vermehrt exprimiert wurden. Zudem war der inflammatorische Signalweg der Nf-kB Kaskade neben anderen Signalwegen dysreguliert. Die Validierung dieser Molekularen Targets als potentiell synthetisch letale Interaktoren steht noch aus. Sobald diese durchgeführt wurden, werden identifizierte Targets möglicherweise dazu beitragen inaktive Tumorzellen zu behandeln und ein erneutes Auftreten des Brustkrebses zu verhindern.

CONTENTS

1 Introduction **13**

1.1 Breast Cancer . . . 14

1.2 Mouse models of human carcinomas . . . 20

2 Unravelling stochastic tumourigenesis, one cell at a time **23**

2.1 Introduction . . . 24

2.2 Objectives . . . 27

2.3 Results . . . 28

2.3.1 Establishing a 3D *in vitro* model of stochastic breast tumourigenesis 28

2.3.2 Live organoid imaging reveals early tumour dynamics in 4D at single cell resolution . . . 32

2.3.3 Development of big-image data compatible analysis workflow allows tracking of cell lineages during early tumour formation . . . 36

2.3.4 Feature analysis of lentivirus-transduced organoids indicates that the origin of tumours may be in cell clusters and not in single transformation events . . . 42

2.4 Discussion . . . 47

2.5 Conclusions . . . 49

3 Quest for the Achilles' heel in breast cancer residual disease **51**

3.1 Introduction . . . 52

3.1.1 Minimal Residual Disease . . . 52

3.1.2 Synthetic Lethality . . . 59

3.2 Objectives . . . 61

3.3 Results . . . 62

3.3.1 Characterization of minimal residual disease *in vitro* using primary organoids . . . 62

3.3.2 RNA sequencing analysis of *in vitro* organoids during the establishment of MRD . . . 65
3.3.3 Characterization of minimal residual disease *in vivo* in mice . . . 75
3.3.4 Validation of selected targets for synthetic lethality during establishment of MRD . . . 79
3.3.5 Imaging tumour organoid regression using light sheet microscopy . . . 88
3.4 Discussion . . . 93
3.5 Conclusions . . . 98

4 Materials and Methods 99
4.1 Animal models . . . 102
4.1.1 Genotyping . . . 102
4.1.2 *In vitro* experimental design . . . 102
4.1.3 *In vivo* experimental design and harvest . . . 103
4.2 3D Cell culture . . . 104
4.2.1 Matrigel culture . . . 104
4.2.2 Organoid transduction . . . 105
4.2.3 Target validation pipeline . . . 105
4.3 Immunofluorescence staining . . . 108
4.4 Immunohistochemistry staining . . . 109
4.5 qPCR analysis . . . 110
4.6 Lentivirus cloning and production . . . 111
4.7 Selective Plane Illumination Microscopy . . . 112
4.7.1 Sample holder preparation and mounting . . . 112
4.7.2 Imaging configurations and conditions . . . 112
4.7.3 Big-image data analsysis workflow . . . 113
4.7.4 Computational feature analysis . . . 113
4.8 RNA Seq data analysis . . . 115
4.8.1 RNA extraction . . . 115
4.8.2 RNA Seq data analysis . . . 115
4.8.3 Clustering and Correlation analysis . . . 116
4.8.4 Gene enrichment analysis . . . 117
4.9 Microarray data analysis . . . 118
4.9.1 RNA extraction . . . 118
4.9.2 Microarray data analysis . . . 118

5 Supplementary Material 119

6 Bibliography 139

7 Lists of Abbreviations, Figures and Tables 157

CHAPTER 1

INTRODUCTION

1.1 Breast Cancer

Breast anatomy

Breasts, medically known as the mammary glands, are composed of two compartments - glandular and stromal. The stromal compartment of the mammary gland consists of the fat tissue and connective tissue that surrounds the glandular compartment. The glandular compartment, in turn, is made up of lobules and ducts. Lobules are milk-producing structures, and a system of ducts transports milk from the lobules to the nipple. Blood vessels along with the lymph nodes, immune cells and nerve endings are located throughout the breast (Figure 1.1) [1].

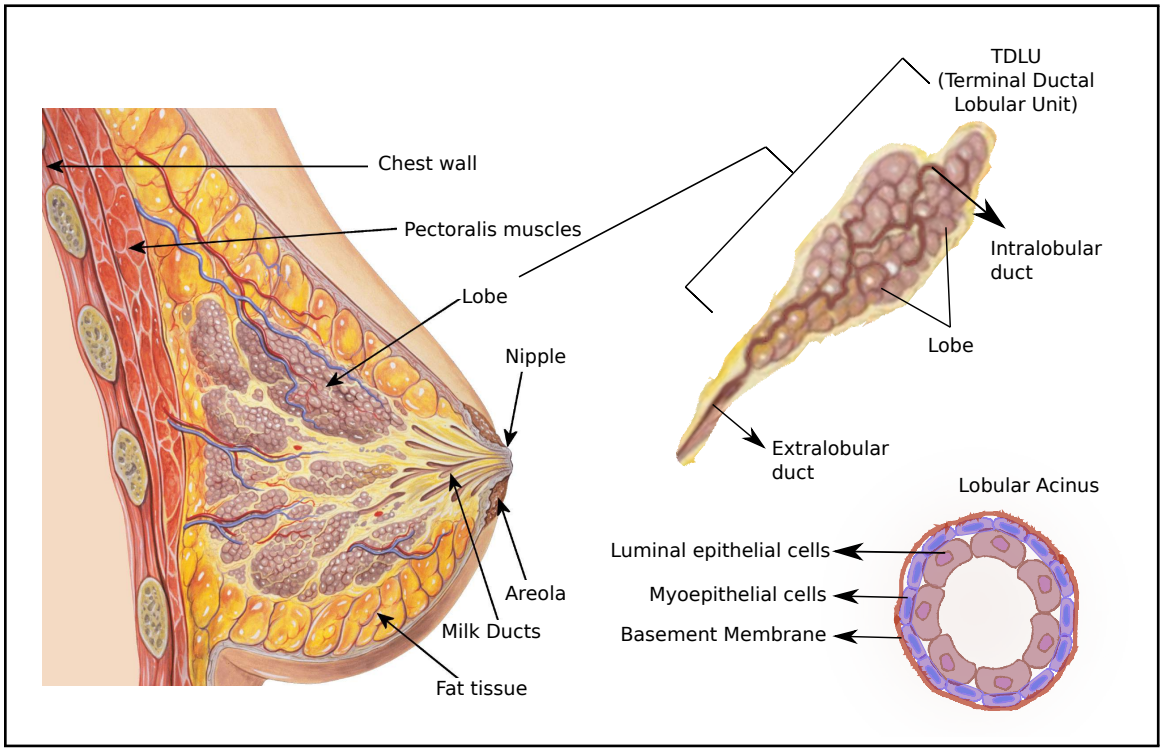


Figure 1.1: Human breast anatomy showing the glandular and stromal cell compartments of the mammary tissue. The functional unit of the mammary tissue is the TDLU that is composed of hundreds of acini- each of which consist of a double epithelial cell layer enclosing the lumen. Figure modified from [2]

Each breast has 12 to 20 sections, or lobes, that surround the nipple in a radial manner, like spokes on a wheel. The lobes are composed of lobules, the structural and functional units of the mammary gland also called terminal ductal lobular units (TLDUs) [3] [4]. The innermost layer of a TDLU is formed of polarized luminal epithelial cells. The lumen formed by these cells is where milk is secreted during lactation (in the lobules) and transported (in the ducts). The second layer of cells surrounding this lumen consists of myoepithelial (basal) cells [5]. These are divided from the surrounding stroma by a base-

ment membrane, which serves as an important barrier and dynamic structure composed of extra-cellular matrix proteins [6]. Breast cancer arises mostly from abnormal epithelial cells contained within the glandular epithelial compartment of the breast.

Breast carcinoma

Transformed epithelial cells in the breast proliferate and start to fill the lumen in the earlier stages of carcinoma known as the non-invasive stage. Ductal carcinoma *in situ* (DCIS) is confined to the milk ducts. It is the most common form of non-invasive breast cancer accounting for up to 80% of clinical cases. Lobular carcinoma *in situ* (LCIS) is confined to the milk-producing glands (Figure 1.2). Unlike DCIS, LCIS is not known to be a pre-malignant condition, but rather a marker that identifies women at an increased risk for invasive breast cancer [7]. As the disease progresses to invasive and metastatic cancer, cells breach the basement membrane and grow into the surrounding normal tissue and eventually home in distant organs [8].

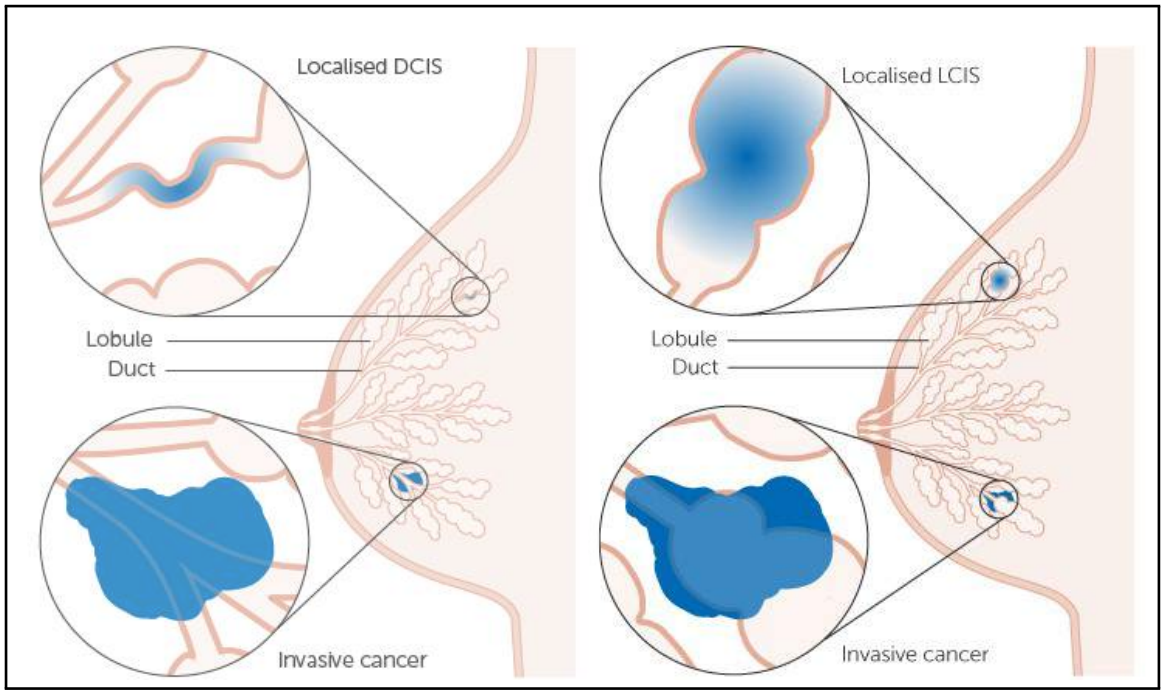


Figure 1.2: Ductal carcinoma (left panel) involves abnormal cell proliferation in the mammary gland ducts, while in lobular carcinoma outgrowths are seen in the lobules (right panel). When the transformed cells are confined to the duct/lobule it is called carcinoma *in situ*. Invasion of surrounding tissue by cancer cells in both cases is seen in metastatic disease. Figure modified from [9].

Breast carcinoma is the leading cause of death among women in most developed countries[10]. Last year, over 2 million new cases of breast cancer were reported in the world

and over 620,000 women succumbed to this disease and died [11]. For decades, we have been trying to limit and eradicate this disease and yet the fatalities continue to persist. This has, primarily, to do with the fact that cancer, irrespective of it's tissue of origin, is a complex collection of diseases which exhibit in biologically different entities and governed by intrinsic cell properties and system factors. Thus, it can have distinct clinical implications on an individual patient basis [12] [13] [14].

Molecular classification of Breast Cancer

Traditional clinico-pathological variables have been conventionally used for patient prognosis and management. These include factors like tumor size, tumor grade and nodal involvement, invasive behaviour and some standard immunohistochemistry (IHC) markers for hormone receptors [15] [16]. Our advances in the genetic profiling of tumors have proven that response to treatment is actually not determined by anatomical prognostic factors but rather by intrinsic molecular characteristics that can be defined by gene expression patterns of specific tumors [17][18].

As such, gene expression signatures are now commonly used to classify breast cancers [12] in five intrinsic sub-types with distinct clinical outcomes:

- Luminal A
- Luminal B
- HER2 over-expressed
- Basal
- Normal-like

Each of the five intrinsic sub-types is nicely mapped to an IHC-defined sub-type [19] [20] (Table 1.1) except for the normal-like tumors which share a similar IHC status with the Luminal A sub-type and are characterized by a normal breast tissue profiling [21]. In addition to this gene signature basis for classification of breast tumors, information such as lncRNA and epigenetic data plays critical roles in tumor progression and classification, providing novel perspectives on breast tumor sub-typing.

The luminal-like tumors are the most common sub-types among breast cancer. They express hormone receptors, with expression profiles reminiscent of the luminal epithelial component of the breast [21]. Luminal A tumors have higher expression of Estrogen Receptor(ER)-related genes and lower expression of proliferative genes than luminal B cancers [19] [22]. Luminal A tumors are commonly treated with endocrine therapy, while luminal B tumors, which are more proliferative, with a combination of anti-proliferative

Molecular Sub-type	Prevalance	IHC status	Prognosis
Luminal A	23.7%	ER+, PR+, HER2 -, KI67 -	Good
Luminal B	52.8%	ER+, PR+, HER2+/-, KI67+	Intermediate/Poor
HER2 over-expressing	11.2%	ER -, P R-, HER2+	Poor
Basal-like	12.3%	ER -, PR -, HER2 -, basal marker+	Poor
Normal-like	7.8%	ER+, PR+, HER2 -, KI67 -	Intermediate

Table 1.1: Classification and prevalence of the five breast cancer intrinsic sub-types with associated clinical parameters

and hormonal treatment.

Basal tumors are composed of Estrogen Receptor(ER) negative, Progesterone Receptor(PR) negative and Human Epidermal growth factor Receptor 2 (HER2) receptor negative (triple negative) tumors with high expression of basal markers (such as keratins 5, 6, 14, 17, EGFR) and proliferation related genes [21] [23]. Tumors characterized by basal cytokeratin expression frequently also have low BRCA1 expression [24] and harbor TP53 mutations [25]. These cancers follow an aggressive clinical course and currently lack any form of standard targeted systemic therapy. Given the triple negative receptor status, basal tumors are not amenable to most conventional targeted breast cancer therapies, leaving chemo-therapy as the only option in some cases in the clinic.

HER2 over-expressing breast tumors

HER2 over-expressing tumors are ER negative, PR negative and HER2 receptor positive tumors, that show either an over expression of the HER2 receptor or mutation-induced activation of the receptor. HER2 over expressing tumors are characterized by the over-expression of other genes in the HER2 amplicon such as GRB7 and PGAP3 or mutation in the TP53 gene [26]. Patients with these tumors have a poor prognosis.

HER2 is a member of the human epidermal growth factor receptor family of transmembrane proteins that respond to extra-cellular ligands by dimerization and trans phosphorylation of their intra-cellular domains, activating various signaling pathways in the cell [27]including cell proliferation and survival. HER2 receptor positive cancers account for 20-30% of all breast carcinomas [28] and show dependence on the HER2 oncogene for survival.

These tumors, like others that show dependence on one or a few oncogenic proteins or pathways to maintain their malignant phenotype, display a phenomenon called “onco-

gene addiction", a term first coined by Bernard Weinstein [29]. It provides a rationale for molecular targeted therapy, whereby acute inhibition of onco-protein function by targeted agents leads to "tumor shrinkage" and regression, associated with proliferative arrest, apoptosis and/or differentiation. This response to targeted therapy is explained by the concept of "oncogenic shock". Oncogenic shock is defined as the acute imbalance of pro-apoptotic and pro-survival signaling from an oncogenic kinase, upon drug-mediated kinase inhibition [30]. A large number of antibodies and drugs that target specific oncogenes are based on this concept.

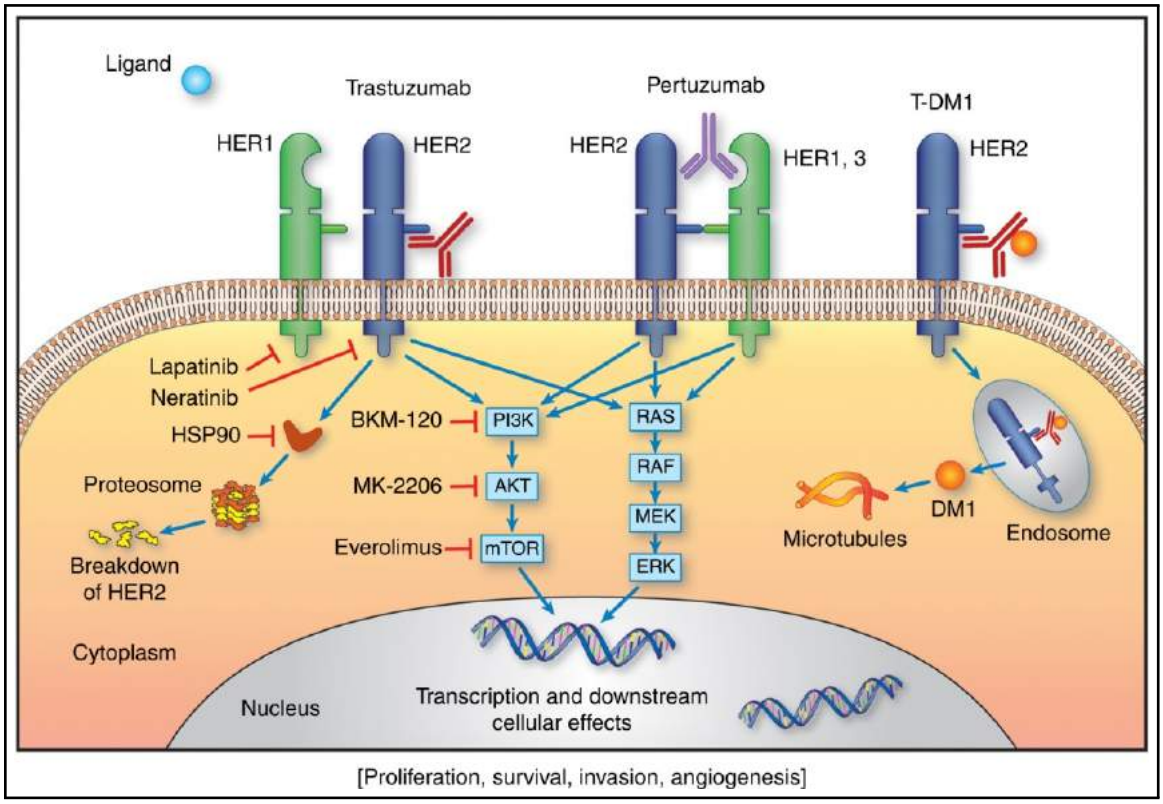


Figure 1.3: Molecular approaches to HER2 targeted therapy. Figure from [31]

.

Several targeted therapies have also been developed for HER2-enriched tumors (Figure 1.3) and have successfully led to dramatically improved prognoses, increasing the overall survival of Her2 positive patients [32]. This is seen especially in the case of the anti-HER2 antibody trastuzumab. But PTEN loss [33] and CXCR4 up-regulation [34] have been implicated in trastuzumab resistance and commonly prevent complete tumor cell eradication. This incomplete anti-tumor activity of targeted therapeutics across a wide range of cancers has been attributed to "residual disease". Residual disease is therefore defined as the population of tumor cells within a mostly therapy sensitive tumor, that survive initial treatment. This leads to accumulation of drug resistant clones that cause an inadvertent relapse that is more often than not, incurable and fatal [35].

Taken together, although great progress has been made in identifying cancer-initiating lesions and effective targeted therapies have been developed, tumor recurrence originating from minimal residual disease is the most common cause of death along with metastatic disease [36].

1.2 Mouse models of human carcinomas

Advances in the genomic analysis of human cancers, has led to a much better understanding of tumour evolution and heterogeneity, aided better classification of cancer subtypes and helped to shed light on the role of the tumour micro-environment. However, large sample numbers have to be obtained to analyse vaguely defined human tumour subtypes, confounding lifestyle factors have to be considered and ethical hurdles have to be overcome in order to obtain human samples and properly characterize disease in humans. Further, a mechanistic analysis of tumour progression and therapy response is hard to achieve with independent patient samples, since they reflect only a snap shot of these dynamic processes. To this end, mouse models have proved to be an invaluable resource to systematically and reproducibly analyse mechanisms in tumourigenesis [37] [38].

Specifically, conditional Genetically Engineered Mouse Models(GEMMS) permit the regulation of cancer inducing genes at a given time-point in a tissue specific manner allowing us to study as well as visualize the outcome of drug treatment. They are engineered to allow normal developmental processes in mice, enabling genomic manipulation that leads to *de novo* tumour formation in the adult tissue. These conditional mouse models allow tissue specific gene regulation either via Cre-ER mediated gene recombination upon tamoxifen administration or through the use of tetracycline inducible transgenes that permit reversible control over target-gene expression [39]. They also serve as a tool to understand late tumour stages by giving access to minimal residual disease following therapy and homing metastatic cells, both cellular populations that largely remain elusive in patient samples [40].

Organoid cultures of primary cells taken directly from GEMM mice are capable of preserving the original architecture and signaling events within the tumour, allowing in-depth mechanistic analysis. The establishment of 3D culture techniques [41] [42] [43] started in the late 1980s and still continues for the development of growth factor cocktails and scaffolds that mimic stem cell niches to allow the growth and long-term maintenance of organoids [44]. Organoids grown from the primary cells harvested from GEMMs can now be used to recapitulate tumourigenesis and study this process in great detail with the help of multifaceted analysis, including 'omics' approaches, and fluorescence imaging – all in a dish [45][46].

TetO-cMYC/TetO-Neu/MMTV-rtTA mouse model

Our laboratory employs the TetO-cMYC/TetO-Neu/MMTV-rtTA tetracycline inducible model of breast cancer together with a 3D culture approach to better understand the mechanisms of HER2 positive breast disease in conjunction with c-MYC activation. GEMM mice in this model have two transgenes – oncogenes that play an important role in breast tumorigenesis: c-MYC (MYC) which contains the truncated human gene with exon 2 and 3 [47] and activated rat *Neu* (Human ERBB2/ HER2 homolog)[48].

HER2 is a member of the EGFR family of receptors that is activated in 20-30% of breast cancers. Her2 has been thoroughly characterized in breast carcinogenesis with many therapeutics made available in the market for HER2 positive cancers [49]. The c-MYC transgene encodes the transcription factor MYC that has been implicated in many primary cancers in humans, mostly for its profound effect on tumor cell metabolism. It can be the primary oncogenic protein in some human cancers like for instance, Burkitt's lymphoma, while in other types MYC is more often an "early-response" gene that is downstream of other activated oncogenic pathways and signaling from ligand-receptor complexes. MYC amplification has been closely linked with HER2 amplification and proliferative activity in breast cancer [50] and by combining both oncogenes at the same time, we can model a breast cancer sub-type with poor prognosis in the clinic [51].

Reverse tetracycline-dependent transcriptional activator (rtTA) regulates temporal expression of the oncogenes which are under control of the tetracycline-dependent minimal promoter (TetO). Mouse mammary tumor virus long terminal repeat (MMTV-LTR) sequence enables spatial control of rtTA gene expression in the mammary gland of these mice. The system is regulated by doxycycline (antibiotic from tetracycline class). Addition of doxycycline to the media (in primary 3D cultures) or in the animal diet allows the rtTA inducer protein to bind to the TetO promoters controlling transcription of both oncogenes. Supplemented with doxycycline, the mammary glands or organoids undergo tumorigenesis, where tumor maintenance is dependent on the action of these two oncogenes, mimicking the phenomenon of oncogene addiction. Upon removal of doxycycline from the animal diet or 3D culture media, the oncogenes are switched off, mimicking perfect targeted therapy aimed at the driver oncogenes. The resulting population of cells in this instance, have all suffered and survived oncogenic insult and thus resemble the residual disease reservoir (detailed review in Section 3.1.1).

Previous efforts in the lab with this tetracycline inducible model of breast cancer have led to the characterization of the mechanisms that these residual cells employed to relapse *in vitro*. These cells showed a deregulated lipid metabolism resulting in ROS production and DNA adduct formation. These features could be verified in the mouse model *in vivo*, as well as in patient derived residual disease samples, collected after successful neoadjuvant treatment [52]. These studies form the basis of the scientific questions asked and answered

during my PhD.

In the following report, this very mouse model has been further adapted to mimic stochastic tumorigeneisis of the breast (Project 1) and unravel mechanisms involved in residual disease establishment (Project 2).

CHAPTER 2

UNRAVELLING STOCHASTIC TUMOURIGENESIS, ONE CELL AT A TIME

2.1 Introduction

Breast cancer, like all cancers, is thought to arise from the aberrant proliferation of a single mutated cell within the epithelial layer [53]. The role of the tumour micro-environment in breast cancer has been shown to be of extreme importance [54] and is explained by the significant alterations in the surrounding stroma or tumour micro-environment of established tumours [55]. However, although much research has gone into studying cell types like the fibroblasts [56], macrophages [57], lymphocytes [58] and even the extracellular matrix [59] that are found in the micro-environment, not much is known about the normal epithelial cells that neighbour neoplastic cells.

Especially in the context of tumour initiation, one would expect that the normal epithelial cells could play a repressive role on tumour formation owing to the protective effect of epithelial cell polarity [60]. Studies in the Drosophila model organism have played a pivotal role in elucidating cell-polarity pathways in cancer. For example, the case of polarity proteins *Lgl*, *Dlg* and *Scrib* is well documented. These proteins, when knocked down in Drosophila, result in neoplastic transformation and tumour-like growth in their larval imaginal discs and brains [61]. Consequently, the re-expression of Dlg and Lgl proteins in human tumour cell lines was shown to reduce viability and aggressiveness [62] [63]. These same polarity proteins have also been implicated, via loss of function, in primary tumours from human patients [64] [65] [66] [67]. These results, taken together, imply that cell polarity is disrupted in cancer, but the question is is this a by-product or a causal mechanism of cancer?

In the context of oncogenes and their downstream activated pathways, there are a few studies that shed more light on this matter. E6 oncogenes found in human papilloma virus (HPV) target cell polarity proteins Dlg and Scrib for proteolytic degradation [68] [69] [70]. The ability to degrade these cell-polarity proteins correlates with the malignant potential of E6 oncogenes. Similarly, in cancers caused by ErbB2 (also known as HER2) [71] [72] [27], the Ras-PI3K-PLC pathway is activated to stimulate cell proliferation but it has been shown to directly disrupt cell polarity through its interaction with the Par6-aPKC protein complex [73]. These results support the idea that along with loss of tumour suppressor genes, the disruption of cell-polarity mechanisms can play a causal role in tumour initiation and is not just a by-product.

Similar to oncogenic pathways, there is an intimate connection between cell-polarity pathways and tumour suppression. The von Hippel-Lindau (VHL) tumour suppressor gene mutations have been uncovered in the development of hemangioblastoma, clear-cell renal carcinoma and pheochromocytomas [74][75]. Wildtype VHL poly-ubiquitinates activated aPKC trageting it for degradation [76] and regulates polarized microtubule growth and

formation of primary cilia [77]. This function makes VHL an important regulator of cell polarity because polarized growth of microtubules is crucial for cell polarization. Similarly, the network regulation of Phosphatase and tensin homolog (PTEN) is also closely intertwined with cell polarity. PTEN negatively regulates the phosphatidylinositol 3-kinase (PI3K) pathway by dephosphorylating the PI3K product, phosphatidylinositol (3,4,5)-trisphosphate [PtdIns(3,4,5)P3] [78]. Spatial membrane segregation of PtdIns(4,5)P2 and PtdIns(3,4,5)P3 is crucial for apical-basal cell polarity [79] because it effects the apical localization of Annexin A2 (ANXA2), Cell division control protein 42 homolog (Cdc42) and atypical PKC proteins.

Clearly, much has been revealed about the disruption of cell polarity in cancer via oncogenes and tumour suppressors within a tumour. However, close to nothing is known about the effect of healthy surrounding cells and their epithelial polarity effects on transformed cells in an epithelial sheet. The major hurdle to studying these interactions between early neoplastic cells and the normal epithelium is the inability to model stochastic tumourigenesis in tissues in a way that it can be followed on a single cell level.

Mouse models, especially in conjunction with 3D organotypic cultures, have been used to study various stages of cancer [52, 80]. However, these mouse models are genetically engineered to express oncogenes in all cells of the epithelial tissue and end up modelling a tissue wide neoplastic transformation that is not analogous to human disease establishment. It also fails to generate a certain heterogeneity at the early tumour initiation stages that is a hallmark of cancer [81]; since all cells start to proliferate in a highly "transformed" tissue environment instead of evolving and acquiring accruing genetic mutations in a dynamic cancer initiation niche. Molecular, phenotypic, and functional diversity within a patient's tumour (intra-tumour heterogeneity) and among tumours from different patients (inter-tumour heterogeneity) are features that can complicate diagnosis and challenge therapy. So it is on prime importance that we uncover the molecular mechanisms underlying early tumour initiation that give rise to heterogeneity.

In this project, we report the establishment of the very first stochastic breast tumourigenesis model using organoid cultures transduced with lentiviral vectors – a first step towards modelling disease from its inception to gain insights into the largely unknown healthy-tumour cell interactions in an intact epithelium.

Further, we report the optimization of Selective Plane Illumination Microscopy (SPIM) to visualize the stochastic breast tumourigenesis events occurring in the 3D organoids. Imaging of delicate primary cell organoids can be tricky and was possible only because SPIM microscopy facilitates rapid imaging, with low photo-toxicity, greater penetration depths and low scattering owing to illumination with a light sheet instead of a point

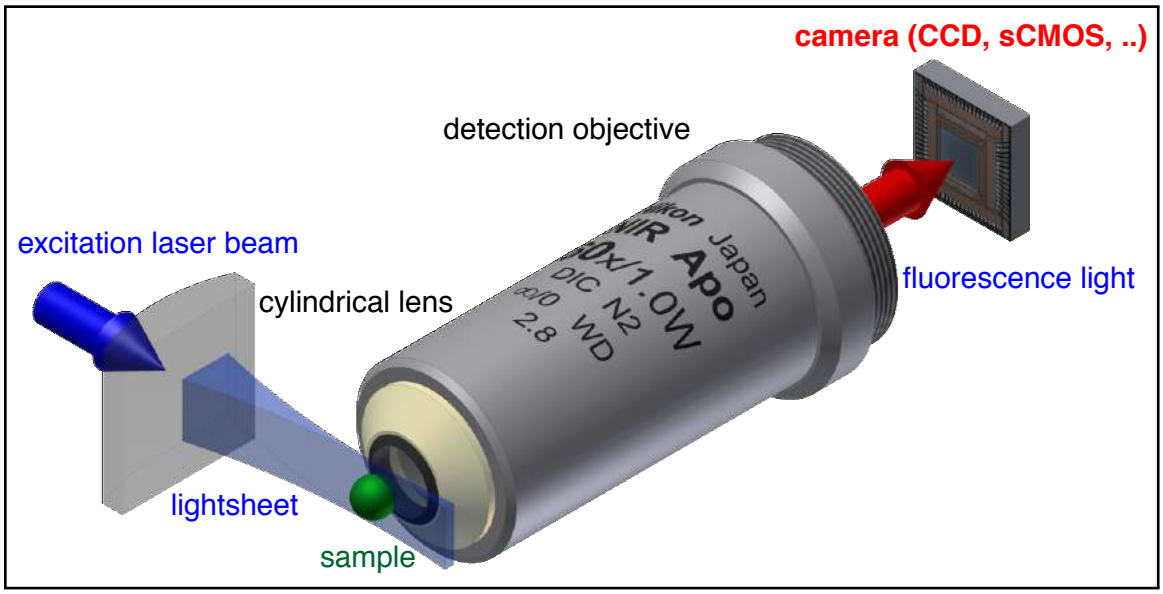


Figure 2.1: Selective Plane Illumination Microscopy provides optical sectioning of the sample using a laser light sheet for sample illumination. Only a thin slice of the sample is illuminated perpendicular to the detection plane and image acquisition with this method is faster and less photo-toxic than traditional confocal microscopy. Figure from [82]

.

source (Figure 2.1). SPIM imaging has a very low photonic load and therefore has been used to visualize tumour dynamics in 3D cultures [83]. But to date, studies have been limited in their cellular and temporal resolution. High resolution imaging can allow us to track single cell dynamics. But this limits the time frame in which organoids can be imaged without phototoxic effects[84]. Conversely, imaging primary organoids for longer time periods requires an offset of temporal and cellular resolution that eventually does not facilitate single cell fate tracking [85].

One of the drawbacks of SPIM imaging in 4D time-lapse is the generation of large amounts of data that needs to be analyzed. Big-image data compatibility is still quite limited in common image analysis tools used in bio-image analysis. Therefore, we developed a novel image analysis pipeline that allowed data handling and processing such that the dynamic cell behaviors in developing tumours could be segmented and tracked over time.

2.2 Objectives

The main aim of this project was to establish a model of breast tumourigenesis that better recapitulates the clinical situation in human disease. By initiating oncogenic transformation in single cells in an intact epithelium, we aimed to better recapitulate tumour initiation that is believed to occur in the human breast epithelium - one rouge cell that acquires a genetic mutation and proliferates to establish tumourigenesis. We planned to employ the 3D culture organoids of primary cells grown from TetO-Myc/TetO-Neu mice. The oncogenic constructs in this mouse model are under the control of the TetO promoter, that can only be activated to transcribe the oncogenes in the presence of the rtTA inducer protein and doxycyxline. In order to activate the oncogenes in single cells of the organoids' epithelial layer, we planned to deliver the rtTA inducer via a lentiviral construct. To further our aim of understanding the mechanisms of tumour establishment in the intact epithelium, we wanted to image the organoids at a resolution in space and time that would allow us to track single cells and their proliferative/apoptotic dynamics while single cells were expressing oncogenes in the otherwise normal epithelial organoid.

To achieve these aims, we established the following objectives for the study:

1. Establish stochastic lentiviral transduction of single cells within organoids such that they have been conferred the ability to express the c-MYC and Neu oncogenes, and can be followed in culture by a fluorescence reporter protein (GFP).

2. Optimize imaging conditions necessary to record stochastic tumourigenesis from the moment the oncogenes are induced in GFP marked cells (via doxycycline addition to the media) and up to 96 hours after.

3. Implement a big image data analysis work-flow that facilitates the visualization of terabytes of image data, further allowing the segmentation and tracking of the nuclei (marked with GFP) within a stochastically transformed organoid.

2.3 Results

2.3.1 Establishing a 3D *in vitro* model of stochastic breast tumourigenesis

The TetO-MYC/TetO-Neu/MMTV-rtTA mouse model is an inducible breast cancer model and has been previously been shown to faithfully recapitulate human breast disease in our group[52]. In this tri-transgenic model of tumourigenesis two oncogenes – Myc and Neu –are spatially and temporally controlled in the mouse breast tissue. The oncogenic constructs have the tetracycline promoter TetO that allows expression of oncoprotein when bound by the reverse tetracycline transactivator (rtTA) protein [86] in the presence of the tertracycline analogue – doxycycline. Spatial control of oncogenes is achieved by the MMTV promoter upstream of the rtTA construct than restricts rtTA protein expression to the epithelial cells of the breast. Temporal control of the oncogenes is achieved by the regulation of doxycycline in the animal diet. This way, the mice can be induced to express oncogenes in their breast epithelial tissue when they reach adulthood (8-10 weeks). Primary mammary epithelial cells from this tri-transgenic mouse model have been grown in 3D cultures and induced to express oncogenes *in vitro* [87, 48].

To model stochastic tumourigenesis and mimic the clinical situation better, we used a modified version of the above-mentioned tri-transgenic system. We only retained the two oncogenic constructs and bred bi-transgenic mice for experiments. We facilitated oncogenic expression in a few cells by delivering the rtTA gene construct to random cells in organoids *in vitro*, using a lentiviral delivery system developed in the lab.

Single cells digested from the mammary glands of the transgenic mice are seeded in a 3D matrix and cultured to form hollow acini. After 3-4 days of growth, these acini are digested out of the Matrigel matrix and mixed with lentiviral particles in solution. The organoids and virus are seeded back into the 3D matrix and after a 3-day recovery period, induced with doxycycline in the media, to facilitate expression of oncogenes in rtTA expressing cells (Figure 2.2).

Organoids grown from the mammary cells of bi-transgenic mice are transduced with the Inducer-reporter (pLenti-rtTA-GFP) lentiviral particles that express rtTA in single cells and mark them with H2B-GFP. Introducing doxycycline into the media expresses c-MYC and Neu oncogenes in the randomly transduced subset of cells in the organoid. When and if tumours are established in these organoids, they develop in the presence and of normal epithelial cells in their micro-environment. The random transduction and subsequent transformation of single cells in these acini (representing normal epithelia) is aligned more closely with the patient situation and represents a stochastic model of tumourigenesis.

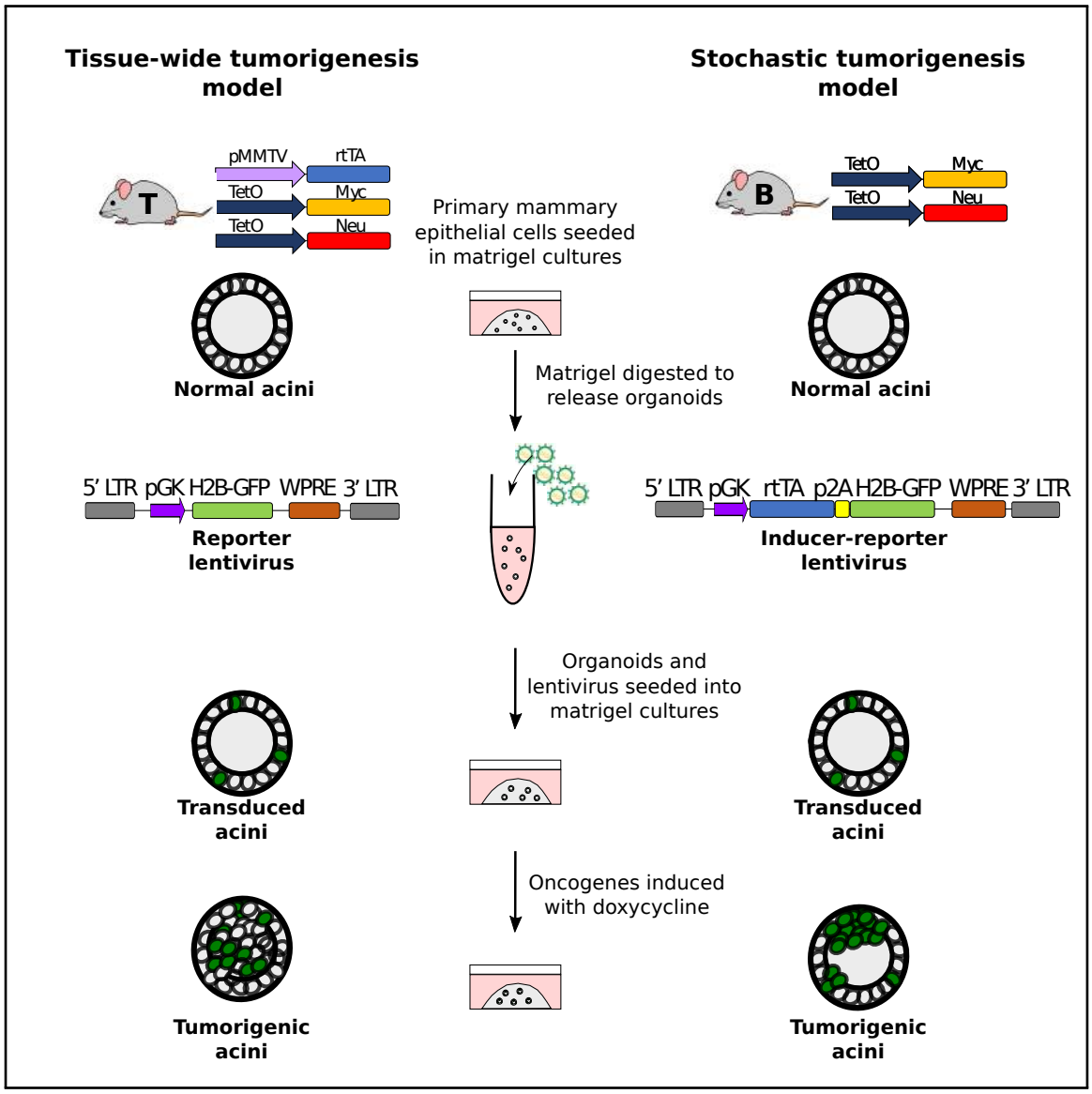


Figure 2.2: Schematic representation of the mouse models and the *in vitro* culture methods used. Organoids are grown from single cells harvested from the mammary glands of either bi-trangenic (B) or tri-trangenic (T) mice, transduced with lentiviral particles in solution and re-seeded into 3D cultures. Doxycycline is added to the media to induce the expression of oncogenes in cells expressing rtTA. B mice have the c-MYC and Neu oncogene constructs in their genome. These oncogenes are activated in single cells infected with the Inducer-reporter (pLenti-rtTA-GFP) lentiviral particles, in the presence of doxycycline — modelling stochastic breast tumourigenesis (bottom panel). T mice have the rtTA transducer construct along with the oncogenes and all cells in T organoids can be induced to express oncogenes in 3D culture in the presence of doxycycline. T mice infected with Reporter (pLenti-NULL-GFP) lentiviral particles are used as infection controls (top panel). Both viral particles mark single cells in the organoids with H2B-GFP.

As transduction controls, organoids grown from the mammary cells of tri-transgenic mice were transduced with the Reporter (pLenti-GFP) lentiviral particles that simply mark single cells with H2B-GFP. Since all cells in the tri-transgenic organoids have the rtTA construct, introducing doxycycline into the media expresses c-MYC and Neu oncogenes in the whole organoid. All cells proliferate rapidly, and the monolayer epithelial organoids form tumours with multiple layer rims. The tumours are established in a completely tumourigenic environment and thereby this system models tissue wide tumourigenesis.

Since the rtTA protein expression is driven by a different promoter in both systems, the doxycycline concentration used in the stochastic system was normalized to the established concentration of doxycycline used to induce organoids of the tissue wide system. Quantification of c-MYC and Neu mRNA in single cells across both systems was performed using qPCR, at various doxycycline concentrations (see Figure 2.3). The doxycycline concentration used to induce oncogenic expression in the tri-transgenic system has been optimized at 800 ng/ml by previous efforts in the group. Based on the qPCR quantification, 600 ng/ml of doxycycline in the stochastic system leads to similar expression of the oncogenes in single cells of the stochastic system. This concentration was used in all further experiments involving the stochastic system.

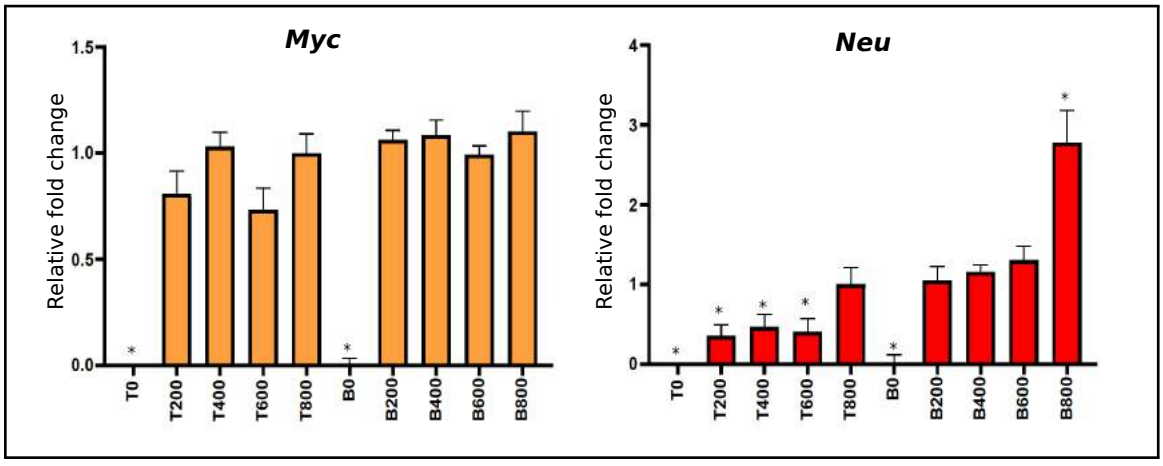


Figure 2.3: Fold changes in the mRNA expression of transgenes, Myc and Neu in transduced mammary epithelial cells of B mice (n=2) infected with Inducer-reporter virus or T mice (n=2) with Reporter virus. The doxycycline dosage of 800 ng/ml (T800) is well established in the T cells and was used as control to normalize the gene expression, and also to determine the dose for transduced B cells (600 ng/ml). Data represented as mean $\pm SEM; *P < 0.05$.

Immunofluorescence staining of PFA fixed 3D gels, shown in Figure 2.4 validate the

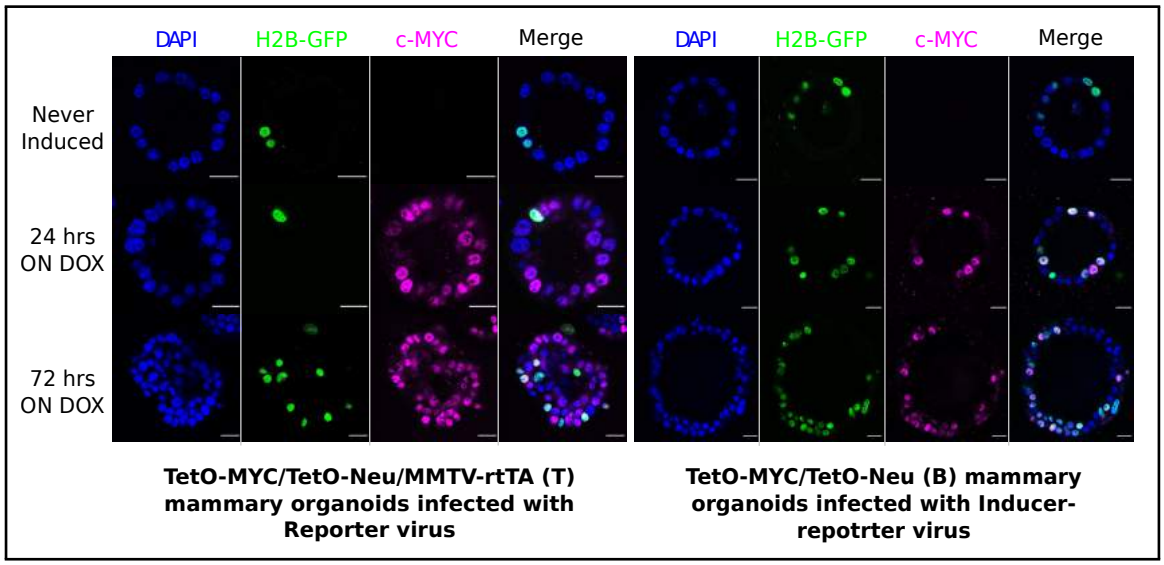


Figure 2.4: Representative immunoflourescence staining images of fixed 3D gels with B organoids transduced with Inducer-reporter virus or T organoids transduced with Reporter virus before induction (top), 24 hours post induction and (middle) 72 hours post induction (bottom) with doxycycline. c-MYC oncogene (magenta), GFP expressing transduced cells (green), DAPI nuclear stain(blue). Scale bar, 10 μm.

expression of MYC in only the transduced cells of the bi-transgenic organoids (stochastic system). A co-localisation of MYC protein is seen in the cells transduced with the Inducer-reporter virus (marked with H2B-GFP). In contrast, the MYC stain is positive for all cells in the tri-transgenic organoids (tissue wide tumourigenesis system). Here the cells expressing GFP are simply marked by the Reporter virus.

2.3.2 Live organoid imaging reveals early tumour dynamics in 4D at single cell resolution

To visualize the establishment of tumours in the otherwise healthy epithelial rim and analyse their evolution on a single cell scale, we crossed the fluorescence reporter H2B-mCherry into the bi-transgenic and tri-transgenic mouse lines using a R26-H2B-mCherry line. Lentivirus-transduced organoids grown from these H2B-mCherry lines were then imaged during induction using Selective Plane Illumination Microscopy (SPIM). Light sheet microscopy employs a sheet of laser light to scan the sample. In comparison to confocal and spinning disc microscopy methods, this offers faster acquisition, high sample penetrance and a low photonic load. All the imaging experiments for this project were performed on the InVi SPIM microscope (Luxendo Light-Sheet, Bruker Corporation). Sample mounting for the InVi SPIM is performed using a custom-made sample holder. The holder has a trough with a slit at the deep end where the sample is placed for imaging. The sample is placed on a 25 μm thin membrane which is glued onto the sample holder using a mould. We transferred gel slivers containing transduced bi/tri transgenic organoids onto the FEP sheet of the holder and induced oncogenic expression by adding media supplemented doxycycline to the sample holder (Figure 2.5).

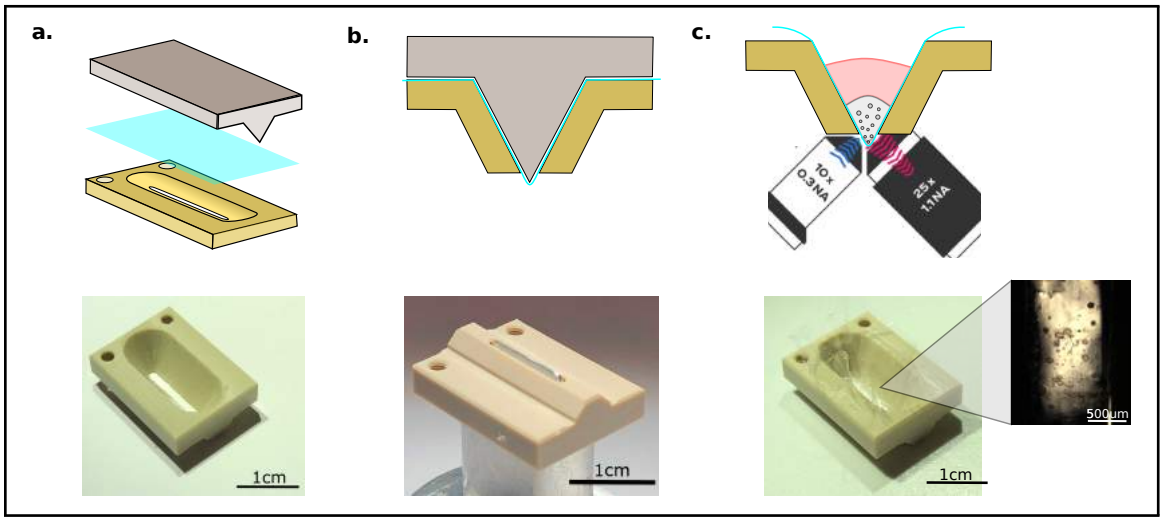


Figure 2.5: (a) The FEP membrane is glued onto the sample holder with the help of a mold and bio-compatible glue. (b) Gel slivers are transferred to the FEP sheet trough in the sample holder (c) Gel slivers are overlaid with fresh matrigel to prevent drift during imaging. Media is added after the upper matrigel layer solidifies.

The process of tumour establishment in both systems was then recorded in 4D. Images were recorded as 2D planes ranging from 100-500 in number, depending on the organoid size. Each 3D stack of planes was recorded in 2 channels - mCherry (all cells) and GFP (transduced cells). Depending on the duration of the time lapse imaging, 450-600 image stacks (equivalent to 72-96 hours) were recorded per organoid at 10-minute intervals. A series of optimization experiments, involving different laser powers, exposure times and

z-step sizes yielded laser powers of 13.1 microWatts for 488 nm and 36.4 microWatts for 594 nm, 100 millisecond exposure time per frame and 1μm z-spacing between frames to be optimal for long term imaging (96-120 hours) without photo-bleaching or photo-toxic effects on growth.

Tri-transgenic organoids transduced with the Reporter virus showed high cell proliferation and rapid expansion when the oncogenes were expressing in the cells. The single layer organoid rims sustained high cell proliferation to result in the filling of the lumen. The tumour organoids have multiple cell layers and pronounced apoptosis that is associated with tumour evolution (Figure 2.6a).

In contrast, the bi-transgenic organoids transduced with the Inducer-reporter virus showed a dual phenotype during imaging. Some lentivirus transduced cells within organoids proliferated under the oncogenic push and established small tumour foci within the organoid. They sustained high proliferative rates in their niches within the normal epithelium and started to fill in the hollow lumens. Other lentivirus-transduced cells did not seem to proliferate at all. They remained integrated in the normal epithelium and did not proliferate or establish any tumourigenic foci within the organoids (Figure 2.6b).

To exclude for imaging artefacts, we performed immunofluorescence staining on PFA fixed 3D sister gels grown in the incubator (Figure 2.7). Consistent with the light-sheet movies, 3D gels grown in the incubator showed a similar dual phenotype for the bi-transgenic (stochastic tumourigenesis) organoids while the tri-transgenic (tissue wide tumourigenesis) organoids consistently formed tumours upon oncogene induction.

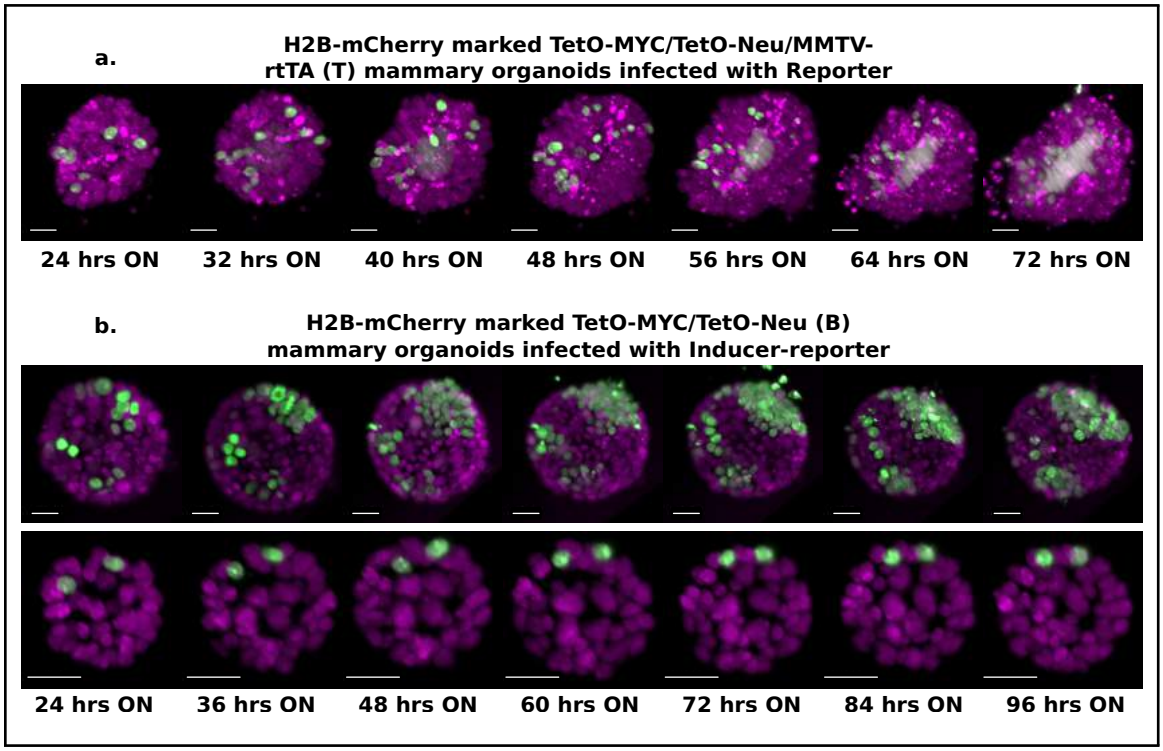


Figure 2.6: 3D images of selected timepoints during live-cell time-lapse microscopy of induced T organoids transduced with Reporter virus (a) or B organoids transduced with Inducer-reporter virus (b). All cells in the organoids express H2B-mCherry (magenta) and only cells transduced with lentiviral particles express H2B-GFP (green). Imaging was started 24 hours after oncogenic induction with doxycycline. In (b) the upper panel shows the proliferative phenotype seen with stochastic transformation, whereas the lower panel shows the non-proliferative phenotype observed in some stochastically transformed organoids. (Imaging conditions: H2B-mCherry 594nm Ex, 610 LP Em; H2B-GFP 488nm Ex and 497-554 nm Em). Scale bar, 20 μm.

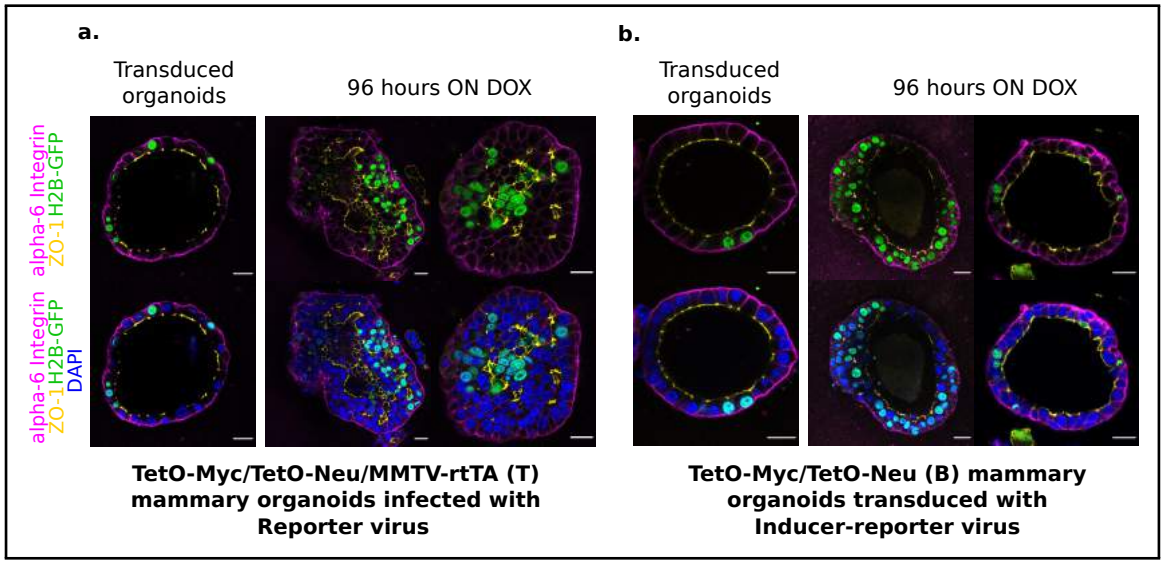


Figure 2.7: Representative immunofluorescence staining images of PFA fixed 3D gels with (a) T organoids (transduced with Reporter virus) and (b) B organoids (transduced with Inducer-reporter virus), before induction and 96 hours post induction with doxycycline. Polarity markers include alpha-6-Integrin (magenta) and ZO-1 (yellow). Transduced cells are marked with GFP (green) and nucleus is counter-stained with DAPI (blue). Scale bar, 20μm.

2.3.3 Development of big-image data compatible analysis workflow allows tracking of cell lineages during early tumour formation

To analyse the 3D time-lapse movies recorded at the InVi SPIM, we developed an image analysis workflow using Fiji plugins[88] and Imaris (commercial software[89]). The workflow was designed and developed with the help of Christian Tischer at the Center for Bioimage analysis (CBA) at EMBL, Heidelberg.

The establishment of stochastic tumours in bi-trangenic organoids transduced with the Inducer-Reporter virus was recorded in two channels - H2B-mCherry—all cells in the organoid, H2B-GFP—transduced cells within the organoid. The raw files from the microscope in .h5 format were streamed using the Big Data Processor (BDP) Fiji plugin [90]. The Fiji plugin provides a virtual viewer that allows lazy loading of the image stacks in 2D for visual inspection. The "Chromatic Shift Correction" tab of the BDP was used to align the two-channel data. The "Cropping" tab of BDP was used to crop out empty/black pixels and remove empty planes. The "Saving" tab of the BDP was used to bin images (3 x 3 x 1 in x, y, z), perform 8-bit conversion and convert .h5 files from the InVi SPIM into an Imaris compatible multi-resolution file format (.ims) for further analysis (Figure 2.8).

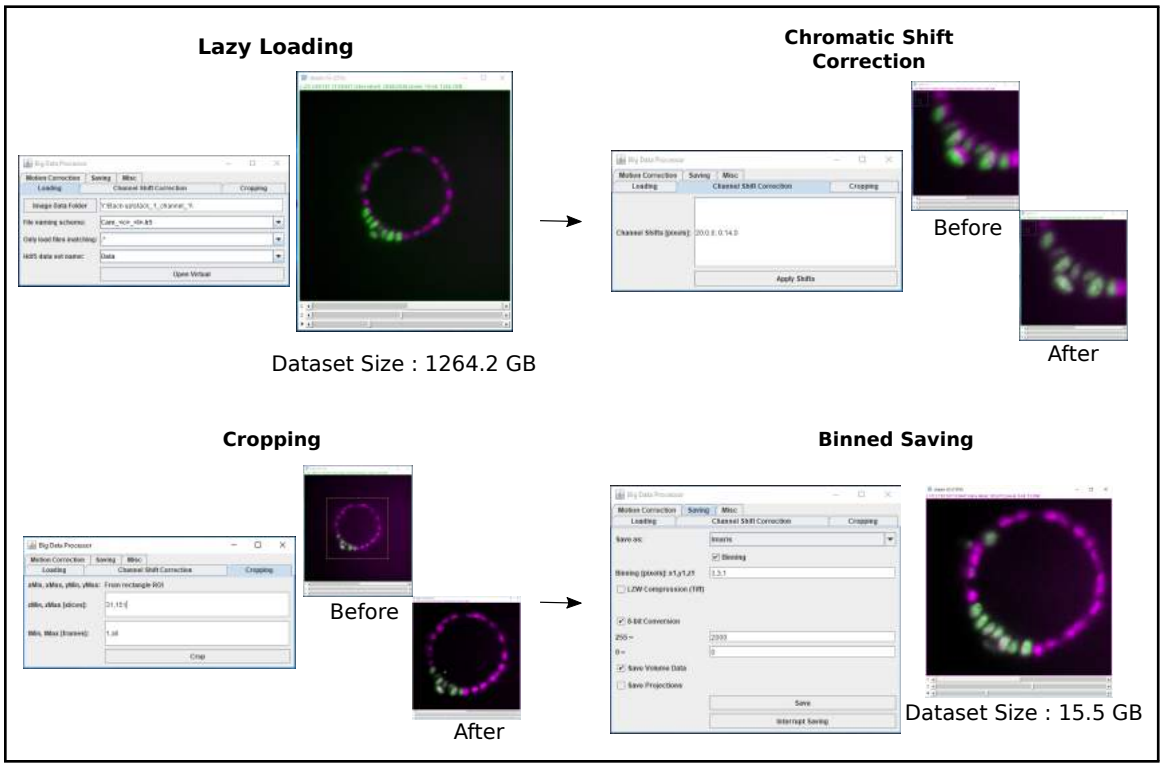


Figure 2.8: The Big Data Processor Fiji plugin was employed for pre-processing light-sheet microscopy images. Two-channel raw images were lazily loaded in 2D slice mode for visual inspection. Channel shift correction was performed to align the two channels. Then, the whole dataset was cropped in x,y,z to remove black pixels and empty planes. The cropped dataset was then saved in an 8-bit Imaris format with 3x3 binning applied in x and y.

The lentivirus-transduced cells marked with H2B-GFP showed heterogeneous morphologies and varying intensity textures in the time-lapse recordings. This makes it difficult to segment them using conventional thresholding approaches, such as the ones in-built in Imaris. Therefore, we used a trainable segmentation approach to convert the raw intensity values into pixel probability maps, using the Fiji plugin Context Aware Trainable Segmentation (CATS) [91]. We trained three pixel classes: background, nucleus center and nucleus boundary on the input files (H2B-GFP channel images). On these we drew about 20 (background), 120 (nucleus center), 100 (nucleus boundary) labels distributed across the different time-frames of the movie for training the classifier. After feature computation and training of a Random Forest classifier the whole dataset was processed on EMBL's high performance computer cluster. The nucleus centre probability maps were then exported from CATS and added as an additional channel to the converted intensity data in the processed .ims files (Figure 2.9).

The data were then loaded into Imaris and visualized, segmented and tracked in 3D (Figure 2.10) The "surfaces" function on Imaris was used to segment the cell nuclei using the nucleus centre probability map channel. Probability maps were manually thresholded using a surface smoothening parameter of 0.3 μm; the minimum quality parameter for seed points was set to 0.1, and object splitting was applied for objects larger than 5.5 μm. Objects with volumes less than 20 μm3 were excluded.

Segmentation accuracy was assessed by analyzing wrongly segmented cells. In Figure 2.11 an example organoid, whose transduced cells were segmented, is shown at 4 time points, counting: True Positives (correctly segmented cells, highlighted in green), False Merges (two cells merged as one, highlighted in orange), and False Splits (one cell split in two, highlighted in purple). Unidentified cells are indicated as False Negatives and Ground Truth indicates the actual number of cells at each time point. Majority of errors in the object segmentation were clearly the false merges, where two cells were segmented as one. This kind of error is frequently not sustained in the previous or following time-points and hence, an increased gap size parameter of the tracking algorithm should not affect tracking. This was verified by tracking the objects identified using Imaris' Lineage tracking function with a maximum distance between objects in subsequent time-points limited to 10 μm and a maximum gap size between identification of the object in a particular track limited to 10 time points. The final track trees of transduced cells within the organoid were corrected manually within Imaris, e.g., excluding apoptotic cells and auto-fluorescent debris.

In Figure 2.12 we see an example bi-transgenic organoid transduced stochastically by the Inducer Reporter virus. The implementation of the analysis workflow allowed us to track the evolution of single cells when induced with doxycycline over 3 days. Clearly, single transduced cells within one organoid show a difference in proliferation and cell fate as indicated in the representative tracks.

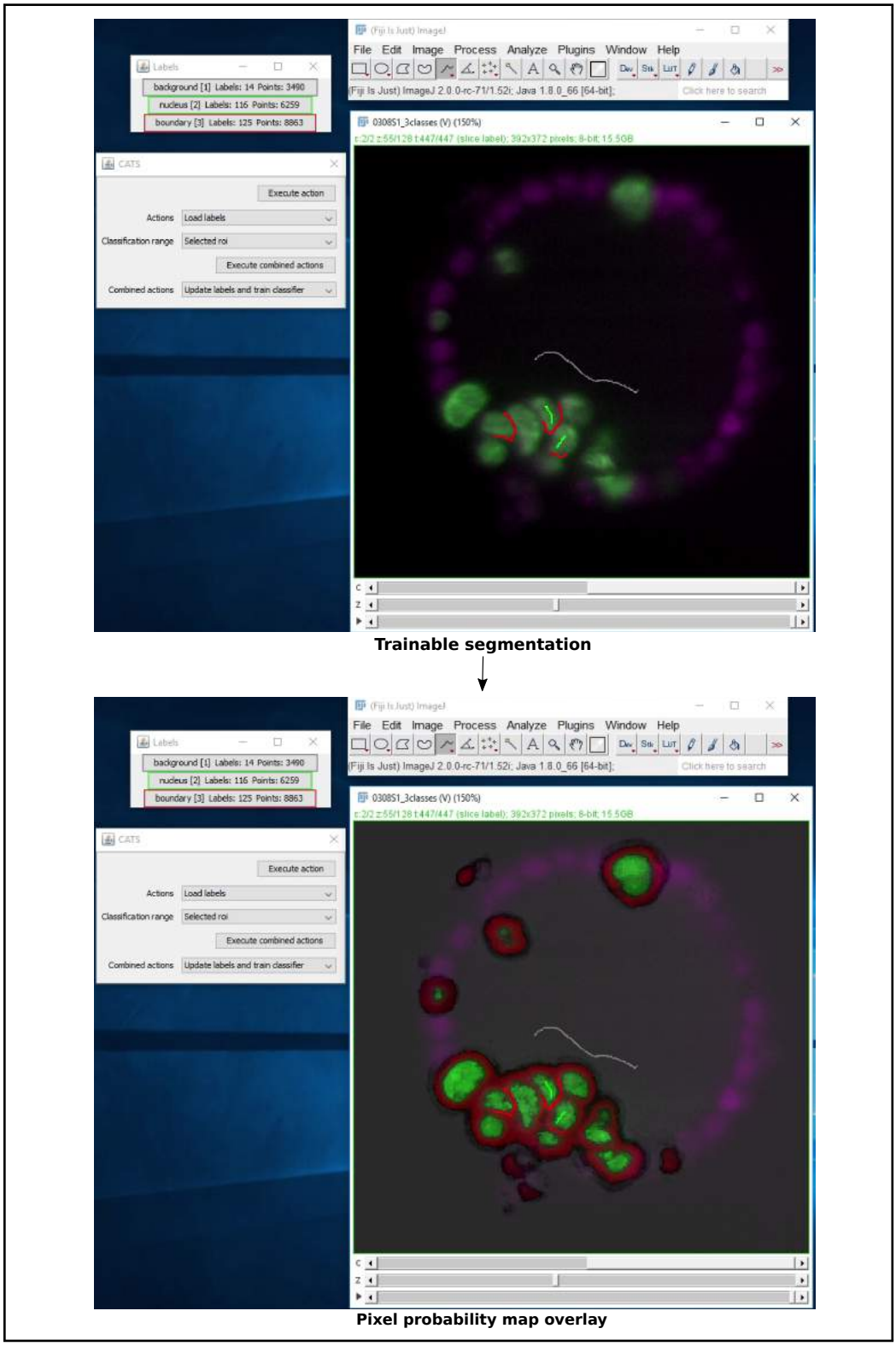


Figure 2.9: The CATS Fiji plug-in was used to generate pixel probability maps for H2B-GFP images. Left panel shows the manual training done by drawing labels on the dataset to classify pixels into 3 classes – background(grey), nucleus boundary (red) and nucleus center (green). The right panel shows the pixel probability output for all three classes overlaid on the intensity data. Only the pixel probabilities from the nucleus center class were exported from CATS and linked to the Imaris data set for further segmentation and tracking on Imaris.

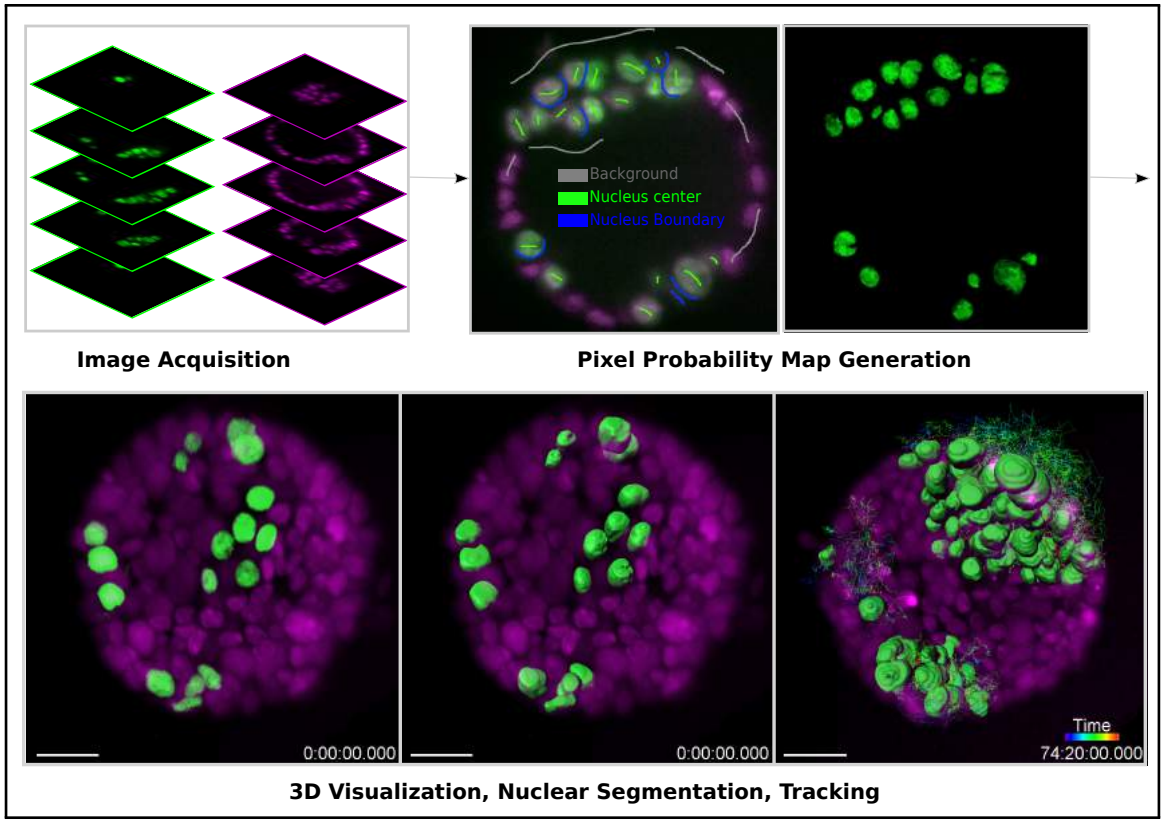
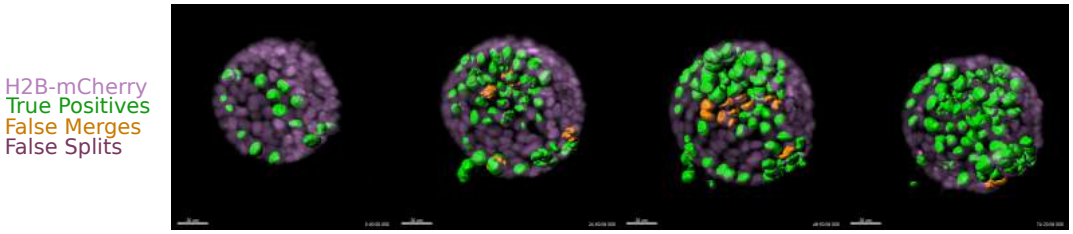


Figure 2.10: Schematic representation of the big-image data analysis pipeline developed to analyze the light sheet microscopy images. Images are acquired in two channels (H2B-mCherry in magenta and H2B-GFP in green) at 10-minute intervals for 3-4 days. Big Data Processor Fiji plugin is used to pre-process the raw images and CATS Fiji plugin is used for generation of pixel probability maps. Image pixels of the H2B-GFP images are classified into background (black), nucleus centre (green), nucleus boundary (blue) classes by manual training. Processed raw images along with the probability maps from the nucleus center channel (green) are exported to Imaris for 3D visualization, nuclear segmentation and single cell tracking.



	T = 1	T = 150	T = 300	T = 447
True Positives	17	52	69	74
False Merges	0	4	6	1
False Splits	0	0	2	0
False negatives	0	1	0	0
Ground Truth	17	61	82	76

Figure 2.11: Image panels show the H2B-mCherry signal (magenta) along with the segmented H2B-GFP cells of an organoid at four equidistant timepoints. Segmentation accuracy was assessed, counting: True Positives (correctly segmented cells, highlighted in green), False Merges (two cells merged as one, highlighted in orange), and False Splits (one cell split in two, highlighted in purple). Unidentified cells are indicated as False Negatives and Ground Truth indicates the actual number of cells at each timepoint. The average True Positive to Ground Truth ratio for the four time-points is 0.92. Scale bar, 20 μm.

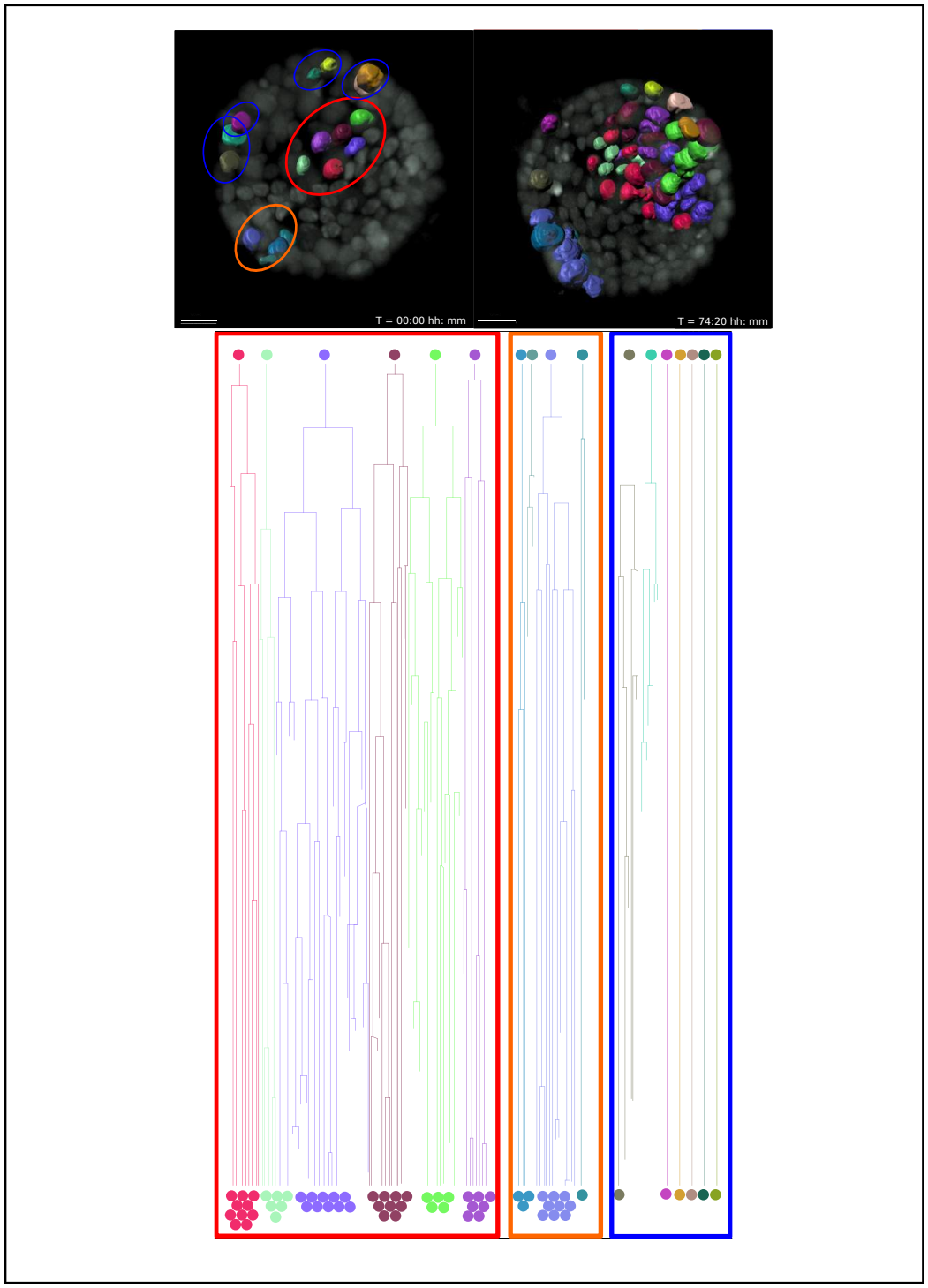


Figure 2.12: Representative B organoid transduced with the Inducer-reporter virus at the beginning of the time-lapse (24 hours post induction with each transduced cell surface rendered with Imaris) and the organoid at the end of the time-lapse (∼76 hours post induction with doxycycline). The bottom panel shows the lineage trees of each individual cell over the time lapse recording. Lineage trees of single cells are grouped into proliferative (highlighted in red, orange) and non-proliferative (highlighted in blue) cell clusters. Color coding of each cell maintained in all panels. Scale bar, 15um.

2.3.4 Feature analysis of lentivirus-transduced organoids indicates that the origin of tumours may be in cell clusters and not in single transformation events

The implementation of the image analysis workflow on the time lapse recordings shows us a variation in proliferation rate and cell fates of bi-transgenic organoid cell transduced with the Inducer-reporter virus. This difference in behavior of single transduction events, even among cells of the same organoid, in turn affects the establishment of tumour foci within the healthy epithelia. Why do single cells expressing the same oncogenes behave differently? Observation suggested that the cells that undergo transformation in proximity with each other seem to be successful in establishing tumour foci. It could also be that the more cells transduced in an organoid, the higher the chances of tumour formation. To better understand the parameters that positively affect a transduced cell in the stochastic tumourigenesis model to start proliferating and establish a tumour within a normal epithelium, we extracted 9 features of the organoids (n=20; Table 5.1) and all the transduced cells in these organoids (n=150) at the start of the imaging.

Basing the analysis of the observation that "proximity" might be the deciding factor for tumour formation, first, we exported the center of mass coordinates of all cells of the organoid so they could be represented in a 3D space for distance calculations. Next, we aimed to cluster the cells that were transduced. We computed the pairwise Euclidean distances between all oncogene-expressing cells in an organoid and applied hierarchical clustering with complete linkage. Clusters were identified automatically by cutting the branches of the trees using the dynamic tree cut algorithm [92]. A cluster was defined as a group of transduced cells that are closer to each other than to other transduced cells within the same organoid. So logically, a cluster can be composed of a single cell if this cell is comparatively isolated from other transduced cells. Mathematically, a "cluster volume" can be calculated as the volume of the sphere centered at the center of mass of the cluster with diameter equal to the distance between the two farthest oncogene-expressing cells of the cluster. This served as boundary parameter for estimation of features involving the state of the immediate micro-environment of single transduced cells.

Then we asked, what features of this organoid could play a role in tumour promotion?
(1) Number of cells in the organoid?
(2) Cell density of the organoid's epithelial layer?
(We expressed "cell density" as the ratio of number of cells to organoid surface area. Organoid surface area was computed by assuming the organoid is a sphere with diameter equal to the distance between the two most distant cells)
(3) Number of oncogene-expressing cells in the organoid?
(4) Number of cells (including both oncogene-expressing and normal cells) in the cluster volume?
(5) Number of oncogene-expressing cells in the cluster?

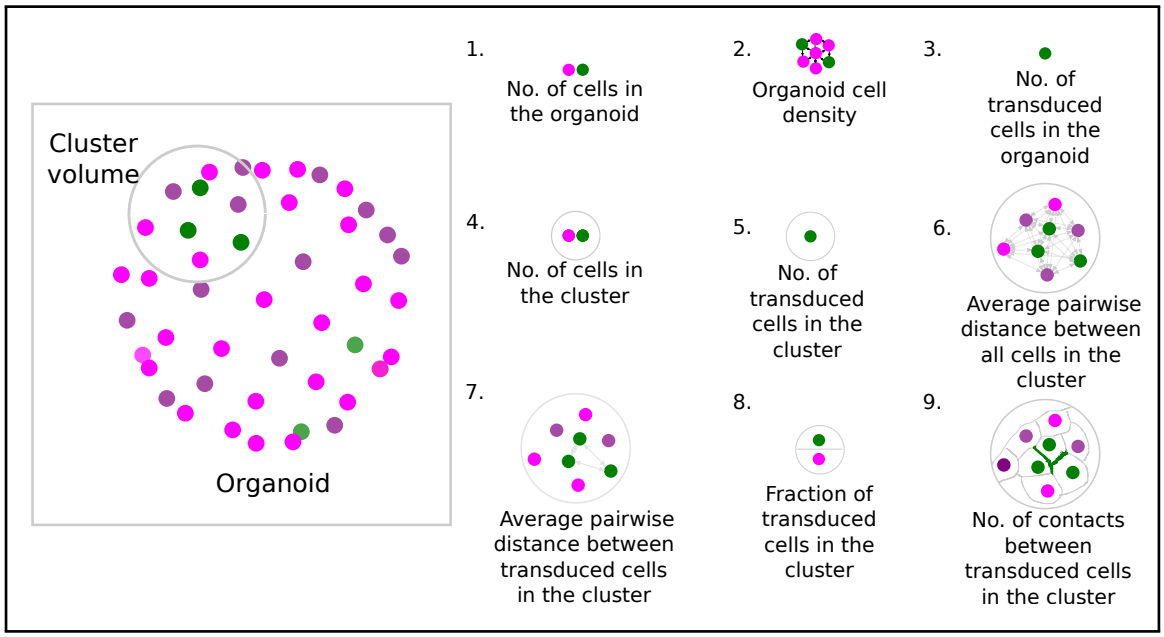


Figure 2.13: Schematic representation of the 9 features of stochastically transformed cells extracted at the beginning of time lapse imaging. These features were assessed for their impact on tumour cell proliferation within B organoids transduced with the Inducer-reporter virus using logistic regression.

(6) Average pairwise distance between all cells in the cluster volume?
(7) Average pairwise distance between oncogene-expressing cells in the cluster?
(8) Fraction of oncogene-expressing cells in the cluster volume
(9) Number of contacts between oncogene-expressing cells in the cluster?
(Two cells are presumed in contact if they are less than the average cell diameter + 2 standard deviations apart).
The above nine features were further tested for their effect on tumour formation (Figure 2.13).

Since all transduced cells were tracked over time, using the previously described big-image data compatible workflow, we were able to associate each cluster with a tumour outcome (if any of its cells lead to tumour formation). To identify which features were linked to this outcome, we took an information-theoretic approach to model selection.

We fitted a logistic regression model for all possible linear combinations of features and selected the best model based on the Akaike information criterion (with correction for small sample sizes) [93]. This model included only three contributing features: (5) Number of oncogene-expressing cells in the cluster (3) Number of oncogene-expressing cells in the organoid and (4) Number of cells in the cluster. Additionally, as indicated in Figure 2.14, p-value calculations show that only the Feature (5) Number of oncogene-expressing cells in the cluster contributed significantly to tumour formation with an odds ratio of 9.1. This means, that each additional transduced cell in a cluster increases the odds of this cluster forming a tumour by 9.

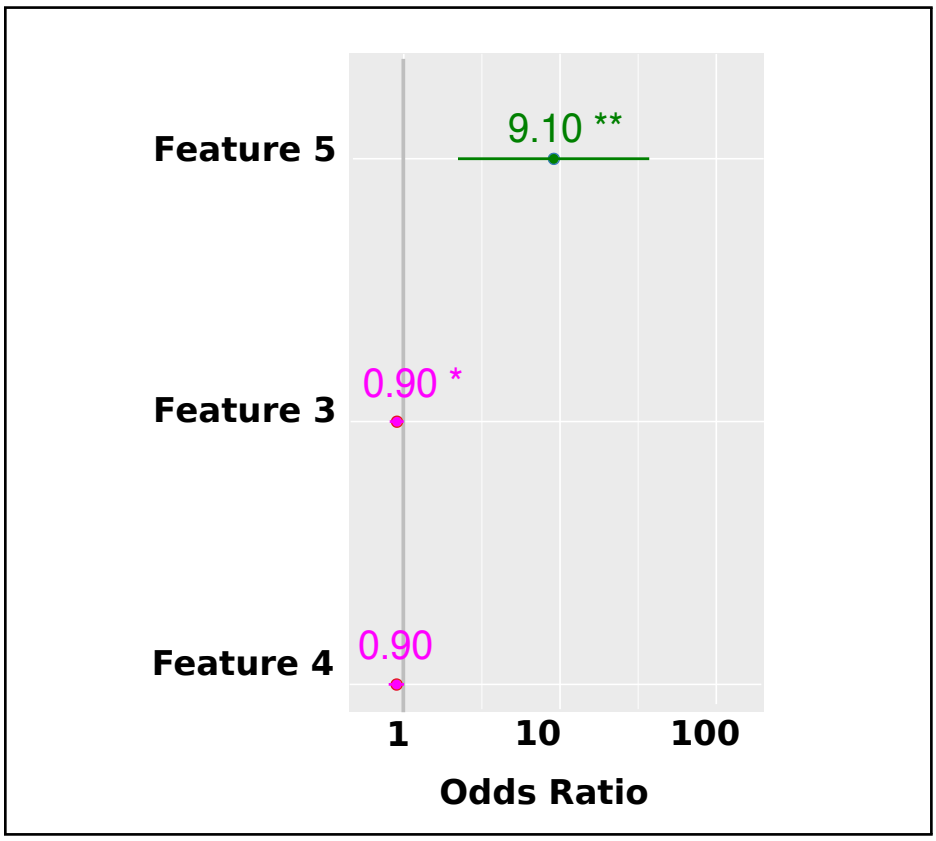


Figure 2.14: Best logistic regression model for all possible linear combinations of features based on the Akaike information criterion. Coefficients (represented as odds ratios) of the three features included in the best logistic regression model, colored horizontal bars represent the 95 percent confidence interval of the estimate. ** indicates p-value (of having no effect) less than 0.01, * indicates p-value less than 0.05. The vertical grey line indicates the position of no effect.

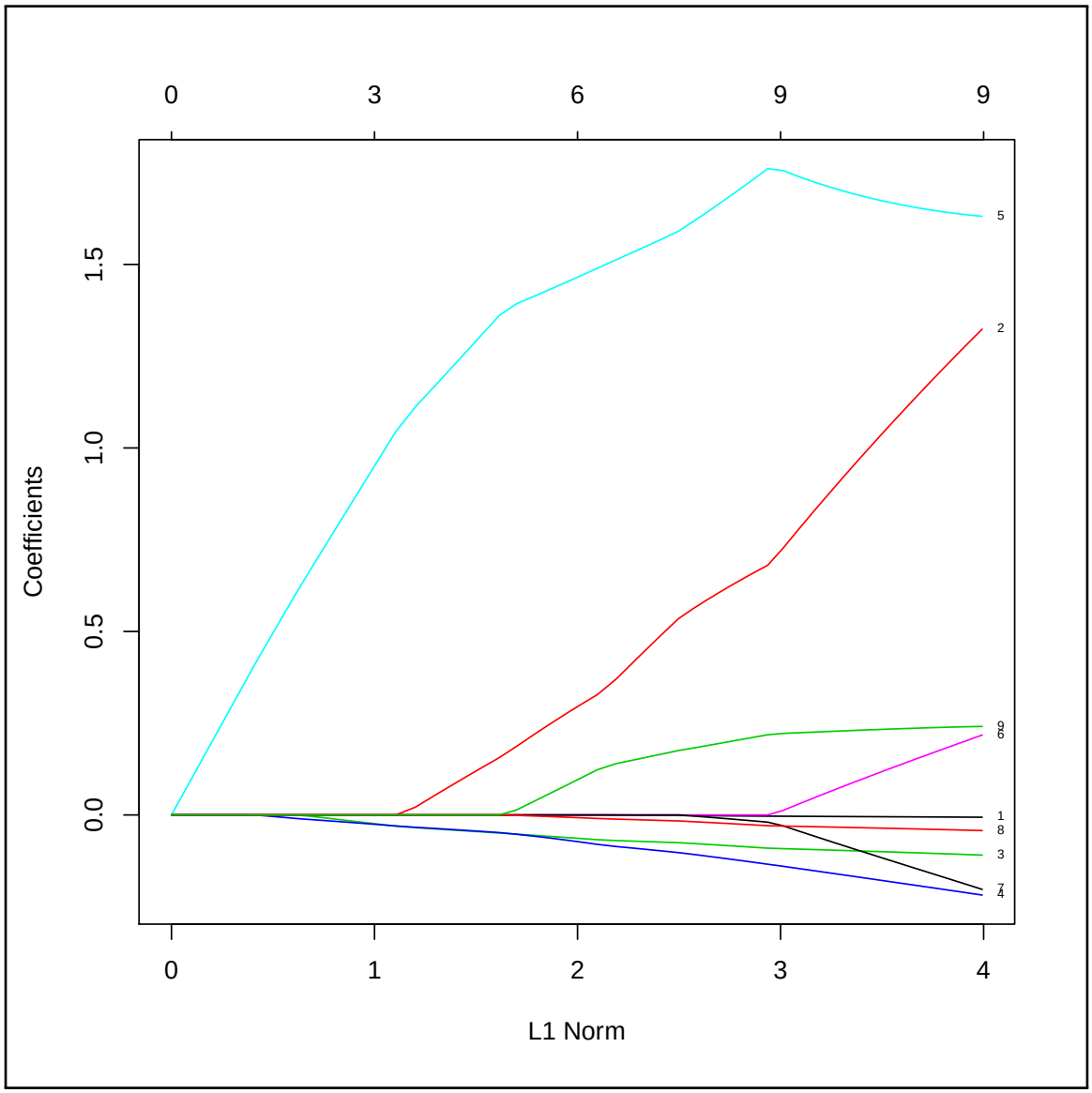


Figure 2.15: Alternate regression model fitted including all features using LASSO regularization. Numbers correspond to features shown in 2.13

Additionally, we also built a logistic regression model including all features (not just the best combination) and using LASSO regularisation to select features. The output of this model shown in Figure 2.15 confirmed our previous results that only the Feature (5) Number of transduced cells in the cluster shows a positive effect on tumour establishment.

This is further demonstrated by representative organoids shown in Figure 2.16. Here, the cells that are likely to proliferate, cluster together at the start of the time lapse imaging and the non-proliferative cells are more sparsely located within the organoid, as shown in the hierarchical clusters. These data confirm our hypothesis that there is indeed a higher chance of tumour establishment by single transduced cells if they are in close proximity to other transduced cells.

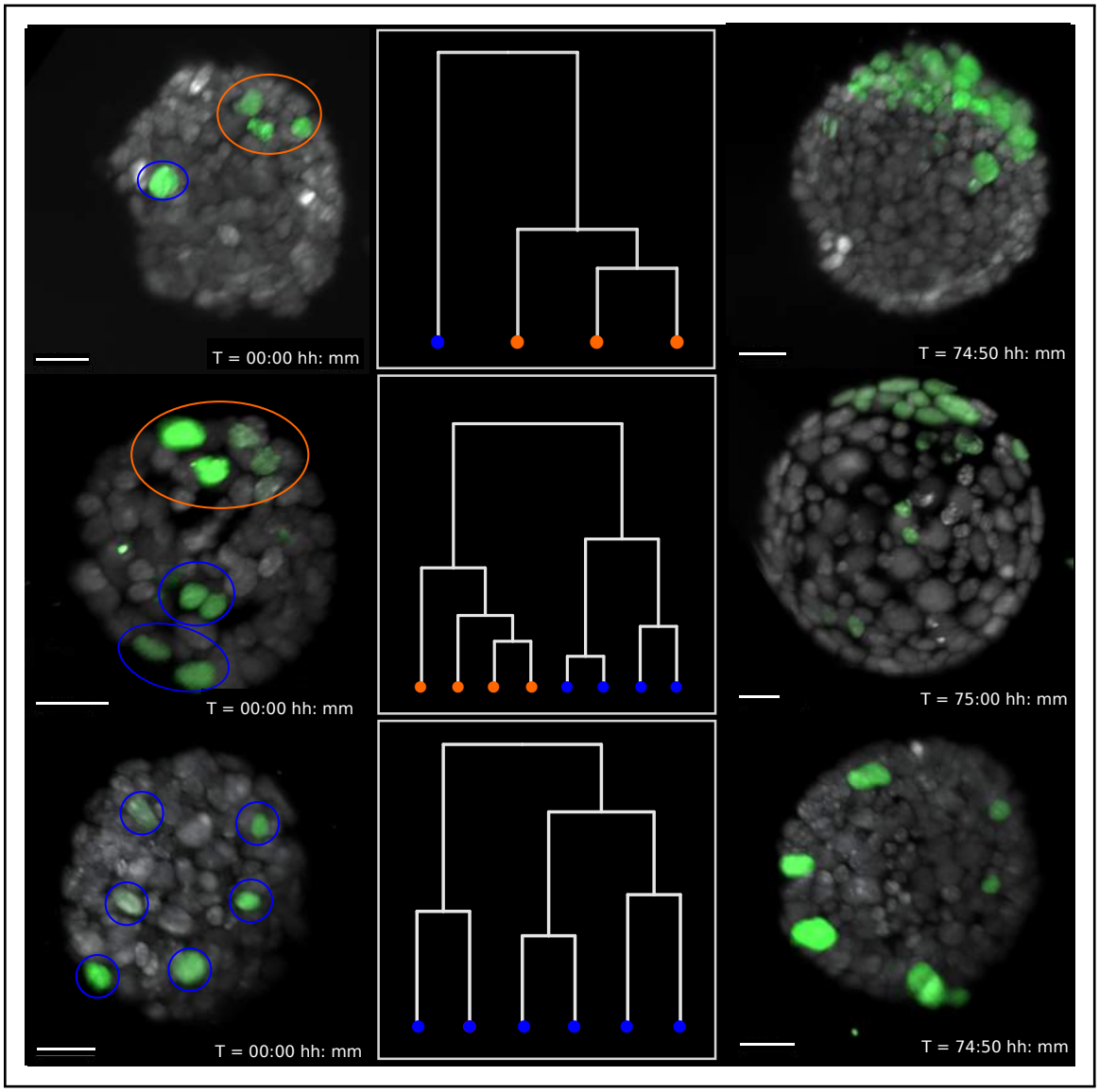


Figure 2.16: Representative B mammary organoids stochastically transduced with the Inducer-reporter virus and induced with doxycycline. Left panels show organoids 24 hours post induction. Color highlights indicate clusters of transduced cells identified from hierarchical clustering (shown in middle panels) with proliferative clusters highlighted in orange and non-proliferative clusters highlighted in blue. Right panels show the same organoids 72-76 hours post induction. Scale bar, 20μm.

2.4 Discussion

The rationale for this study, as explained earlier, is the inaccessible mechanisms of early breast cancer initiation in an intact epithelium. In humans, the detection of the earliest events in cancer initiation is limited by its cellular scale and the lack of detection technology. Even in mouse models of breast cancer we are able to achieve only a tissue wide neoplastic transformation, where the oncogenes are being expressed in all cells of the mammary epithelium. To truly understand how single transformed cells survive and proliferate in a healthy epithelilal cell layer, we established the first ever *in vitro* stochastic model of breast cancer using lentiviral transduction of single cells in murine mammary organoids (Objective 1). Further, to explore the processes underlying tumor initiation and the role of normal epithelium in these processes, we established live cell imaging of these organoids such that we could acquire images at the resolution and time interval needed to track single cells in a complex 3D system (Objective 2). Finally, we developed and optimized an image analysis workflow that could process big image data files generated during light sheet microscopy, and allowed for nuclear segmentation and tracking (Objective 3).

Modelling stochastic tumourigenesis in murine mammary organoids has shown, for the first time in 3D organoid cultures, that the normal epithelium has a profound effect on early tumour establishment. Light sheet imaging over four days, following every cell, in conjunction with feature analysis and mathematical modelling bolster the observation that small groups of oncogene-expressing cells within the healthy organoid are much more successful at establishing tumourigenic foci in comparison to isolated oncogene expressing cells. This indicates that a possible proximity controlled paracrine signalling network is established within the transduced cells that allows them to proliferate and evolve as tumours. Conversely, the inability to proliferate, that seen in the isolated oncogene expressing cells, is indicative of the "repressive" effect of the normal epithelium that has been reported before via paracrine soluble factors and even microRNAs [94].

The interaction of tumour cells with the immediate microenvironment has been subject of extensive studies with regards to immune cells [95] and other tumour associated cell types [54], however, the interaction with the normal neighboring cells has not been explored in real time using an organotypic mammalian model system. Rather, questions of cell competition in heterogenous tissues have mainly been addressed in either 2D culture systems or in the Drosophila wing [96]. Organ-specific mammalian cell types that self-organize in a manner similar to the *in vivo* situation can now be studied in specialized organoid 3D culture conditions *in vitro* [44]. Furthermore, these organoids can be used to model disease and serve as an alternative system for drug testing that better recapitulate effects as compared to conventional 2D cell culture [80] [97].

The established paradigm that cancers in epithelial tissues arise from single mutated cells has been challenged by these findings and the cell of origin of tumours has been found to be cell clusters of origin. Further characterization and multi-omics analysis of this model will validate this hypothesis in the future. Apart from the tumour cell of origin debate, this system has the ability to further investigate the importance of epithelial polarity proteins on the establishment of breast tumour progression and metastasis. For example, the loss of important polarity proteins like Par3 and Par4/LKB1 have highlighted their function as non-canonical tumour suppressors in breast tumourigenesis [98, 99]. However, other reports on the tumour initiating potential of Par3 [100] and the alternative pathways triggered by the loss of the Par4/LKB1 polarity protein[101], indicate that they could be drivers of tumour formation. Clearly, to better interrogate such conflicting reports, there is a need for a more detailed analysis, employing a model system that does not show modification of all cells in the tissue. To this end, our new stochastic tumourigenesis system can be employed to settle conflicting reports. Introducing a knockout phenotype of any of the epithelial polarity proteins into the mouse model via somatic mutation or additional lentiviral delivery is an option that can be explored.

The amenability of this system to interference with small molecule inhibitors, viral shRNA vectors and genomic editing has the potential to further our understanding of the mechanisms important during tumour initiation. The ability to distinguish marked tumour cells from the normal epithelium will now allow us to perform single cell RNA sequencing analysis on select sorted cells. This will help delineate the signalling networks within the immediate tumour micro-environment.

Taken together, we strongly believe that our integration of a true stochastic tumour model with the ability to image single-cell fates will successfully bridge the gap between genetically modified model systems and the clinical situation, helping gain novel insights on early breast cancer events .

2.5 Conclusions

The in-depth characterization of early tumor initiation events in breast cancer, that give rise to tumor heterogeneity and early dissemination of metastatic cells, have not been fully characterized. The gap in our knowledge of these early initiation events is a result of lack of samples at this stage - a black box of patient material. Modelling the conditions of early tumor initiation using mouse models also fails to recapitulate the human situation, driving a tissue wide neoplastic phenomenon instead.

Adapting an already well characterized oncogene inducible mouse model of breast cancer, we developed an *in vitro* organoid system that is truer to the human situation and can activate oncogenes in single cells of the healthy organoid epithelium via lentiviral transduction. Characterization of this *in vitro* stochastic tumorigenesis system helped us observe that not all transformed single cells in the organoid epithelium had the capacity to proliferate and establish tumorigenic foci. Some cells resisted proliferation over many days of oncogene exposure unable to expand despite confirmed onco-protein localization.

In an effort to better understand this dual phenotype of single transformed cells we then established live cell imaging of these stochastically transformed organoids in 3D over 3-4 days using light sheet microscopy. Overcoming light sheet data limitations, we also implemented an image analysis workflow that allowed us to track single cells as they evolved within the organoid. Feature analysis of of over 150 transformed cells in 20 organoids showed that the cells in clusters were more likely to form tumors than isolated oncogene-expressing cells.

These results point to a proximity controlled signalling network between single transformed cells that is imperative to the growth and establishment of tumors within healthy epithelia. As a result, this study provides a whole new perspective on cell of origin of tumors, postulating rather a cell cluster of origin governed by signals from its healthy and neoplastic neighbours. What these signals are and how they are transmitted can now be explored using this model. Clearly, the role of the neighbouring healthy epithelial cells via cell polarity mechanisms needs further investigation as does the transcriptomic profiles of all cells in the cancer initiation niche, so that we can better understand early tumor events in the breast.

CHAPTER 3

QUEST FOR THE ACHILLES' HEEL IN BREAST CANCER RESIDUAL DISEASE

3.1 Introduction

3.1.1 Minimal Residual Disease

Definitions from the lab and the clinic

In the present age of precision medicine, when molecular dependencies of various tumours are being unraveled by sequencing and functional assays, a panoply of therapeutics that target kinases, transcriptional modifiers, immune checkpoints and other cancer vulnerabilities are being used in the clinics with excellent short-term prognosis. Consequently, the response to therapy has greatly improved for patients whose tumours have been profiled extensively, such as those with chronic myeloid leukaemia (CML), KIT proto-oncogene receptor tyrosine kinase (KIT)-mutated gastrointestinal stromal tumours [102], epidermal growth factor receptor (EGFR)-mutated or ALK receptor tyrosine kinase (ALK)-rearranged lung adenocarcinoma [103] [104][105] and BRAF-mutated melanoma [106][107]. Most patients with these diseases now achieve what is known as a complete remission (CR). CR in patients with cancer is traditionally defined as the absence of a visible tumour by use of sensitive radiological imaging (PET, MRI or CT). This means, CR, whether achieved by chemotherapy, targeted therapy, radiation, surgery or a combination, typically requires >99% (that is, >2–3 log10) reduction in tumour burden [108][109]. In a hypothetical patient with five metastatic lesions averaging 2 cm^3 each, this would equate to a reduction from approximately 10^{10} tumour cells to $<10^8$ tumour cells [110]. This implies that despite the high sensitivity of various cancers to treatment, some residual cancer cells persist anyway. These persistent cells have been reported to lead to tumour recurrence and treatment failure. As such, this cell reservoir surviving these precision therapies is called "minimal residual disease". These residual cancer cells can persist locally, in the bloodstream as circulating tumour cells or in distant organs as disseminated tumour cells [111]. But just because cells remain in the patient body after therapy, does not necessary mean they will cause relapse. In some cases, residual cells share phenotypic similarity (for example, histologic appearance and lineage markers) and genetic heritage (for example, truncal mutations and chromosomal rearrangements) with the original tumour cells, but they are not fully malignant [110]. It is therefore probably only the "malignant" reservoir of residual cells that harbor somatic alterations and/or phenotypic alterations resembling a tumour that need to be targeted to prevent tumor recurrences. This is the therapeutic challenge of the precision age.

Mechanisms of establishment of minimal residual disease.

Based on its definition using the complete remission(CR) theory mentioned above, the residual disease cells left behind at CR stage may seem an arbitrary selection – stochastically unaffected by the therapy method. However, breast cancer minimal residual disease studies in mouse models [112] for BRCA1-deficient breast cancer, show that repeated

treatments with the drug could not limit the residual cell reservoir completely even at titrated doses and higher therapeutic duration. This hints towards there being defined biological mechanisms that favor the survival of residual tumour cells.

Early observations in the clinical situation suggested there was a subset of tumour cells that had somehow mutated and was insensitive to drugs. These insensitive cells were postulated to be the only survivors of therapy. This was supported by findings in lung cancer patients [113] where residual cells developed a mutant allele of the oncogene being targeted. Another common mechanism of therapy evasion was the activation of an alternative signaling pathways as decribed in great detail in Niederst and Engelman, 2013 [114]. These mechanisms were coined as "secondary resistance" mechanisms and are still being unraveled today. For example, mutant BRCA1 or BRCA2 breast and ovarian cancer relapses have been found to have secondary mutations of the BRCA1/2 genes that apparently restore BRCA1/2 function [115][116]. These relapse tumours with secondary resistance no longer respond to the previously successful Poly (ADP-ribose) polymerase (PARP) Inhibitors [117][118] [119]and DNA cross-linking drugs [120].

Indeed, the identification of secondary resistance mechanisms is of extreme importance to develop therapeutic interventions for late stage patients. But therapy evasion cannot be the only mechanism used by cells to survive in a tumour under attack. This is proven by clinician statements that patients with recurrent disease respond again when the initial therapy is repeated. These cases cannot be explained by the "drug insensitive cell subset" theory of residual disease establishment. These observations point to cell intrinsic molecular mechanisms – for example, entering senescence to evade drugs targeting proliferative cells — that may cause residual disease in these patients. In addition, dynamics of the tumour micro-environment such as fibroblast signaling, immune cell intervention, extra-cellular matrix effects and the vascular network may play a role. Therefore, residual disease establishment mechanisms may be attributed to cell intrinsic and cell extrinsic factors [121]

Cell-intrinsic mechanisms of establishment of minimal residual disease.

CSC conferred resistance

The cancer stem cell (CSC) tumour model postulates that there is an inherent hierarchy within tumours[122] – growth depends on a few cancer stem cells that are capable of self-renewal. According to this model these self-renewing CSCs are inherently more drug- or radiotherapy-resistant and are therefore responsible for local tumour recurrence and distant metastases appearing after initial treatment [123][80]. Pro-survival mechanisms in CSCs include: elevated apoptosis resistance, drug-efflux pumps, enhanced DNA repair efficiency and epithelial to mesenchymal transition (EMT)[124]. Another interesting observation is the increased re-population rate of CSCs during radio- or chemotherapy [125]. Following initial therapy response, an accelerated re-population during successive cycles of therapy causes tumour relapse in the absence of any change in the intrinsic sensitivity of the tumour cells. Accelerated growth of tumour-repopulating cells here is attributed to re-oxygenation of cells upon reduction of tumour burden or the release of prostaglandinE2 (PGE2) from tumour cells killed by chemotherapy [126].

Metabolic dependency Studies in an inducible mouse model of pancreatic ductal adenocarcinoma (PDAC) has shown that the residual cells were less glycolytic and more dependent on mitochondrial pathways for energy production. Accordingly, surviving cells showed high sensitivity to oxidative phosphorylation inhibitors, which successfully blocked tumour relapse in this model. Intriguingly, residual cells had a distinct metabolic profile and this was considered critical for the survival of these cells[127].

Autophagy

Autophagy or "self-devouring" is the natural, regulated mechanism of the cell that removes unnecessary or dysfunctional components [128]. It allows the orderly degradation and recycling of cellular components. Surprisingly, this process has been shown to help the establishment of residual disease in various cancers. For example, in an osteosarcoma xenograft model, the dormant state correlated with enhanced autophagy and the presence of extracellular glutamine [129]. Inhibiting autophagy or depleting glutamine was sufficient to increase chemotherapeutic sensitivity in this model for osteosarcoma. Since autophagy is executed by the activity of lysosomes, sequestration of hydrophobic weak base chemotherapeutics in the lysosomes may be the mechanism of therapy evasion [130].

Cell Senescence

Therapy Induced senescence (TIS) is another concept that could explain MRD establishment [131]. It is frequently seen in the form of giant polyploid cells that occur at the

primary site after initial intervention. These cells are formed because of cancer cells that enter a cell cycle arrest by suspending their mitotic activity while maintaining DNA duplication [132]. The DNA endoreduplication generates polyploid cells which can survive for weeks as non-proliferating mono-or multi-nucleated giant cells. A study also showed that viable progeny can arise from polyploid TIS cells in a process that strikingly resembles that of daughter cell budding exhibited by many protozoans. This could be because of an increased expression of cyclin-dependent kinase 1 (CDK1)[133].

Chromatin remodelling mediated drug tolerance

Studies done in non-small-cell lung cancer (NSCLC) cell lines with tyrosine kinase inhibitors (TKIs), proved the existence of a surviving population of cells that survived drug concentrations 100-fold above the IC_{50}. Gene expression profiling of these persister cells revealed KDM5A/Jarid1A and a histone H3K demethylase to be up-regulated. These genes have functions spanning chromatin modification and the authors compared them to persisting bacterial cells that survive antibiotic therapy. Perhaps the comparison is valid, considering that the bacterial survivors were a result of stochastic gene modifications that allowed the population subset to survive. Although, persistence in cancer cells could be due to many individual single cell mechanisms [134].

Cell-extrinsic mechanisms of establishment of minimal residual disease.

The profound effect of the tumour micro-environment during all stages of tumourigenesis and therapy has been investigated for decades, so naturally, the effect of the tumour micro-environment on the establishment of minimal residual disease has also been revealed to a large extent. Micro-environmental factors like oxygenation, interstitial fluid pressure, acidity and proximity to fibroblasts or immune cells vary greatly in the context of a large solid tumour and as such may result in gradients in cell proliferation and more importantly a variation in the efficacy of anticancer drugs[135]. More specifically, cancer cells or cells within the tumour micro environment have been shown to secrete growth factors [136] or tyrosine kinase ligands that then activate pathways mediating cell survival and drug resistance of cancer cells[137].

The concept of cancer immunoediting [138] probably best describes the effect of the immune system on the tumour micro-environment and growth, unraveling how cancer cells use this to survive therapeutic intervention. The sequential phases of immune editing include elimination followed by establishment of equilibrium and eventually escape. In the elimination phase, tumour cells are destroyed by cytotoxic CD8+ T lymphocytes (CTLs) by inducing tumour cell lysis [139]. Despite this efficient clean-up of transformed cells, some rare cells survive in a dormant state [140]. The theory behind their survival rests on the fact that they evade the immune system via molecular changes and this allows them to lay dormant over a long period of time (equilibrium). As a consequence of constant immune selection pressure on genetically unstable tumour cells, changes that favor growth of the surviving tumour cells are eventually selected. These are no longer recognized by the immune system and thus "escape" immunosurveillance.

Another crucial component of the micro-environment that effects tumour progression is the vasculature. It provides the growing tumour mass with nutrients and oxygen and as such can cause profound effects if it grows abnormally or does not penetrate the tumour mass efficiently. Abnormal vasculature may result in a dynamic metabolic landscape for cancer cells due to tumour patches where there is hypoxia and malnutrition. This kind of "angiogenic dormancy" can prevent tumour growth [135]. Studies have shown that apart from the diffusion of nutrients and gases, the clearance of acidic metabolites is also a vital function of the vasculature [141]. Decreased clearance at distal tumour lobules that are not vascularized efficiently can cause pH changes that in turn effect the activity of anti-cancer drugs [142] aiding therapy evasion. Another cellular mechanism used by the surviving population of cancer cells with regards to vasculature patterns and hypoxia, is the activation of oxygen sensitive transcription factors – HIF1 and HIF2. These transcriptional modulators activate multiple programs to combat apoptosis, including induction of autophagy as the ultimate survival process before cell death [70].

Following this discussion of the many cell-intrinsic and cell-extrinsic mechanisms, used by residual disease cells to persist at the site of tumour after therapy, it is probably

not completely unreasonable to hypothesize that multiple of these mechanisms may be employed by different cells in the same tumour. This variation could depend on factors like the type of tumour, the location of the tumour and the micro-environment it grows in. Encompassing this variability however, is one of the central mechanisms used by residual cells to survive – tumour dormancy. All cells surviving initial insult therapy lay dormant for a variable period (ranging from months to decades) of time before re-establishing malignant disease [143][144] at the initial site or at a distant metastatic site.

Tumour dormancy is a phenomenon observed in clinical practices which refers to a temporary halt in tumour progression, with a prolonged latency. This means, the patient is rendered disease-free during the tumour dormancy period and complete remission (CR) is coupled with absence of clinically detectable tumour recurrence[145]. The presence of tumour dormancy in patients has been shown in many cases of breast cancer[143], prostate cancer [144], melanoma renal cell carcinomas [146] and other tumour sub-types. Recurrences in these patients are sometimes seen after decades of disease-free survival.

Tumour dormancy can be on a cellular level or more on a whole tumour mass scale. tumour mass dormancy happens due to the factors that constrain tumour expansion despite cell proliferation. The balance between cell division and apoptosis is achieved through the conditions in the microenvironment, be it the lack of nutrients and oxygen or immune system responses that efficiently eliminate the tumour cells [147]. The cellular dormancy, on the other hand, is a true mitotic arrest in the tumour cells. Residual disease establishment is quite possibly a combination of both these dormancy inducing processes. In this project, I aim to explore the cellular dormancy mechanisms that cause residual disease so as find ways to target the cellular reservoir of dormant cells that will revert to refractory disease eventually.

Rationale for targeting regressing tumours and the residual disease state

Unraveling so many mechanisms by which residual cells can survive leads to hypotheses that targeting the establishment and persistence of residual cells may be the ultimate way to cure cancer –eradication of all tumour cells capable of driving relapse. In the clinic, targeting residual disease is a logical choice due to the following main reasons:
First, there is a reduction of tumour burden at this stage. This confers many advantages to the treatment of residual disease, for example, a presumably smaller number of treatment refractory clones persist at this point. This implies that there is a reduced likelihood of sub clonal resistance to one or more therapeutics [148]. Reduced tumour burden also weakens the actual cancer cells that have survived. They are fewer in number and less effective at remodelling microenvironments, reprogramming infiltrating hematopoietic cells and inducing chemoprotective niches [149]. As a result, certain drugs may have greater efficacy against MRD than against the same cancer at the time of clinical relapse.
Second, if the drugs for treating the residual cells are toxic and lead to many side effects (as they commonly do), a patient treated for residual disease eradication at this stage has a better chance of survival than for example in the relapse tumour stage, when the body is under active tumor stress.

Despite these many mechanistic and practical reasons mentioned above, the most convincing evidence that targeting residual disease may be the best option for pateints with cancer, is the proof of principle published by researchers and technicians using neoadjuvant therapy. Assuming the prevention of relapse, is a measurable surrogate for cure, patients following resection of epithelial tumours and sarcomas, treated with adjuvant (and to some extent neo-adjuvant) therapy intended to eradicate MRD achieve long-term disease-free survival with the combination of surgery and adjuvant (or neoadjuvant) therapy than with surgery alone [150] [151] [152] [153] [154] [155] [156]. Similarly, for haematological cancers such as acute leukaemias or aggressive lymphomas, a single cycle of intensive chemotherapy can induce complete remission , but virtually no patients are cured without additional therapy to eradicate residual disease[157]. Additional therapy aimed at residual disease in combination with targeted tumor therapy could thus have the potential to limit relapse tumor prevalence and the concept of synthetic lethal interactions within tumour cells could help us identify the most promising combinational therapies to achieve this goal, as discussed in the following chapter.

3.1.2 Synthetic Lethality

Synthetic lethality was discovered by fruit fly researchers in the early 20th century[158]. They discovered that, certain non-allelic genes were lethal only in combination, even though the homozygous parents were perfectly viable. Synthetic lethality is thus defined as an interaction that occurs between two genes when a perturbation (a mutation, RNA interference knockdown or inhibition) that affects either gene alone is viable but the perturbation of both genes simultaneously is lethal [159].

The question of why synthetic lethal interactions occur is probably best answered by studying biological processes like evolution, developmental canalization (ability to produce the same phenotype regardless of genotype) and cancer (also other multifactorial diseases)[160][161]. The need to maintain homeostasis in the face of diverse genetic and environmental challenges in these processes, frequently leads to synthetic lethal interactions. These interactions make the organism or process more genetically robust[162] [163].

Robustness is mostly provided by two well studied mechanisms of buffering:

1. Via non-homologous genes operating in the same cellular process or in back-up pathways[164].

2. Via "capacitors" like heat shock proteins and chromatin regulators that can mask the effects of many different mutations [165] [166].

Thus, cellular systems maintain homeostasis partly by ensuring that processes do not depend on any single component, which could easily be perturbed by mutation or environmental effects, setting the scene for synthetic lethal interactions[167].

Exploiting synthetic lethal interactions to identify new anticancer drug targets has been pursued by researchers for over 20 years now [168]. In cancer cells, the process of "oncogene addiction" is well studied and exploited for therapy by targeting the oncogenes that the cells get addicted to[169]. But although small-molecule and antibody-based inhibitors of oncogenes have proved to be effective for some tumour genotypes, not all tumours have targetable gain-of-function oncogenes, and therapeutic resistance is a common outcome.

However, these cancer cells got through a lot of genotoxic stresses and have been shown by us in the lab and others, that they have distinct profiles from their wild type counterparts[52][127]. This points to a "re-wiring" process that could expose new genetic vulnerabilities. Synthetic lethal interaction partners of cancer-associated molecular changes – "non-oncogene addiction" interactions – have therefore become of great value for therapeutic opportunities [170].

The protein product of a gene that has a synthetic lethal interaction with a frequently occurring tumour-specific somatic mutation would be an excellent anticancer drug tar-

get, because a therapeutic that exploits the synthetic lethal interaction should result in favourable therapeutic indices, in which only tumour cells that harbour the mutation would be sensitive to the therapeutic (Figure 3.1) .

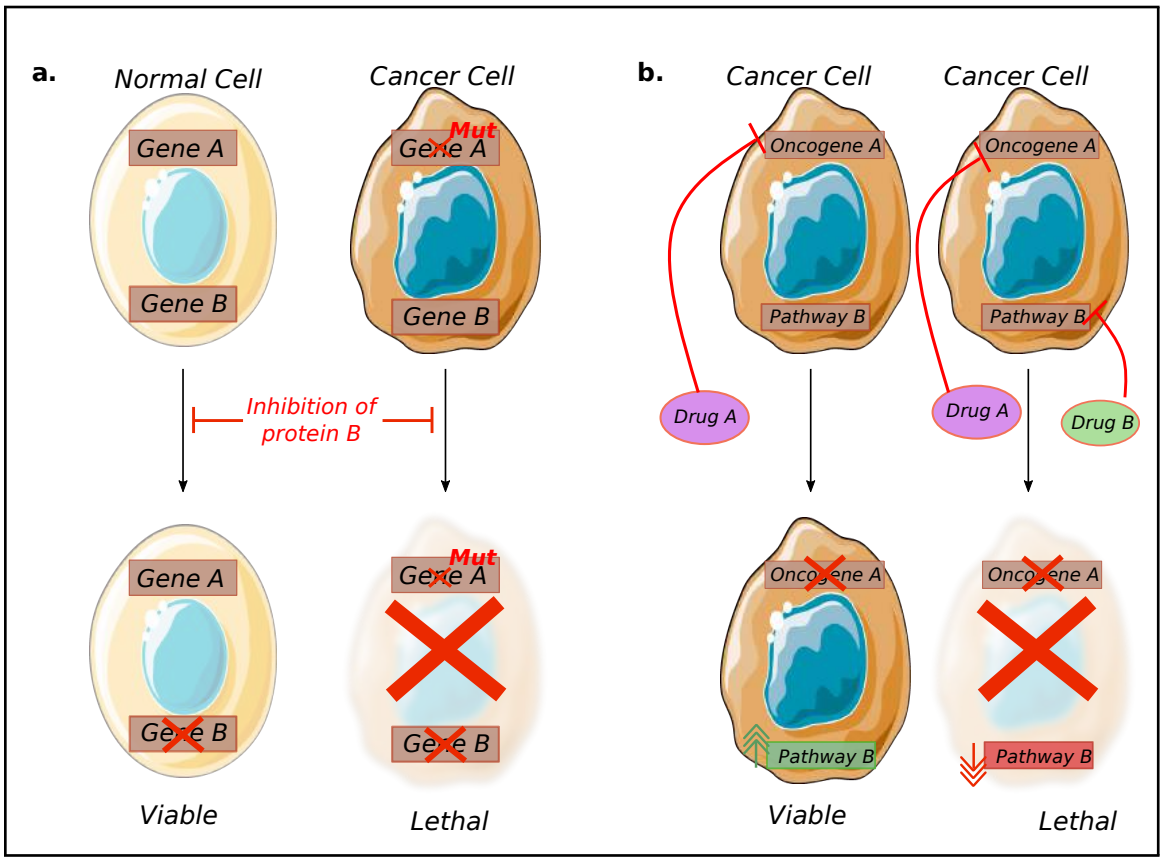


Figure 3.1: Exploiting synthetic lethality for anti-cancer therapy

One of the early and possibly still the best examples of this kind of therapy is seen in the case of DNA repair enzyme called PARP that was proved to be synthetic lethal with DNA repair genes BRCA1 and BRCA2, mutations in which can cause breast and ovarian cancer. tumours of patients carrying these mutations could be successfully treated using a chemical PARP inhibitor with remarkably mild side effects [171] [172].

3.2 Objectives

The main aim of this project was to better understand the molecular basis of residual disease in mouse mammary carcinoma and gain insight into the apoptotic pathways activated in the regressing tumour mass as well as understand the senescence, growth arrest and repolarization pathways activated in the residual cells. The goal was to understand the survival mechanisms invoked by residual cells and target the key molecular players involved in them. This would allow us to identifying potential synthetic lethal targets for clearing residual disease.

To achieve this aim, we established the following objectives for this study:

1. Characterize the *in vitro* reductionist model of the TetO-MYC/ TetO-Neu/ MMTV-rtTA organoids during tumour regression and analyze the kinetic transcriptome signature during this process.

2. Characterize the analogous *in vivo* path to residual disease in TetO-MYC/ TetO-Neu/ MMTV-rtTA mice with induced tumours and establish reproducible timelines for tumour progression and regression both *in-vitro* and *in-vivo*

3. Employ an unbiased data analysis approach to identify genes specifically up-regulated during regression to rescue surviving residual cells both in the reductionist *in vitro* system and immune proficient mouse model. If successful, this objective extends to testing the efficacy of inhibiting these targets using an appropriate experimental setup.

4. Perform live cell imaging of regressing tumours, *in vitro*, using Selective Plane Illumination Microscopy (SPIM) in a more biased approach, to understand regression dynamics and draw inferences on potential survival mechanisms used by cells reintegrating into the residual rims.

3.3 Results

3.3.1 Characterization of minimal residual disease *in vitro* using primary organoids

The TetO-MYC/ TetO-Neu/ MMTV-rtTA mouse model, described in detail in the Introduction section of this report, was employed to obtain a correlate of minimal residual disease *in vitro*.

3D matrigel cultures were established from single cells, isolated from the mammary glands of adult TetO-MYC/ TetO-Neu/ MMTV-rtTA female mice. Subsequent to 4-7 days growth in embedded matrigel matrix, these single cells grow to resemble acinar-looking, polarized structures with a central lumen (Figure3.2a, left panel). Continuing on, these hollow acinar structures without any exposure to oncogenic proteins will be referred to as the "Normal" or "Never Induced" acini. Owing to their resemblance in structure and function to the basic ductal units of mammary glands, they are considered the healthy mouse correlate of the healthy human breast morphology. The left panel in Figure 3.2b shows this stage in an immunofluorescence staining of 3D cultures with the key polarity markers : Alpha 6 integrin (*ITGA6*, shown in red) has a basal localization; Zonula Occludens (*ZO-1*, shown in green) is apical, staining occludins at the tight junctions and the Golgi matrix protein (*GM-130*, shown in magenta) is restricted to apical part of the cell, marking the Golgi.

Upon introduction of doxycycline into the culture media, bright field microscopy shown in Figure 3.2a, middle panel shows the establishment of a tumourigenic phenotype after five days of doxycycline exposure. The activation of oncogenes c-MYC and *Neu* in the hollow acinar cells results in high proliferation rates and crowding of cells in the lumen. Subsequent breaking of polarized epithelium phenotype results in filling of the lumen and expansion of these acini to several times their original dimensions. Several lobules are formed and the growth is prolific, accompanied by a higher apoptosis rate, as is commonly observed in proliferating tumour cells. This tumourigenic acinar state is referred to as "5 Days ON DOX" and is the correlate of "Tumour" in human disease. The immunofluorescence staining of 3D cultures at this this stage vary greatly from what is seen in the "never induced" state as seen in the middle panel of Figure 3.2b.

Owing to the principle of oncogene induction discussed earlier, withdrawal of doxycycline from the media results in rapid death and lumen clearing of the "tumour" acini. Figure 3.2a, right panel shows the resulting acinar phenotype "7 Days OFF DOX". Despite oncogenic exposure, the cells remaining in the reintegrated rims at this stage could be compared to the human correlate of "Minimal Residual Disease" and will be alluded to as such in this study. As shown by the IF staining of 3D cultures at this stage, residual disease structures are phenotypically very similar to "never induced" state as seen in the Figure 3.2b, right panel.

The advantageous use of 3D cultures established from primary murine mammary ep-

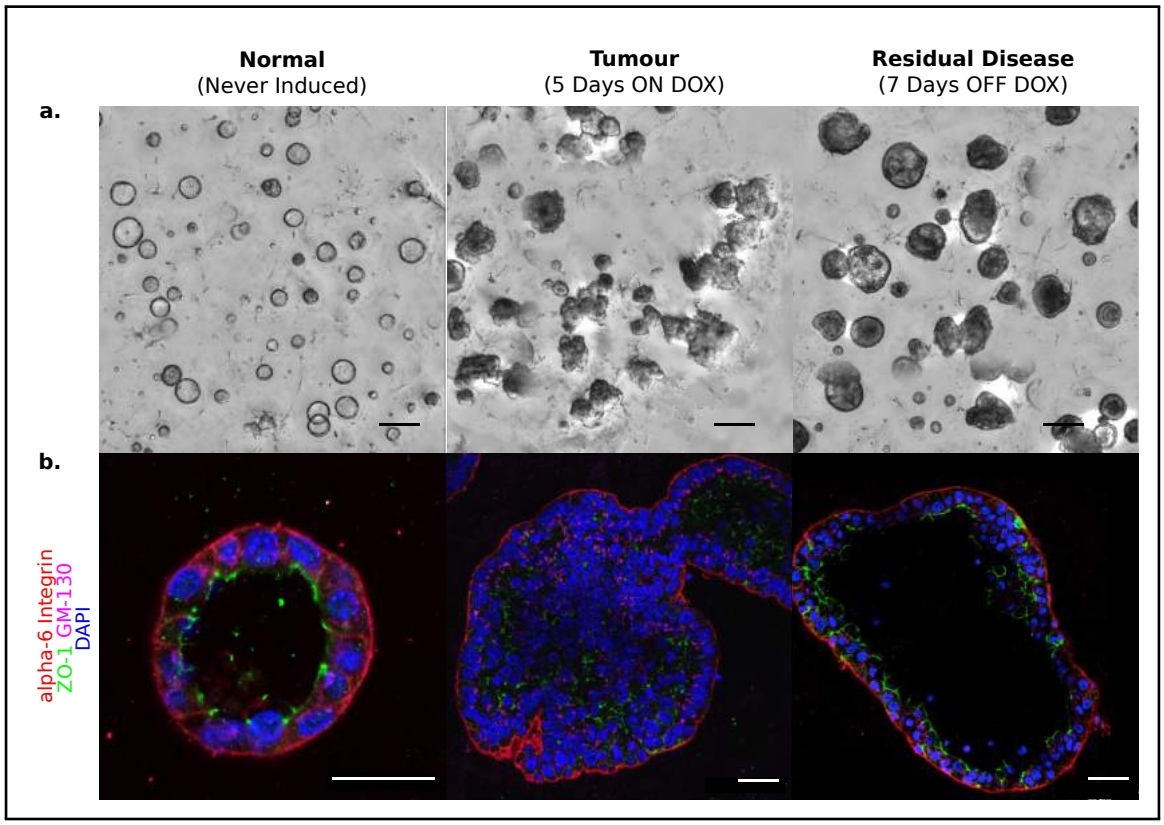


Figure 3.2: Representative (a) Bright field microscopy images (Scale bar, 100 μm) and (b) Immunofluorescence staining images (Scale bar, 25 μm) of fixed 3D gels with TetO-Myc/TetO-Neu/MMTV-rtTA organoid culture. Left panels show the normal mammary acini grown from primary cells harvested from the mice. Middle panels show the tumour organoids that develop post induction with doxycycline for 5 days in culture. Right panels show the regressed tumours after 7 days of doxycycline withdrawal. Polarity markers include alpha-6-Integrin (red), GM-130 Golgi marker (magenta), ZO-1 (green) and DAPI nuclear stain (blue).

ithelial cells allowed close monitoring of the c-MYC and *Neu* expressing tumour phenotype. This included observing initiation and regression closely under the microscope and establishing the timeline for oncogene activation and dependence during tumour formation and similarly "oncogenic shock" upon acute onco-protein inactivation. Immunofluorescence staining experiments utilizing 3D gels harvested and fixed at multiple time points during the kinetic processes of apoptosis, repolarization and establishment of residual rims during the tumour regression phase have shed much light on the molecular characteristics of these cells over the seven days of *in vitro* regression. After 5 days of induction with doxycycline media, oncogenic protein degrades within 12 hours of de-induction. The inner tumour mass undergoes apoptosis rapidly and re-establishment of the polarized epithelial rim is seen within 72 hours. The surviving rim enlarges and apoptosis is near complete by day 7 after removal of doxycycline from the media.

Detailed characterization of the regression phase after de-induction of oncogenes has

allowed for better characterization of the cellular events that occur in this phase: oncoprotein depletion (6-8 hours after de-induction), apoptosis (as early as 8 hours after de-induction), clearing of the tumour core (24-36 hours after de-induction), re-polarization of epithelial cells around a central lumen (36-72 hours after de-induction) and establishment of a dormant residual rim (by 7 days after de-induction). Figure 3.3 shows representative IF stains demonstrating these findings.

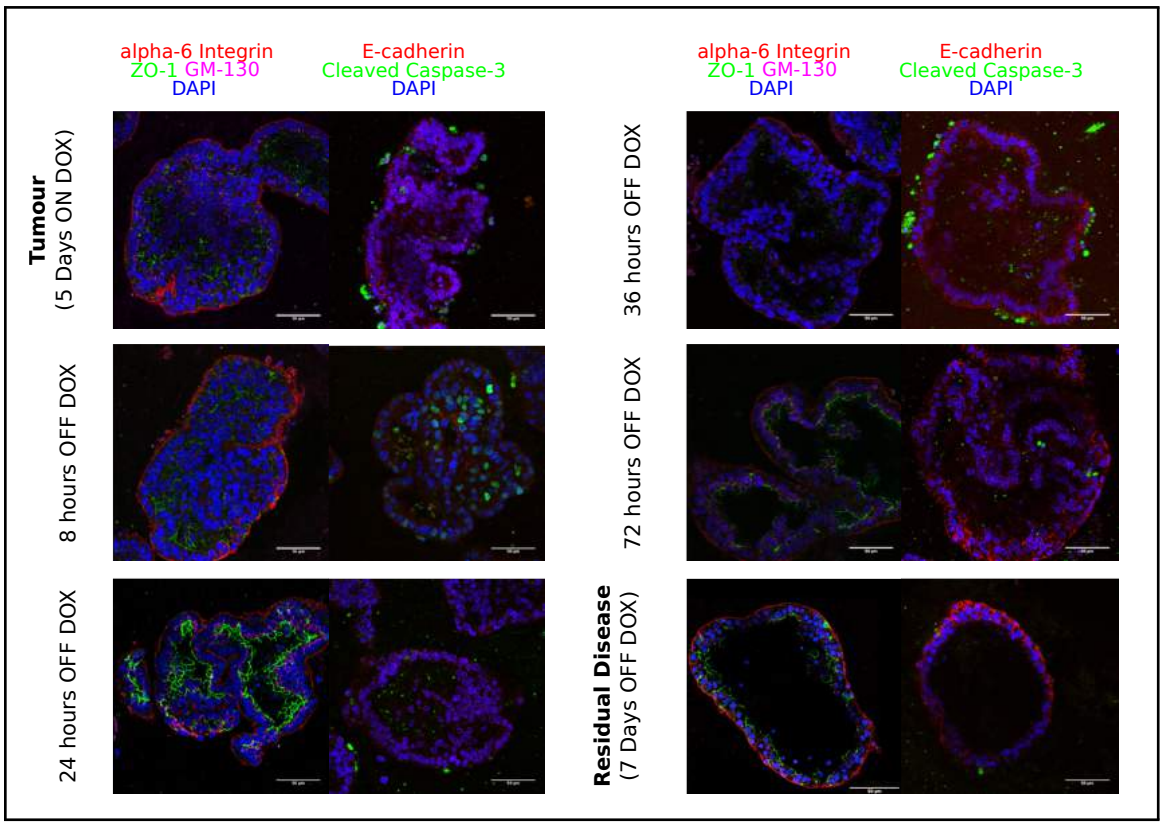


Figure 3.3: Representative Immunofluorescence staining images of fixed 3D gels with TetO-Myc/TetO-Neu/MMTV-rtTA organoids harvested during tumour regression(upon doxycycline withdrawal). Left panels show polarity staining with markers including alpha-6-Integrin (red), GM-130 golgi marker (magenta), ZO-1 (green) and DAPI nuclear stain(blue). Right panels show cell death with the help of the cleaved Caspase-3 stain (green). E-cadherin(red) is used as membrane marker and DAPI is the nuclear stain(blue). Scale bar, 50 μm.

3.3.2 RNA sequencing analysis of *in vitro* organoids during the establishment of MRD

Equipped with the aforementioned findings, the objective to unravel the kinetic transcriptome during tumour regression and minimal residual disease establishment was furthered, in that we could now ascertain the hallmark timepoints at which "transcriptomic snapshots" would be worth exploring in these *in vitro* cultures. We established organoid cultures from single cells isolated from the mammary glands of four different adult TetO-MYC/ TetO-Neu/ MMTV-rtTA female mice, induced and then de-induced oncogenic expression in the resulting mammary acini and collected RNA samples at the hallmark points on the regression time-scale (Figure 3.4).

Cdh1(E-cadherin membrane marker) and c-MYC immunofluorescence staining performed early during tumour induction with doxycycline, revealed that the expression of oncogenes is relatively quick, already visible with c-MYC co-localization in some acinar cells at the 4 hours ON DOX timepoint (Figure 3.4, right upper panel). To capture the very early tumour signaling networks, we included 1, 2 and 4 hours ON DOX timepoints into the scheme. Similarly, *Cdh1* (E-cadherin membrane marker) and c-MYC immunofluorescence staining at early timepoints after doxycycline depletion from the culture media, showed that the oncoprotein had a very short half-life and the protein was no longer visible as early as 8 hours OFF DOX(Figure 3.4, right middle panel). Coupled with a *Cdh1* (E-cadherin membrane marker) and cleaved *Casp3* (Caspase-3 apoptotic marker) immunofluorescence stain at the same time point (Figure 3.4, right lower panels), we see that Caspase-3 has already been cleaved in some cells implying that the depletion of the oncogene and decision to apoptose has already been taken at this point. To gain as much insight into the transcriptomic signatures during this important "survive-vs.-apoptose" stage, we included 4, 6, 8, 10 and 12 hours OFF DOX timepoints into the scheme. Because all the transcriptomic snap shots would be compared to the "Never Induced" transcriptomic snap shot, the evolution of the transcriptomic signature of the normal acini were taken into account. Cultures grown without doxycycline alongside other cultures on doxycycline were collected at the 5 days ON DOX timepoint and similarly after 7 days at the 7 Day OFF DOX timepoint. I performed immunofluorescence staining in parallel 3D matrigel cultures for oncogenic, apoptotic and polarity markers to ensure the organoid characteristics did not vary too much from the expected. RNA samples were sequenced at the Genomics Core Facility at EMBL Heidelberg using NGS HiSeq protocol – 75bp read length, 30x coverage.

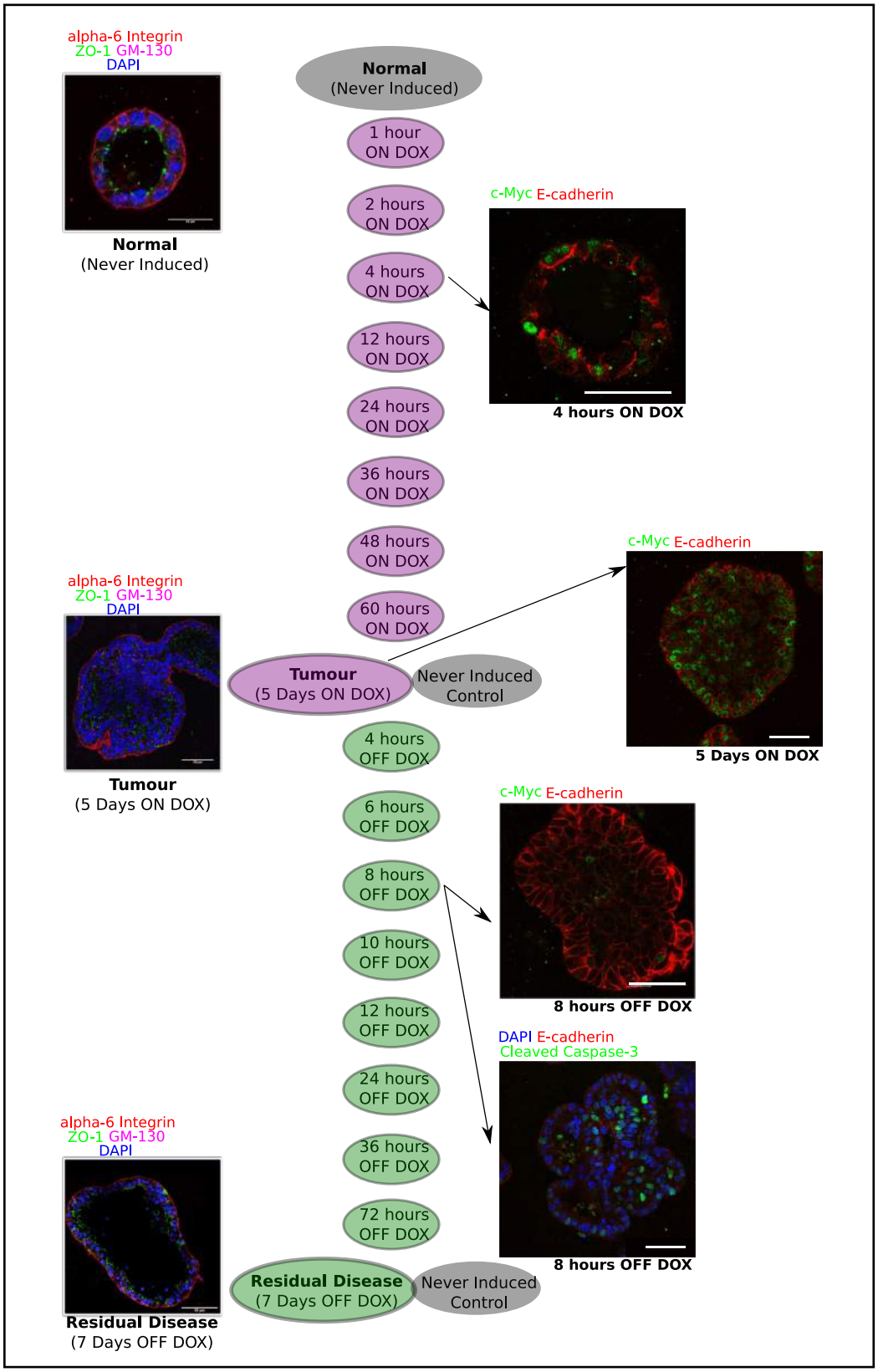


Figure 3.4: Hallmark time points during tumour progression (purple) and tumour regression (green). RNA was collected from *in vitro* matrigel cultures at these timepoints and RNA sequencing was performed to unravel the dynamic transcriptome. Scale bar, 50 μm.

Principal Component Analysis (PCA) plotting of data points across the time series and subsequent bioinformatics analysis was performed by Katharina Zirngibl from the Patil Group at EMBL, Heidelberg. The first and the second principal components accounted for 38% and 22% of the variance, respectively. It is apparent (Figure 3.5) that the "7 days OFF DOX" timepoint (represented by the darkest green dots in the green regression kinetic arrow) indeed had a distinctive transcriptomic profile that distinguished it from both the "5 days ON DOX" (represented by the darkest purple dots in the purple regression kinetic arrow) and "Never Induced" (shown in gray) timepoints. This clearly indicates that although the residual disease stage is phenotypically similar to the normal acini as seen in the staining experiments, there is, in reality, a significant difference between them on a transcriptomic level. These findings correlate with data published formerly by the lab showing a distinct transcriptomic profile between "Never Induced and "7 days OFF DOX" regressed acini on microarray chips. The residual disease substrate in this study was also found to have an increased lipid metabolism and elevation of ROS species [52].

DeSeq2 expression analysis showed that over 5300 genes were significantly dysregulated along the tumour progression and regression time course (compared to never induced control, q value0,05), with a false discovery rate (FDR) of 1%. This is far too large a number of genes to allow untangling of network connections using standard GO (Gene Ontology) enrichment analysis. It was therefore decided to perform clustering of the genes based on their expression trajectories along the time course.

Clustering would proved to have a twofold advantage :

1. Provide information regarding the broad gene expression trajectories during oncogenesis/regression and discussion on which trajectories might belong to potential synthetic lethal targets.

2. Serve as guidelines for classifying genes into fewer clusters for efficient network analysis.

Clustering of the differentially regulated genes (over 5000) across the time series using the GP Clust algorithm yielded 75 unique clusters (Figure 3.6). Each tiny graph in the Figure 3.6 represents the average gene expression trajectory of a subset of genes. The midpoint of the Y axis (at the marking '10') indicates the point of de-induction along the time course. Since the induction was performed for 5 days and de-induction for 7 days, the gene expression trajectories have been normalized to fit an equidistant axis. The gene expression levels all start at 0 (represented by the X axis) at the start and then either increase or decrease. An inflection point normally occurs at the center (at the timepoint when the organoids were de-induced) and not all gene expression values return to baseline.

In Figure 3.7, the grey clusters on the left represent gene expression clusters that seem to be upregulated (top) or downregulated (bottom) during tumour induction and then return to normal along the course of regression. Conversely, the other four colorful clusters in Figure 3.7 represent gene expression profiles that seem to be up-regulated temporarily or permanently in varying intensities and peaks during tumour regression. Closer inspection of the gene cluster trajectories is informative, therefore, in hypothesizing the gene clusters that are simply a consequence of oncogene activation and addiction from the gene clusters that seems to be upregulated during the establishment of minimal residual disease and could therefore be potential synthetic lethal targets.

The clustering of over 5000 differentially regulated genes into 75 unique clusters based on their expression trajectory, paved the path for us to manually curate the interesting trajectory clusters and has helped us classify interesting trajectories into 5 broad clusters:

1. Sustained targets – upregulated consistently during regression up to residual disease state

2. Early responders – upregulated early during regression before apoptosis decisions are made (up to 8 hours after de-induction)

3. Intermediate responders – upregulated after apoptosis decisions are made to potentially salvage residual cells/aid in tissue rearrangement to increase survival (12-24 hours after de-induction)

4. Late targets – upregulated during re-polarization of epithelial cells into residual rim (36 hours after de-induction onwards)

5. Shoulder trajectories – upregulated in tumours and continued sustained expression during regression (at least up to 34 hours after de-induction)

To better understand the gene networks linked with the expression of oncogenes, gene trajectories like the grey ones on Figure 3.7 were classified into "Expected Up" or "Expected Down" clusters. Genes in the "Expected Up" cluster seem to be upregulated during tumour induction and then return to normal along the course of regression. Genes in the "Expected Down" cluster, conversely, seem to be downregulated during tumour induction and then return to normal along the course of regression.

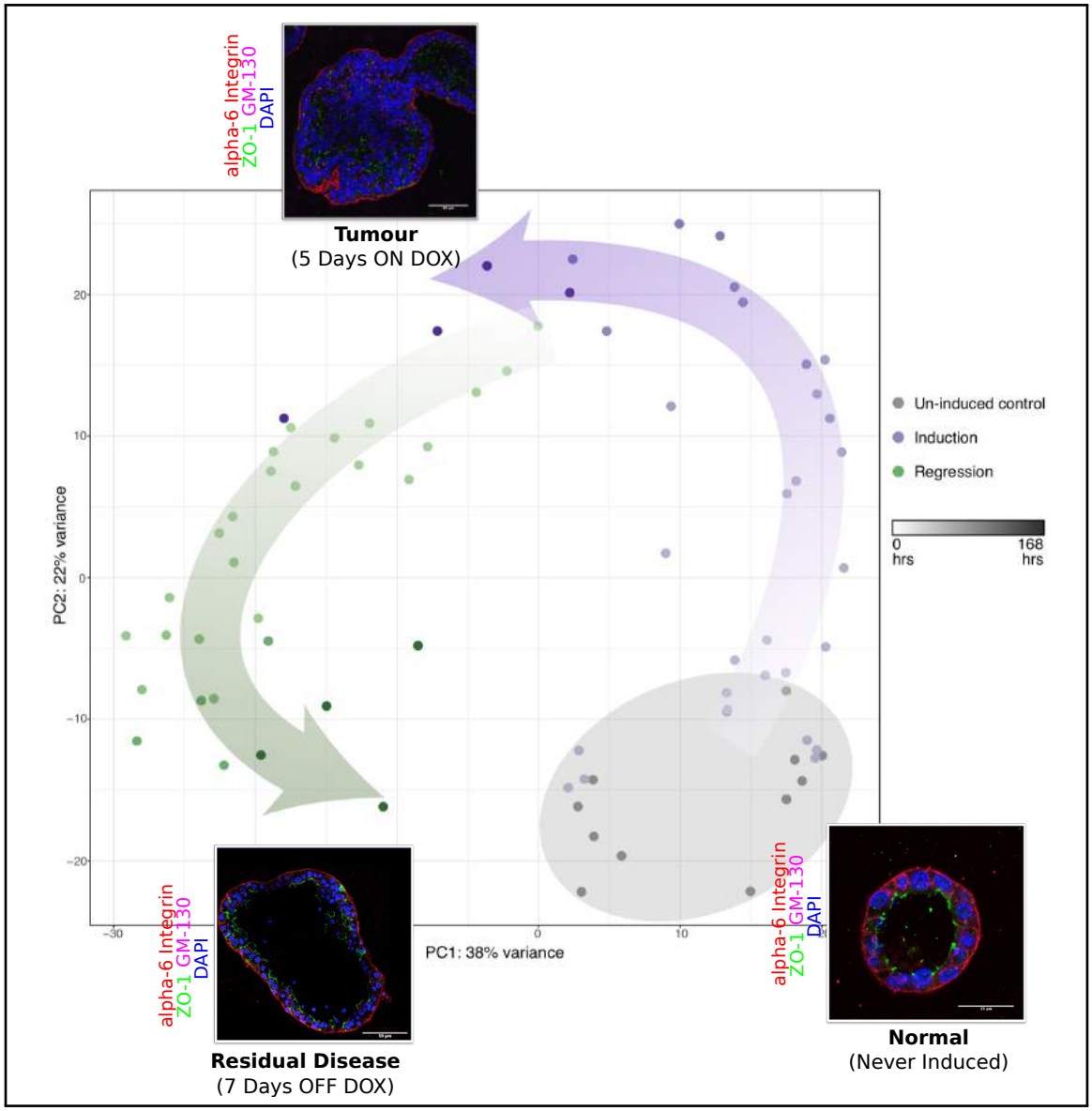


Figure 3.5: PCA plot for RNA Sequencing data of samples collected at hallmark timepoints during tumour progression and regression (See Figure 3.4). Every dot along the arrows is a data point for a biological replicate for the particular time point (arrows and points colored purple for tumour progression, green for tumour regression and grey for never induced controls). In-lay images show immunofluoresence staining for polarity markers at the 3 endpoints. Scale bar, 50 μm.

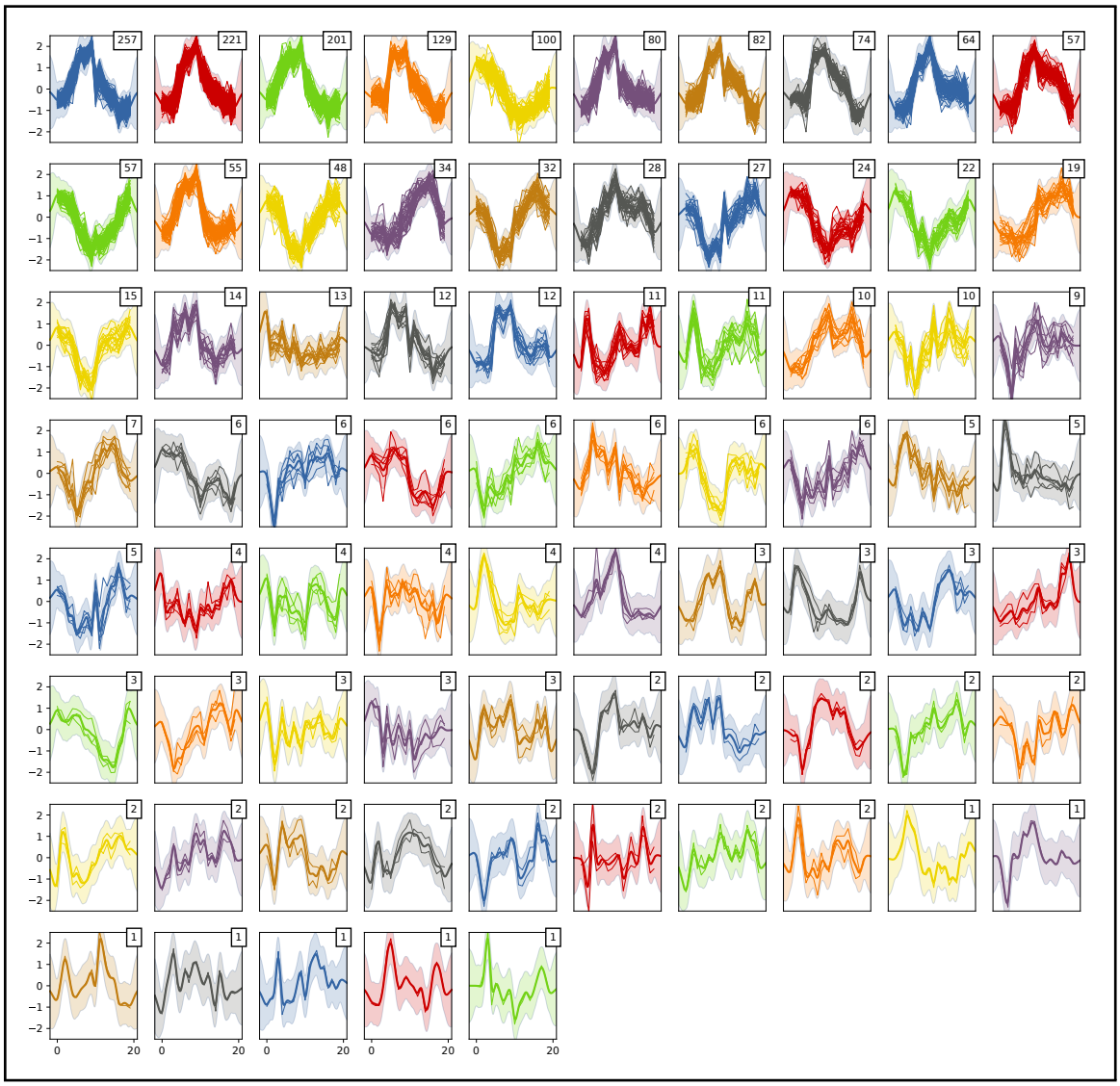


Figure 3.6: Genes found to be significantly deregulated along the time course shown in Figure 3.5, with a false discovery rate (FDR) of 1%, were clustered across the time series using the GP Clust algorithm to yield the 75 unique clusters shown in this figure.

The genes (number indicated in the box on the right-hand corner of each graph in Figure 3.6), that belonged to the clusters with trajectories that looked similar to the selected seven interesting trajectories, were pooled into manually curated clusters. Further, to prevent loss of gene targets that did not meet stringent quality criteria but show similar expression trajectories over the time course, we developed a correlation analysis algorithm. Only the trajectories from the five potentially "synthetic-lethal" clusters (Figure 3.8 a) were used to pull out other genes with similar trajectories from the whole data-set (Figure 3.8 b). Stringent filtering on fold change and significance of gene expression has yielded a total of 473 potential synthetic lethal targets among these 5 interesting trajectories. Including the "Expected Up" and "Expected Down" clusters the number of genes totals 1796 (Table 3.1).

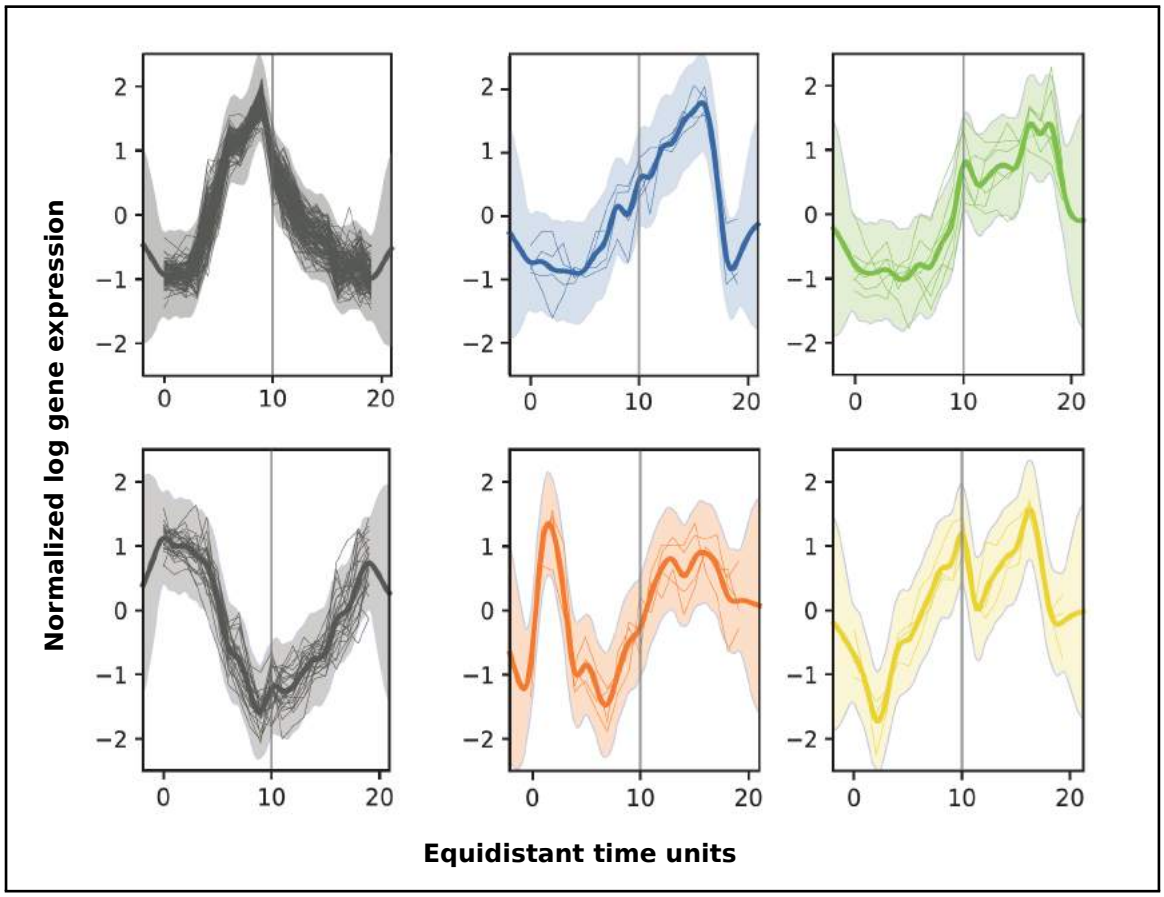


Figure 3.7: Example gene cluster trajectories from the 75 unique clusters (see Figure 3.6). Grey gene trajectories show gene expression deregulated due to oncogene expression. This aberrant expression seems to return to normal upon oncogene de-induction. Colored trajectories show genes with aberrant activity before and/or after oncogene withdrawal that are hypothesized to help residual disease cells survive tumour regression. Time units are normalized for graphical representation and grey lines at position 10 in each graph show the point of oncogene de-induction.

Collaboration with Matt Rogon, Center for Biomolecular Network Analysis, EMBL, Heidelberg was established so as to perform Gene Ontology Enrichment analysis of these 1796 genes in 7 manually curated clusters. Enriched GO terms for each of the 5 potentially synthetic lethal clusters is shown in Figures 5.3, 5.4, 5.5, 5.6 and 5.7. A combined multi-cluster gene enrichment analysis on all these potentially synthetic lethal gene clusters is shown in Figure 3.9. The pathway enrichment information from the five potentially synthetic lethal clusters – individually and in comparison – allowed us to narrow the list of molecular players vital to establishment of residual rims.

This was done in a two-step process: first, we manually picked interesting GO terms that seemed to be novel yet logical in the residual disease establishment landscape. These included GO terms shown in Table 3.2

Next, we mapped each of these networks based on protein-protein interaction databases.

Curated Cluster	FDR Clustering	Correlation Analysis	Total
Expected-Up	1087		
Expected-Down	236		
Sustained Target	18	109	
Early Response Target	33	9	
Late Response Target	11	86	
Intermediate Response Target	58	94	
Shoulder Target	31	24	
			1796

Table 3.1: Gene numbers and sources for manually curated and potentially synthetic lethal clusters that were used for gene enrichment and network analysis.

GO Term identifier	GO Term
GO:0002819	Regulation of adaptive immune response
GO:0098742	Cell-cell adhesion via plasma-membrane adhesion molecules
GO:0002009	Morphogenesis of an epithelium
GO:0001213	Formation of the cornified envelope
GO:0031663	Lipopolysaccharide-mediated signaling pathway
GO:0017017	MAP kinase tyrosine/serine/threonine phosphatase activity
GO:0000097	Downregulation of ERBB4 signaling
GO:0000618	Oxidative Stress Induced Senescence
GO:0008593	Regulation of Notch signaling pathway
GO:0000571	FAS pathway and Stress induction of HSP regulation

Table 3.2: GO Terms enriched in the five potentially synthetic lethal clusters manually curated from RNA Seq analysis.

As shown in Figures 5.8 to 5.17, we mapped the network proteins for each pathway and further expanded the net by including the first neighbor proteins of all the proteins in the particular network. First neighbors are proteins that have been shown, experimentally and/or theoretically, to interact in a cellular environment. Filtering the GO-term -protein + first-neighbor network based on availability of drugs for inhibition of gene/ protein target further helped in narrowing the list of potential synthetic lethal targets to interfere with minimal residual disease. Further, manual filtering for biological relevance was postulated to provide a comprehensive list of targets for validation.

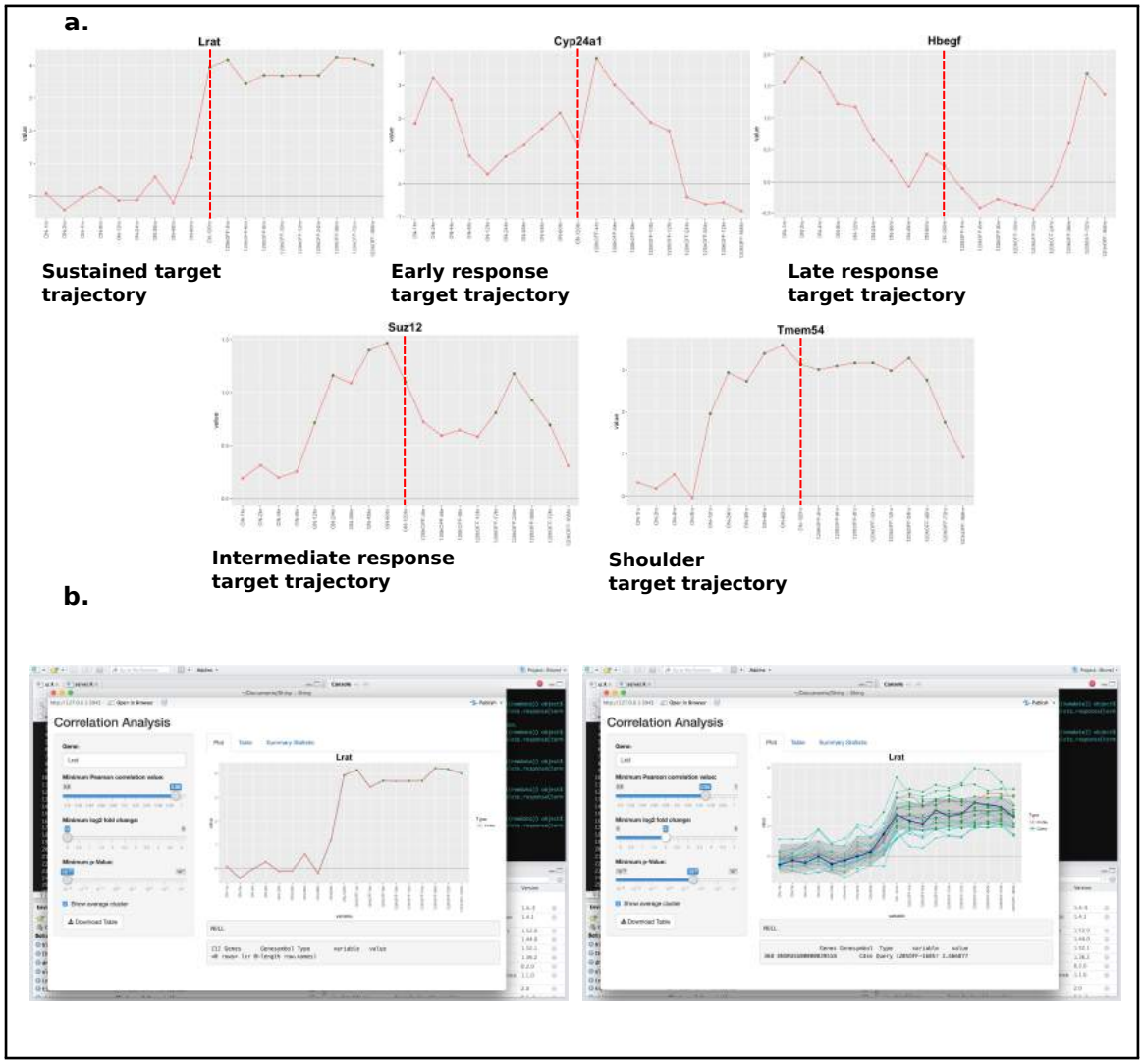


Figure 3.8: (a)Five broad, interesting gene expression trajectories that could belong to potential synthetic lethal targets – picked manually based on biological criteria. (b) Correlation analysis algorithm used to pull out genes from the whole RNA Seq data set to enrich single gene trajectory clusters in (a)

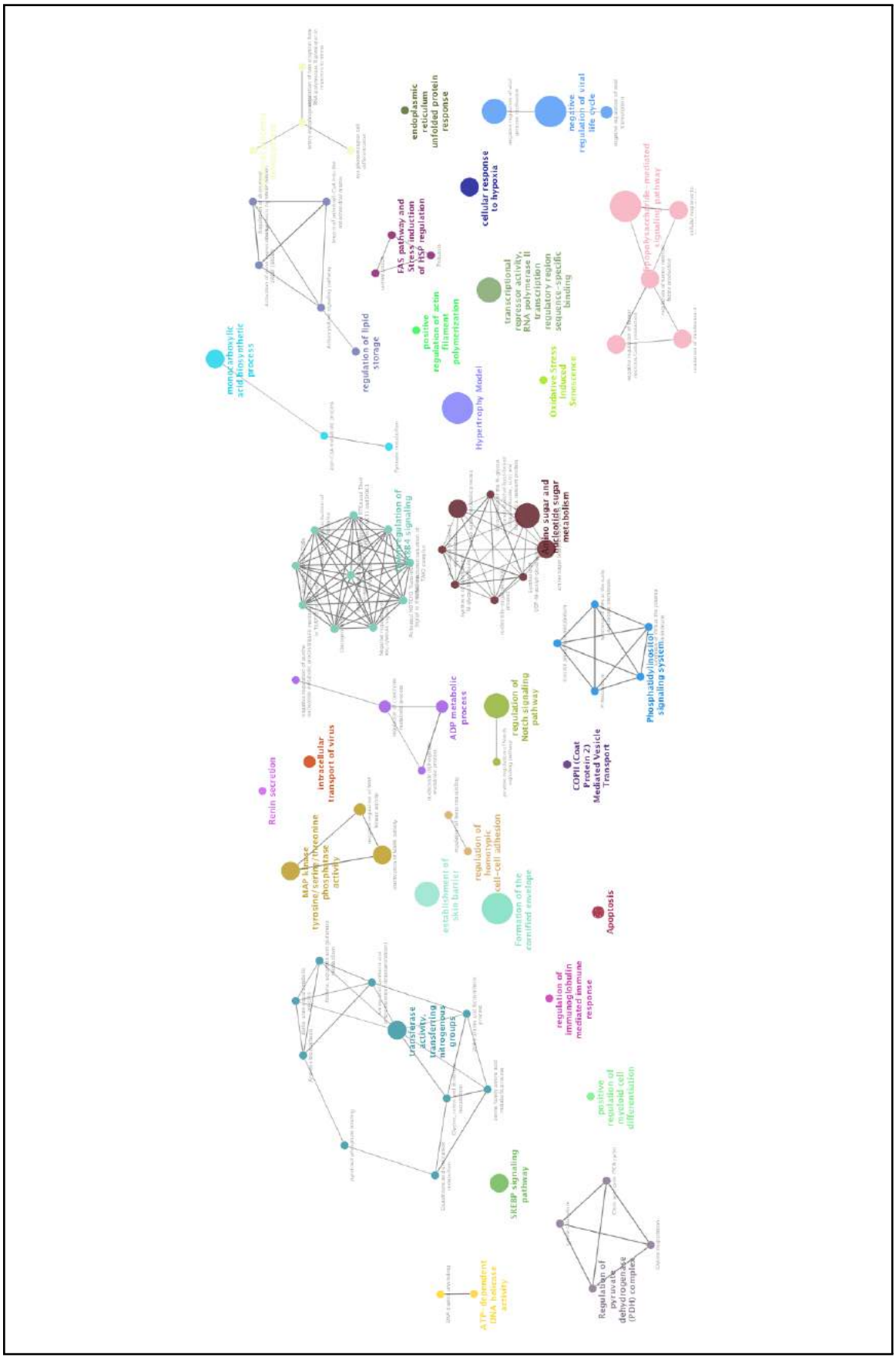


Figure 3.9: Schematic representation of enriched GO terms within the genes in clusters with potential synthetic lethal trajectories. GO term groups that are connected with grey lines and shown in the same color represent grouping based on broader a Molecular Function(MF) analysis.

3.3.3 Characterization of minimal residual disease *in vivo* in mice

Establishing a pure correlate of minimal residual disease using mouse mammary glands and organoid cultures has proved useful in elucidating mechanisms of tumour regression as seen in the previous findings of the lab [52] and the findings so far in this report. However, although the 3D culture system is clean and synchronous it provides a rather reductionist approach, owing to the fact that key cell compartments of the breast tissue are missing from this system: fat cells, myoepithelial cells, fibroblasts and most importantly, a functioning immune system. Without the complexity encountered in living organisms, it was deemed important to validate our *in vitro* findings in an *in vivo* context, using mouse experimentation. Before we started to explore certain synthetic lethal targets and the effect of their inhibitors *in vitro*, we wanted to know that these targets were indeed upregulated in mice with full organ morphology and a proficient immune environment.

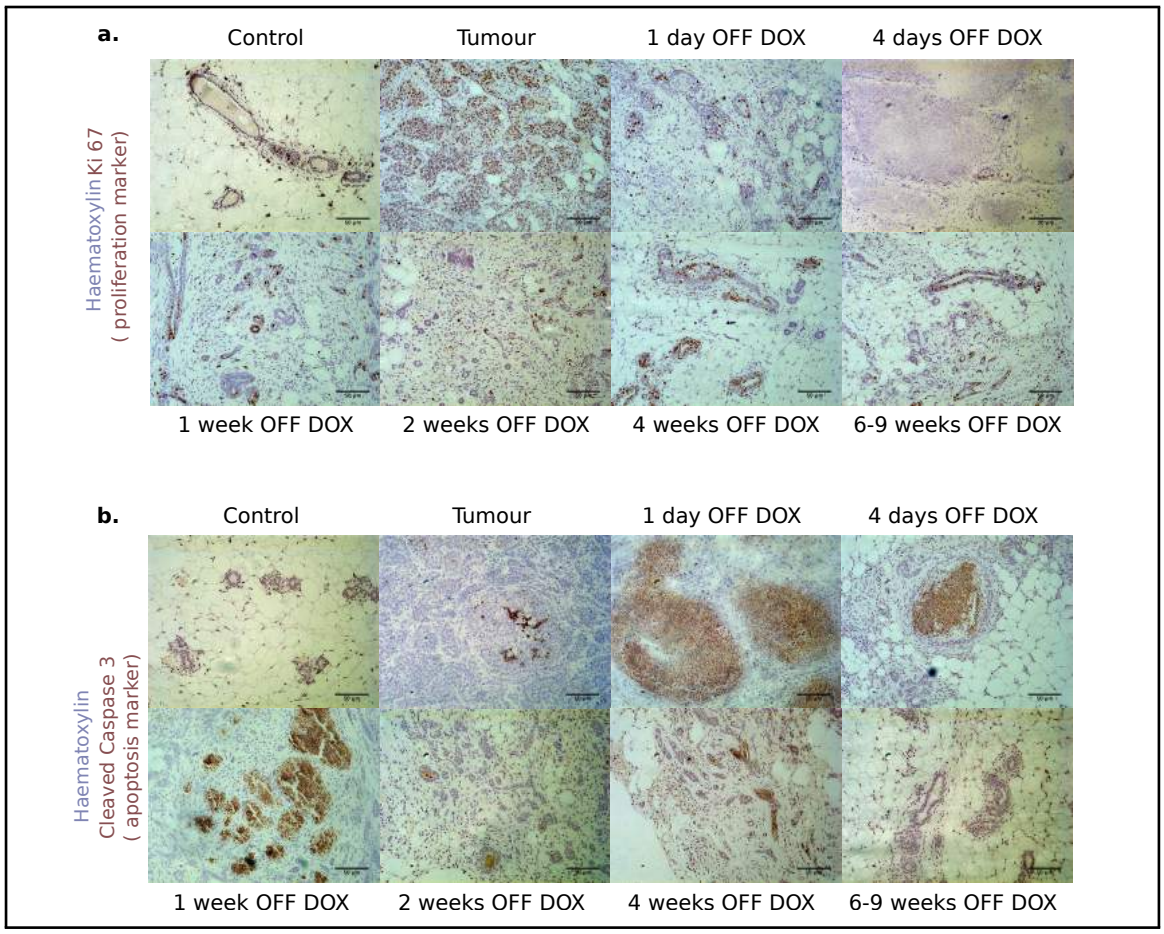


Figure 3.10: Immunohistochemistry (IHC) staining of mouse mammary tissue sections during the regression of mammary tumours upon withdrawal of doxycycline from the diet of TetO-MYC/ TetO-Neu/ MMTV-rtTA mice. (a) IHC stains for *Ki67* proliferative marker and (b) cleaved *Casp3* apoptotic marker in mammary glands at the indicated times during tumour regression. Counter stained with haematoxylin. Scale bar, 50 μm.

The first step was naturally characterizing the inducible tumour model *in vivo*. TetO-

MYC/ TetO-Neu/ MMTV-rtTA adult female mice (between 8-10 weeks old) were administered food with doxycycline so that they developed tumours. Palpable tumours were observed within 5-6 weeks (median time to tumour = 38 days) when oncogenes are induced in their mammary glands by administration of food with doxycycline. Tumours were allowed to grow until the tumour burden was considered too high (any tumour lobule is more than 2 cm in diameter) and then the mice were switched to normal food without doxycycline. Thereafter, rapid tumour regression follows, wherein tumours are no longer palpable as early as 2-3 weeks after removal of doxycycline from the animal diet. However, upon sacrificing and dissecting mice 2-3 weeks OFF DOX we observed micro tumours and regressing nodules in the mammary glands. Consequently, the animals were allowed to regress for longer to allow complete regression. IHC staining for apoptosis and proliferative markers was performed at various timepoints during *in vivo* regression (Figure 3.10) and 9 weeks OFF DOX was ascertained as the "residual disease" state. Animals showed continued apoptosis in their mammary gland tissue, until 6 weeks after removal of doxycycline from the animal diet. Analogous to the *in vitro* regression, IHC stains early after doxycycline withdrawal from the animal diet, showed massive cell death and depletion of oncogene within 48 hours. This contributed to deciding the hallmark timepoints during *in vivo* mammary tumour regression in this mouse model.

At the hallmark time points shown in Figure 3.11, tissue samples were collected from 3 mice (biological replicates) per time point. An age matched control was also harvested with each tumour/regressing tumour mouse. This was frequently a wild type or uninducible genotype sibling. Typical harvest from these mice included:

1. Fixed tissue for paraffin embedding and histology
2. Frozen tissue for DNA, RNA, protein extraction
3. Estrous smears to control for estrous cycle effects on gene expression/protein levels

RNA samples from mouse mammary tissue were run on microarray chips for gene expression profiling. Analysis of the microarray data and differentially regulated genes and pathways compared to control mice was performed using the Transcriptome Analysis Console (TAC) 4.0 software. Principal Component Analysis (PCA) plotting of this microarray dataset (Figure 3.11) follows the trend observed in the *in vitro* dataset with the residual disease correlate clustering distinctly from the control samples despite phenotypic similarity in tissue morphology.

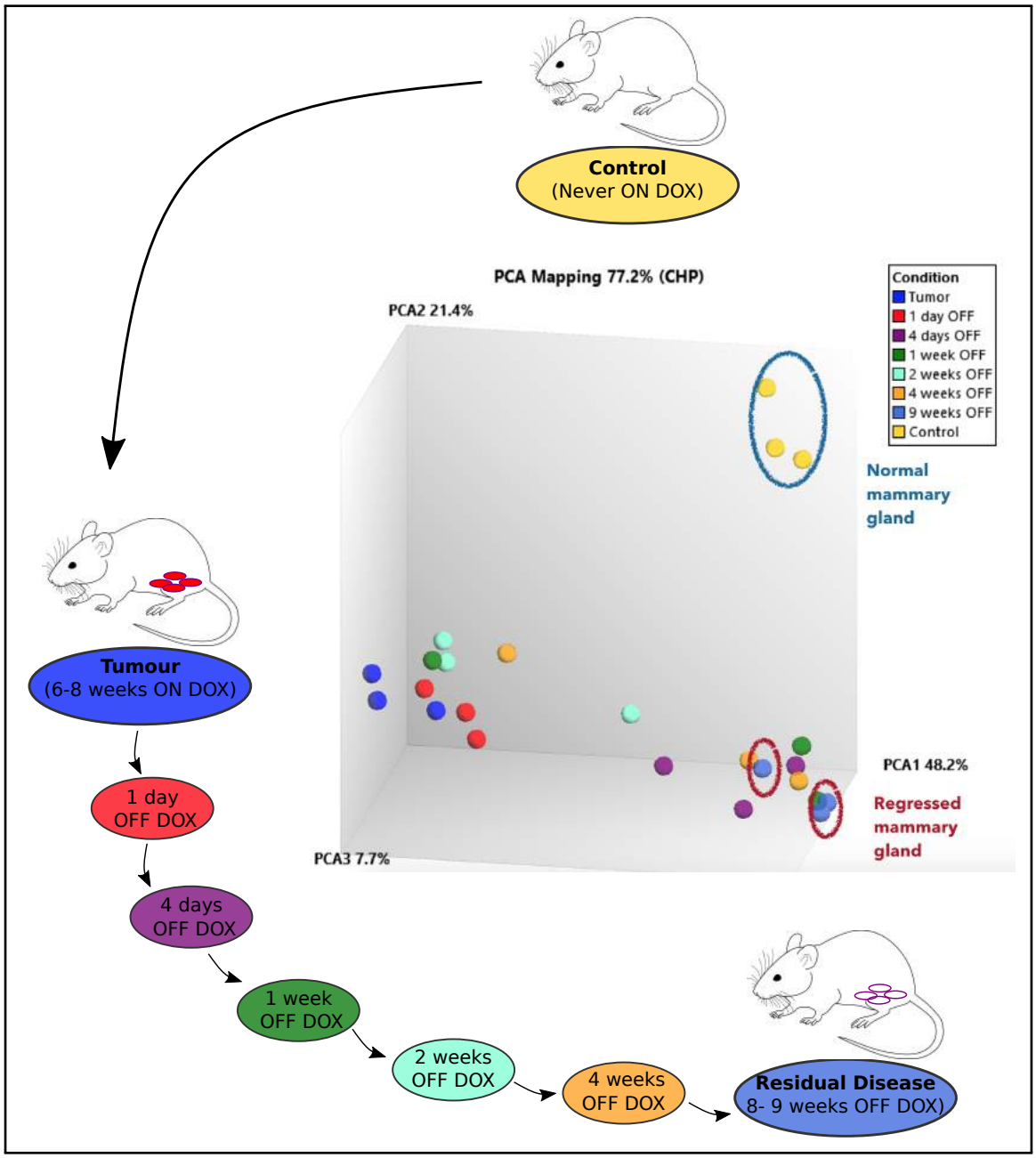


Figure 3.11: PCA analysis plot for micro array data from samples collected at hallmark time points during tumour regression in mice. Each dot on the plot represents a biological replicate and the stage of harvest is color coded with the schematic bubbles.

Now equipped with the gene expression changes in the mouse model with organismal complexity, we performed a manual overlay of gene enrichment analysis for both *in vivo* and *in vitro* datasets. This yielded the following interesting subgroups of pathways that seem to be altered both in reductionist 3D cultures and immune proficient mice during tumour regression:

(a) Focal Adhesion PI3K- AKT-mTOR Sinalling Pathway

(b) MAPK Signalling (specifically: c-Jun Kinase/Stress-activated Pathway)

(c) TNF-alpha and NF-κB signalling pathway

(d) Tgf-Beta Signalling pathway
(e) DNA damage response (ATM dependent)
(f) Apoptosis pathways.

Key molecular players from these pathways were selected for validation, including c-Jun, BCl-6 and CCl5.

3.3.4 Validation of selected targets for synthetic lethality during establishment of MRD

Key molecular players that were identified and hypothesized to be involved in the establishment of residual disease— c-Jun, BCl-6, CCL5 — were taken forward for validation. These targets were identified from network analysis on *in vivo* and *in vitro* transcriptomic datasets and so the logical next level of validation would be the protein level.

Using IHC and immunofluorescence staining, c-Jun protein was shown to be enriched significantly in the MRD correlate of the *in vitro* 3D cultures (Figure 3.12a) as well as the tissue sections of the regressed mammary glands harvested from mice (Figure 3.12b). As indicated in the graphs in Figure 3.12a & b, c-Jun was upregulated on the transcriptomic level. It is an important transcription factor involved in numerous cell activities, such as proliferation, apoptosis, survival, tumourigenesis and tissue morphogenesis[173]. It has already been shown to be relevant in the human breast cancer landscape because it cooperates with NF-B to prevent apoptosis induced by TNF [174]. Protein validation of c-Jun warranted further functional validation of this target. We aimed to functionally validate c-Jun first in the *in vitro* context by testing the effect of its inhibition on the establishment of residual disease rims.

Sylwia Gawrzak in the group collaborated with me to adapt our primary murine mammary cultures to her inhibitor screening pipeline using 3D cultures, high throughput imaging and cell viability and toxicity assays. Inhibitors of potential synthetic lethal targets were to be tested in *in vitro* cultures in the context of combinatorial treatment — inhibitors were to be added during removal of doxycycline ("ideal therapy"). The hypothesis was, that if the cells surviving therapy and reintegrating to form residual disease rims were dependent on the synthetic lethal target, then an inhibitor for the target applied during oncogene withdrawal would reduce the number of cells surviving and limit the reservoir of cells establishing the minimal residual disease correlate. As such, we expected to see reduced cell viability and increased cell death in regressing organoids that were taken off doxycycline ("ideal therapy") along with an inhibitor against the synthetic lethal target as compared to simply organoids regressing because of doxycycline withdrawal.

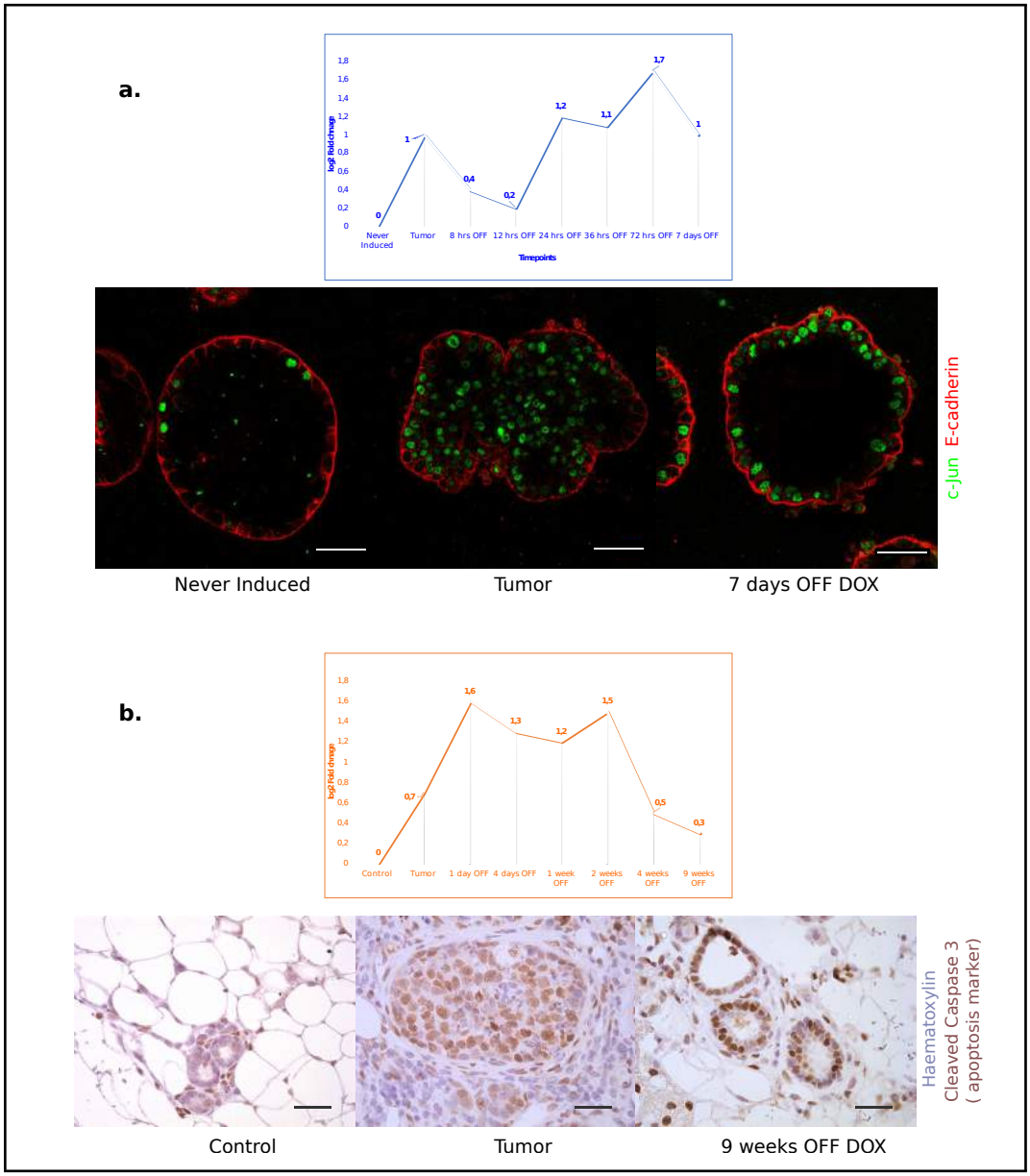


Figure 3.12: (a) Gene expression profile for *c-jun* in 3D cultures undergoing tumour regression are shown in the blue graph. Y-axis indicates the log2 Foldchange values of gene expression at each timepoint as compared to the Never Induced timepoint. The IF staining panels below show protein localization of *c-jun* (green) in 3D acinar cultures at Never Induced(control), 5 day ON DOX (Tumour) and 7 days OFF DOX (residual disease) timepoints. *Cdh1* (E-cadherin) used as membrane marker in red. (b) Gene expression profile for *c-jun* in mammary tissue of mice undergoing tumour regression are shown in the orange graph. Y-axis indicates the log2Foldchange values of gene expression at each timepoint as compared to the Control. The IHC staining panels below show protein localization of *c-jun* in mammary tissue sections of control mice , tumour mice and mice harvested after 9 weeks OFF DOX (residual disease stage). Counterstained with haematoxylin. Scale bar, 50 μm.

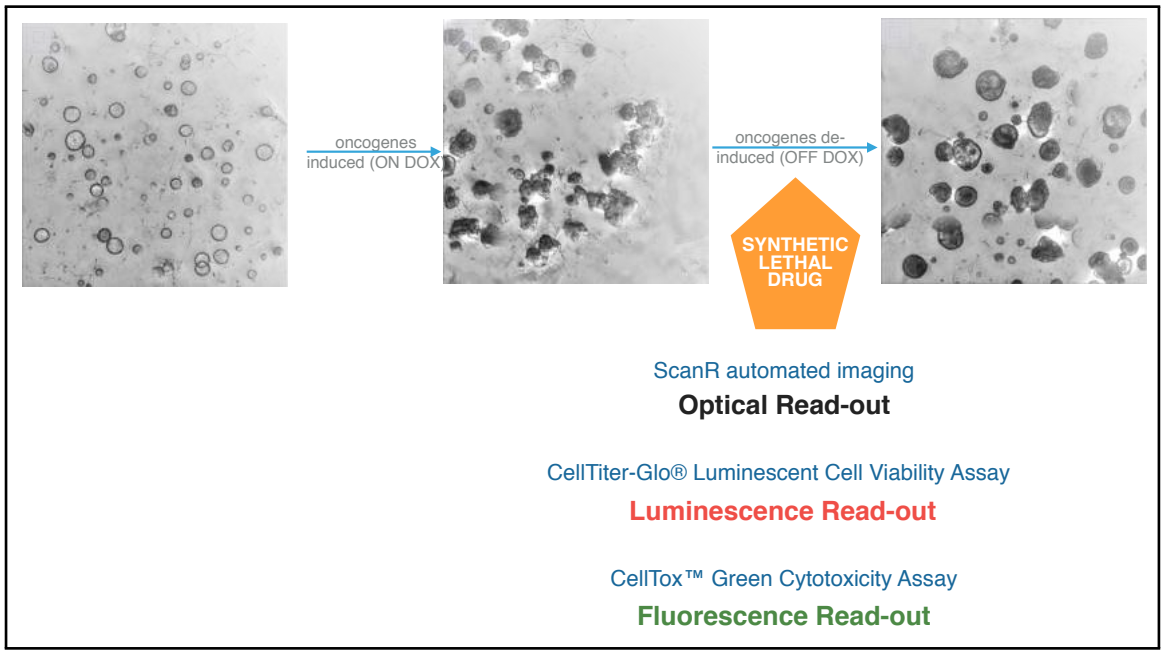


Figure 3.13: Experimental design for the *in vitro* target validation. To test the synthetic lethality of identified genes, their inhibitors would be supplemented in the culture media starting from the point of doxycycline withdrawal. Optical, Luminescence and Fluorescence readouts were incorporated to evaluate phenotype variation, cell viability and cell death.

To test this hypothesis, we adapted our 3D culture method to mini 3D gels (5ul) in a 96 well format (experimental design summarized in Figure 3.13). This format allows for experimental reproducibility and reduced variability – because one can test varying concentrations of the inhibitors on the same plate with the appropriate controls and enough technical replicates to perform statistical testing. The 96- well mini-gel format also allowed for:

(a) Optical readout via high throughput ScanR imaging

This facilitated following whole gels over the time course visually to demarcate phenotypic changes.

(b) Cell Viability read-out using CellTiter-Glo® Luminescent Assay

This outputs a well-wide viability parameter that is ATP dependent and can be used to compare metabolically active cells in each well using a luminescence read-out.

(c) Cytotoxicity and Cell Death readout via CellTox^{TM} Green Fluorescence Assay This outputs a well-wide cell toxicity parameter that is biomarker dependent and can be used to compare apoptosis in each well using a fluorescence read-out.

Table 3.3 shows the list of inhibitors tested using this pipeline on 3D cultures grown from TetO-MYC/ TetO-Neu/ MMTV-rtTA mouse mammry glands. We tested inhibitors for c-Jun namely Tanshinone IIA and JNK inhibitor SP600125; BCL-6 namely FX-1 and 79-6 inhibitor and two targets that form integral parts of the NF-κB cascade, CCL5 with it's antagonist Maraviroc and the IKK2 kinase with the IKK2 -IV and IKK2-VIII inhibitors.

Inhibitor tested	Inhibitor target	Concentration range(μM) tested
Tanshinone IIA (AP-1 inhibitor)	*c-Jun*	0.1 - 80
SP600125(JNK inhibitor)	*c-Jun*	0.02 - 80
FX-1 Inhibitor	*Bcl6*	15.6 - 4000
79-6 Inhibitor	*Bcl6*	1.95 - 1000
IKK2 Inhibitor IV	*Ikbkb*	3.125 - 50
IKK2 Inhibitor VIII	*Ikbkb*	0.39 - 100
Maraviroc	CCL5 antagonist	6.25 - 100

Table 3.3: Inhibitor list tested for synthetic lethality in regressing mammary matrigel cultures using the pipeline outlined in 3.13.

Bcl6 is a strong transcriptional repressor implicated in B-cell lymphomas and other solid cancers. It has recently been implicated to have a role in human breast cancers using cell line models and bioinformatic analysis [175]. It has also been shown that higher BCl-6 expression reduces relapse free survival in HER2+ patients (Online KM plotter[176]). In our model system, commonly repressed proteins like the ones controlling DNA damage sensing (eg: *Atr*, *Chek1*, *Trp53*, *Cdkn2a*) and proliferation checkpoints (eg: *Cdkn1a*, *Cdkn1b*, *Cdkn2b*, *Pten*) are found to be downregulated during tumour regression. These are all known to be repressed by *Bcl6*. In addition, *Bcl6* upregulation could be validated on a protein level in the 3D cultures where it was found to be enriched in regressed rims as compared to never induced acini (Figure 3.14). Combined with the elevated transcriptomic signatures from the *in vitro* and *in vivo* experiments, it qualified for functional validation with the inhibitor screening pipeline.

Similarly, targets from the NF-κB cascade were also tested for synthetic lethality using this pipeline. The NF-κB complex proteins such as *Rel* and *Relb* were shown to be upregulated *in vitro* and *in vivo* during regression. Previous literature suggests that TNFα-induced NF-κB activation induces transcription and expression of genes encoding anti-apoptotic factors *Birc2*, *Birc3* and *Bcl2* homologue *Bcl2l1* [174]. Transcriptome levels of all these anti-apoptotic factors along with the *Tnf* RNA are also upregulated in the minimal residual disease correlates in our system (See schematic in Figure 3.15).

As hypothesized earlier, inhibiting a true synthetic lethal target is expected to have a pronounced deleterious effect on regressing tumours. Controlling for inhibitor toxicity was performed using a dose curve method of experimental design and by applying the

inhibitor on Never Induced acini in gels grown on the same plate. As such, upon addition of a synthetic lethal target inhibitor, we would expect the regressing tumour acini to show decreased viability and at the same time we would expect no effect on the Never induced acini. A target cannot be synthetic lethal if it's inhibitor, at a given concentration, is less toxic towards regressing organoids in comparison with normal mammary acini.

Figures 3.16 and 3.17 show the readouts from the CellTiter-Glo® Luminescent Assay (viability) and CellTox^{TM} Green Fluorescence Assay (cell death) for two c-Jun inhibitors – Tanshinone IIA and JNK inhibitor SP600125. Both experimental readouts for both inhibitors indicate that although the inhibitor severely reduces residual disease viability at higher doses, this correlate is far more refractory than normal never induced mammary organoids. Similar results were obtained during the testing of all the inhibitors mentioned in Table 3.3 (See Supplementary Figures 5.18, 5.19, 5.20 and 5.21).

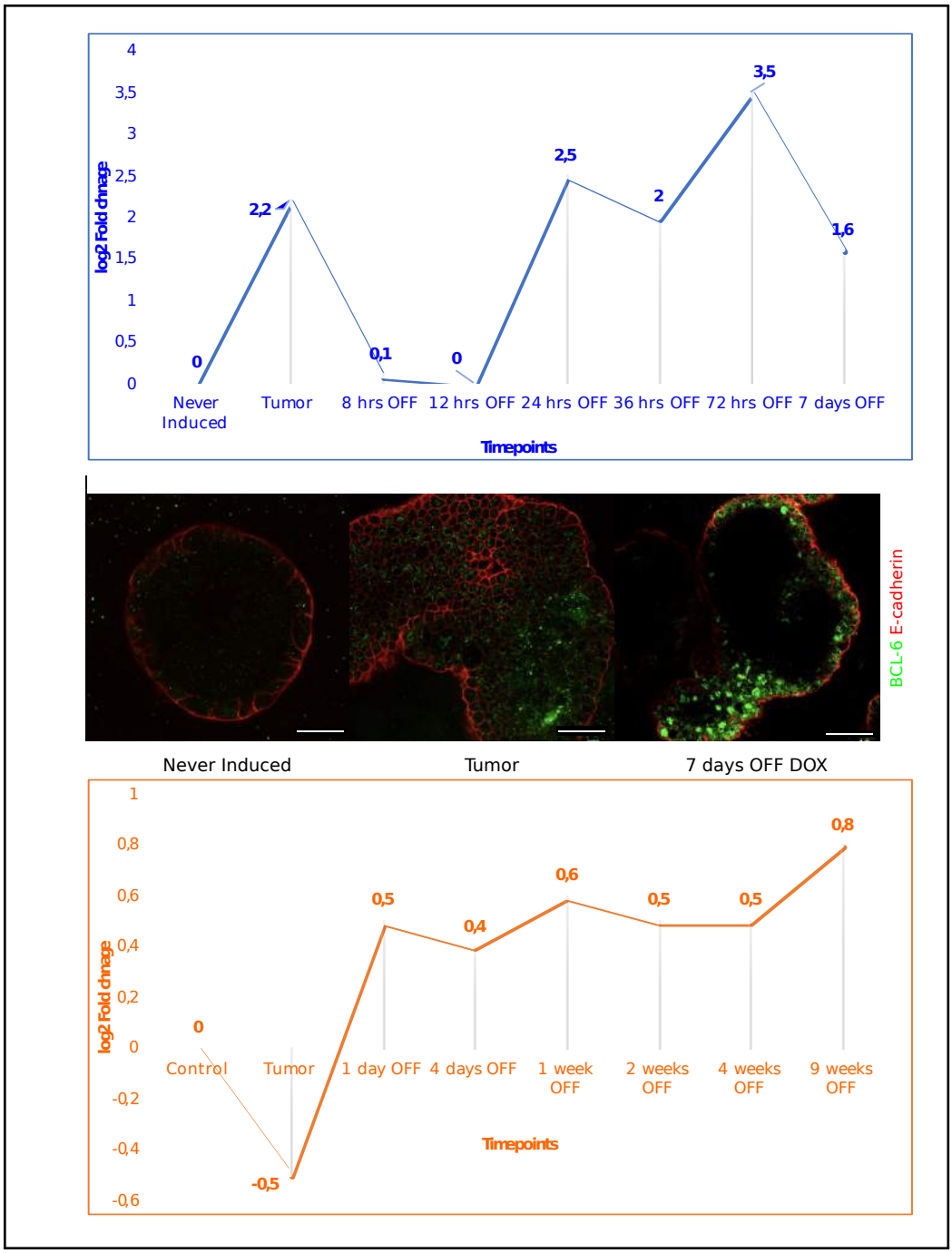


Figure 3.14: Gene expression profile for *Bcl6* in 3D cultures undergoing tumour regression are shown in the blue graph. Gene expression profile for *Bcl6* in mammary tissue of mice undergoing tumour regression are shown in the orange graph. Y-axis indicates the log2 Foldchange values of gene expression at each timepoint as compared to the Never Induced timepoint or Control mice. IF staining panels show protein localization of *Bcl6*(green) in 3D acinar cultures at Never Induced (Control), 5 day ON DOX (Tumour) and 7 days OFF DOX (residual disease) timepoints. *Cdh1* (E-cadherin) used as membrane marker in red. Scale bar, 50 μm.

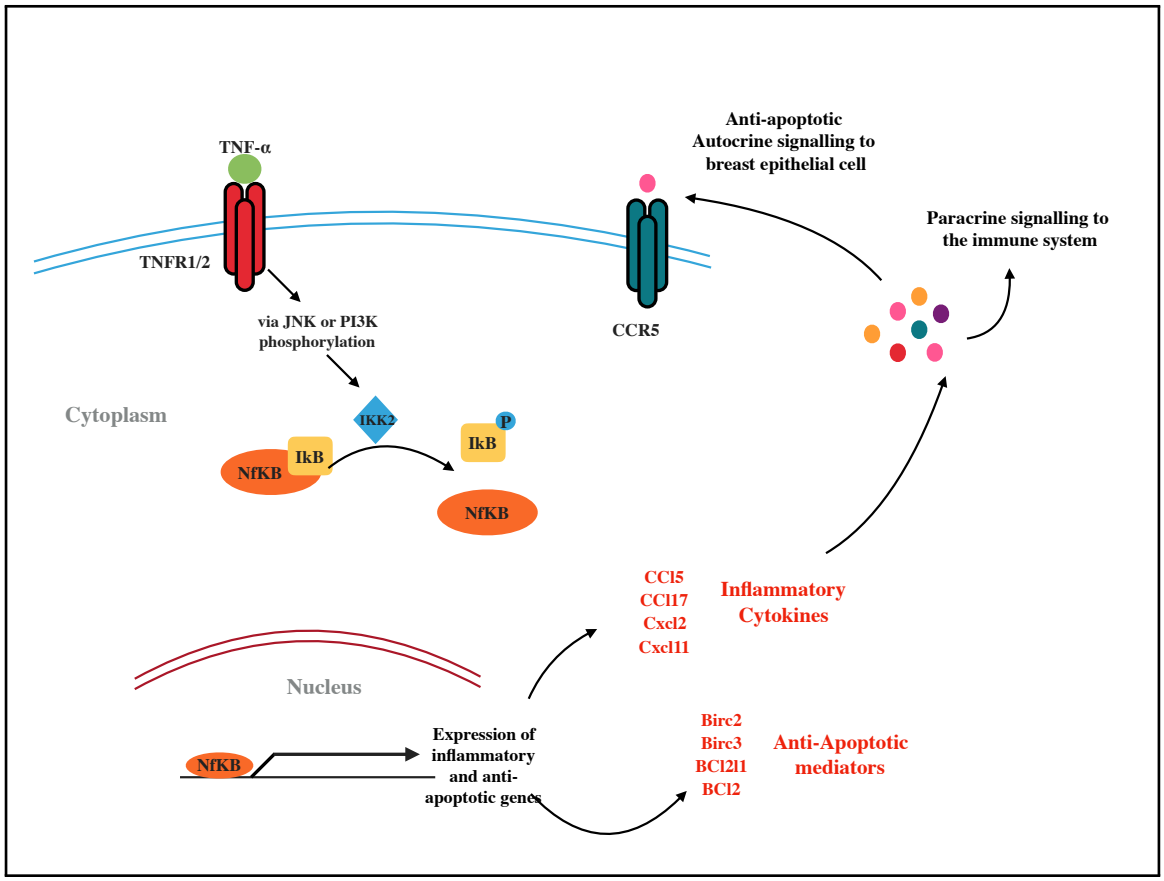


Figure 3.15: Schematic representation of the molecular players in the NF-κB cascade and their feedback signalling loop with CCR5 receptor anti-apoptic signalling during *in vitro* breast cancer regression. Molecular targets highlighted in red were up-regulated along the regression transcriptome of our 3D culture system.

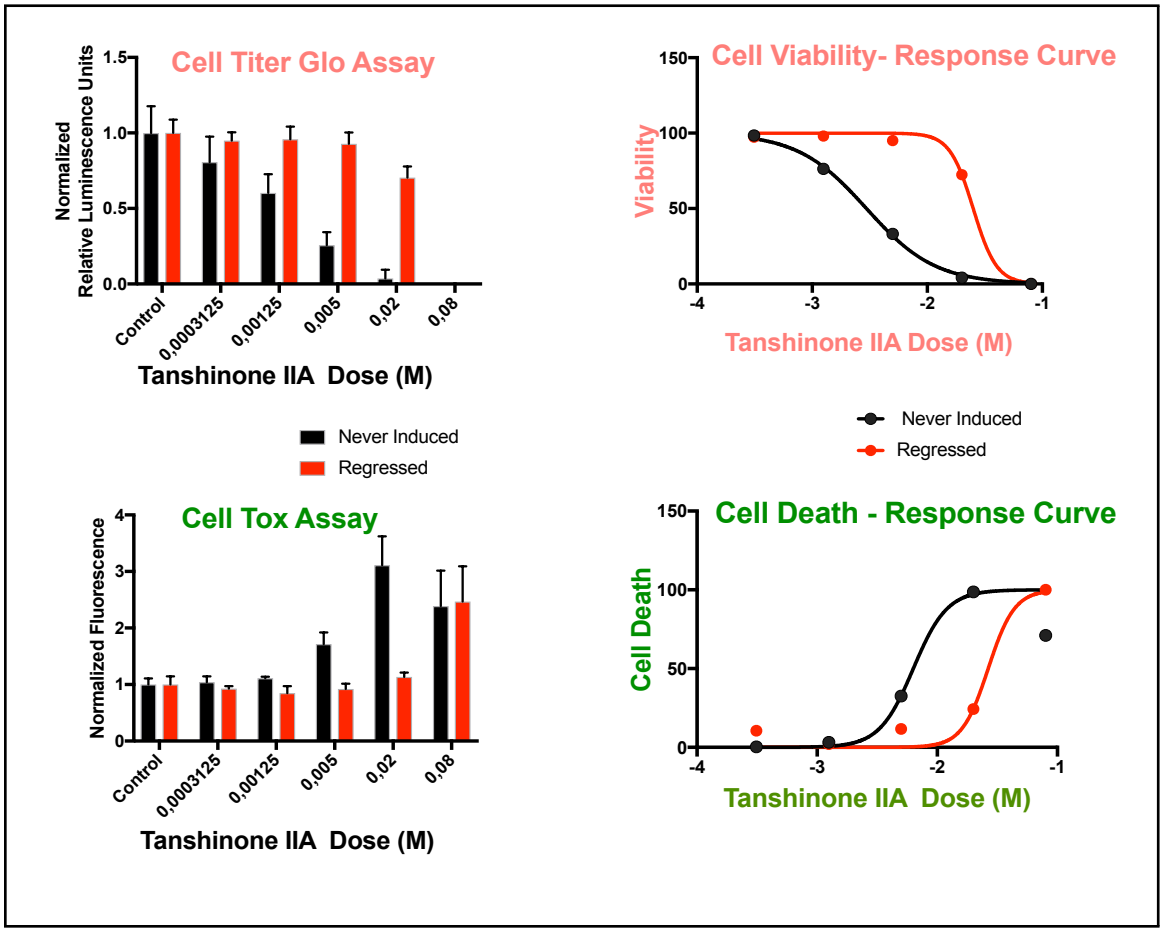


Figure 3.16: Cell Viability and Cell Death readouts from synthetic lethality testing of inhibitor Tanshinone II A against the identified traget *c-Jun*. Upper left graph shows the readout from the CellTiter-Glo® Luminescent Assay. Cell viability readouts for the 5 doses of Tanshinone IIA tested and normalized against the control are shown. Derived from this graph is the upper right graph depicting the Cell Viability response curve. IC_{50} values for Tanshinone IIA in the Cell Viability response curve are 0.0029 mM for Never Induced and 0.025 mM for Regressed cultures. Lower left graph shows the readout from the CellTox^{TM} Green Fluorescence Assay. Cell death readouts for the 5 doses of Tanshinone IIA tested and normalized against the control are shown. Derived from this graph is the lower right graph depicting the Cell Death response curve. IC_{50} values for Tanshinone IIA in the Cell Death response curve are 0.0063 mM for Never Induced and 0.0265 mM for Regressed cultures.

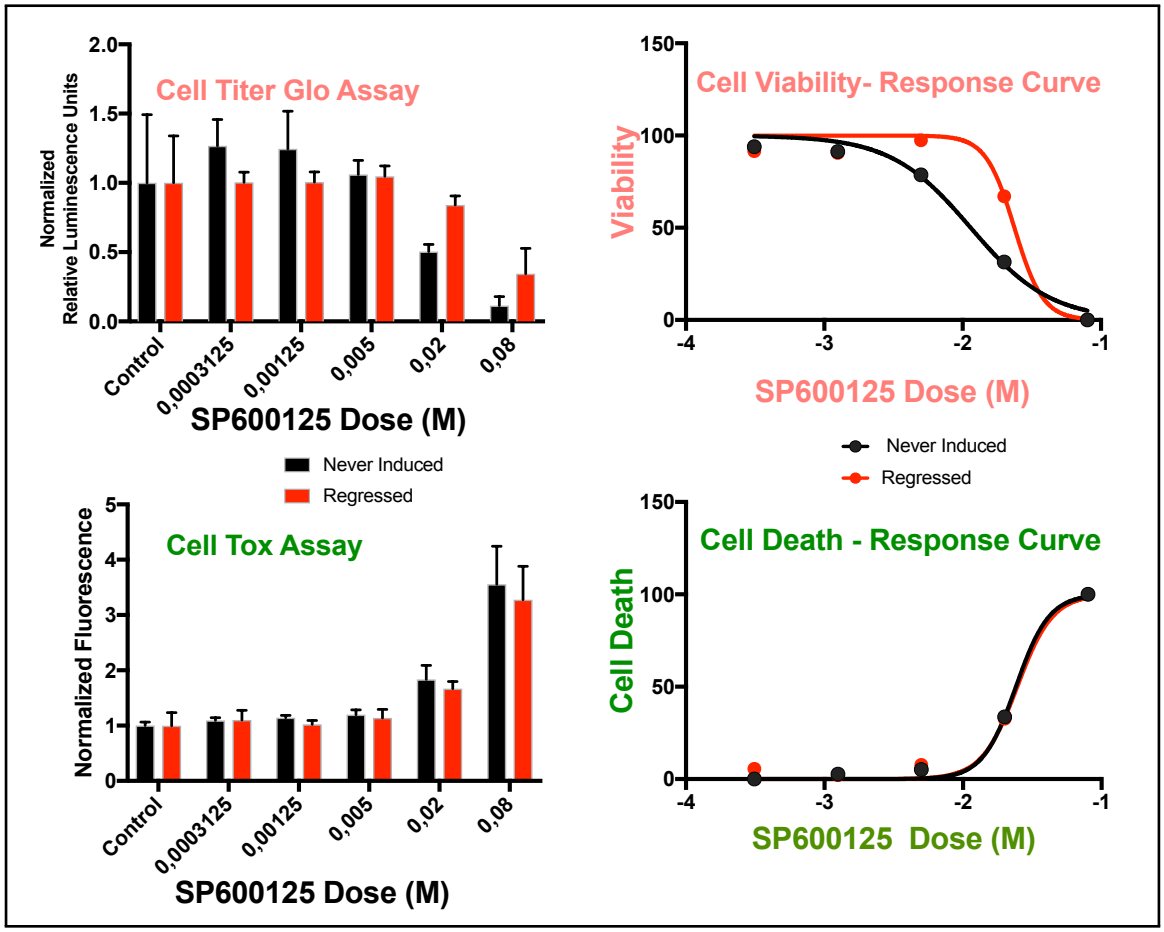


Figure 3.17: Cell Viability and Cell Death readouts from synthetic lethality testing of inhibitor JNK Inhibitor SP600125 against the identified traget *c-Jun*. Upper left graph shows the readout from the CellTiter-Glo® Luminescent Assay. Cell viability readouts for the 5 doses of JNK Inhibitor SP600125 tested and normalized against the control are shown. Derived from this graph is the upper right graph depicting the Cell Viability response curve. IC_{50} values for JNK Inhibitor SP600125 in the Cell Viability response curve are 0.011 mM for Never Induced and 0.023 mM for Regressed cultures. Lower left graph shows the readout from the CellTox^{TM} Green Fluorescence Assay. Cell death readouts for the 5 doses of JNK Inhibitor SP600125 tested and normalized against the control are shown. Derived from this graph is the lower right graph depicting the Cell Death response curve. IC_{50} values for JNK Inhibitor SP600125 in the Cell Death response curve are 0.024 mM for Never Induced and 0.0246 mM for Regressed cultures.

3.3.5 Imaging tumour organoid regression using light sheet microscopy

The primary focus of this study was to identify key molecular players in minimal residual disease of breast cancer and the transcriptomic method described so far is an unbiased approach to elucidate the mechanisms of establishment of minimal residual disease using gene expression analysis (RNA Seq). In a more biased approach, we aimed to image the special dynamics during regression to analyze mechanisms of survival of residual disease cells in our 3D culture system.

To better understand the spatio-temporal dynamics of regressing tumours and visualize the establishment of a minimal residual disease correlate, we exploited the 3D culture method combined with the inducible TetO-MYC/ TetO-Neu/ MMTV-rtTA mouse model. The reporter protein H2B mCherry was bred into the mouse line to fluorescently mark the nuclei for imaging –to image regression *in vitro*. SPIM microscopy was ascertained to be the most adept technique for this objective because it facilitates rapid live cell imaging, with low photo-toxicity, greater penetration depths and low scattering owing to illumination with a light sheet instead of a point source.

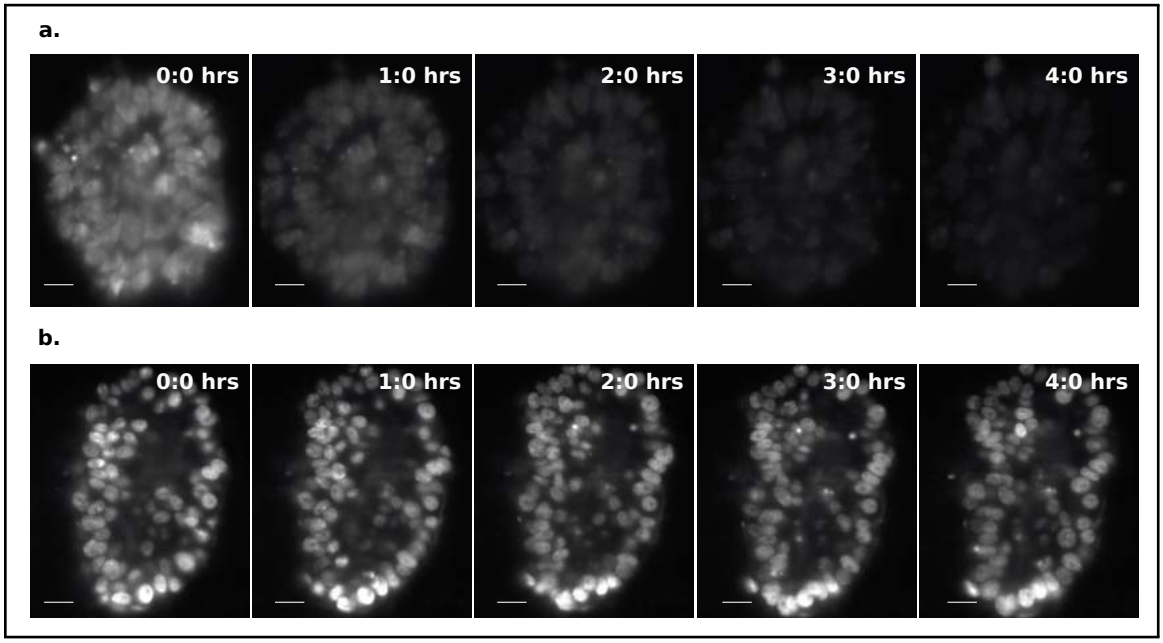


Figure 3.18: tumour organoids recorded at the In-Vi SPIM (Luxendo) for short periods to test phototoxicity. (a) shows image acquisition with a z-step size of 0.5 μm, while (b) shows image acquisition with a z-step size of 1 μm. Time stamps on the upper left corner of panel images show times since the start of image acquisition at 10-minute intervals. H2B-mCherry fluorescence in cells shown in grey. Scalebar, 10 μm.

However, one of disadvantages of this alternate optical arrangement- illumination with a light sheet perpendicular to detection – is that SPIM microscopes often require special sample mounting for imaging. Mounting procedures are sometimes incompatible with optimal growth conditions for the 3D mammary epithelial cell cultures and must be estab-

lished differently. For example, during my efforts to image mouse mammary acini with the Mu-Vi SPIM (Multi-view SPIM), growth seemed to be impaired due to lack of diffusion of nutrients with the media because the 3D cultures had to be seeded in 1mm diameter FEP tubes. Similarly, I also experienced sample mounting incompatibility with the Leica SP8 Confocal-SPIM for our 3D system. To suit the mounting needs of the SP8 SPIM, we seeded matrigel cultures on 2 mm wide FEP strips and mounted them at an elevation to allow the light sheet formation through the sample. Although we could image a few movies on this system, the gel cultures were increasingly unstable over time and we experienced frequent sample drift. Another disadvantage of this system was the inability to image at multiple positions in the gel culture efficiently. However, the InVi SPIM made by Luxendo was eventually shown to be ideal for our sample as detailed in the previous project in this thesis.

Imaging on the InVi SPIM set up was optimized to reduce phototoxicity and so that the imaging conditions would not affect cellular fates. In Figure 3.18, the upper panel shows a movie recorded at the InVi SPIM over 4 hours, every 10 mins, with 0.5 μm z-step size. The panels, as they progress from left to right, show bleaching of the fluorophore and eventual phototoxicity in the sample. The lower panel of Figure 3.18 shows a similar recording of organoid growth wherein the parameters were optimized to prevent photobleaching: laser powers of 13.1 microWatts for the 488 nm laser and 36.4 microWatts for the 594 nm laser, 100 millisecond exposure time per frame, 1m z-spacing between frames and 10-minute imaging intervals. A series of optimization experiments, involving different laser powers, exposure times and z-step sizes yielded these parameters to be optimal for long term imaging (96-120 hours) without photo-bleaching or photo-toxic effects on growth.

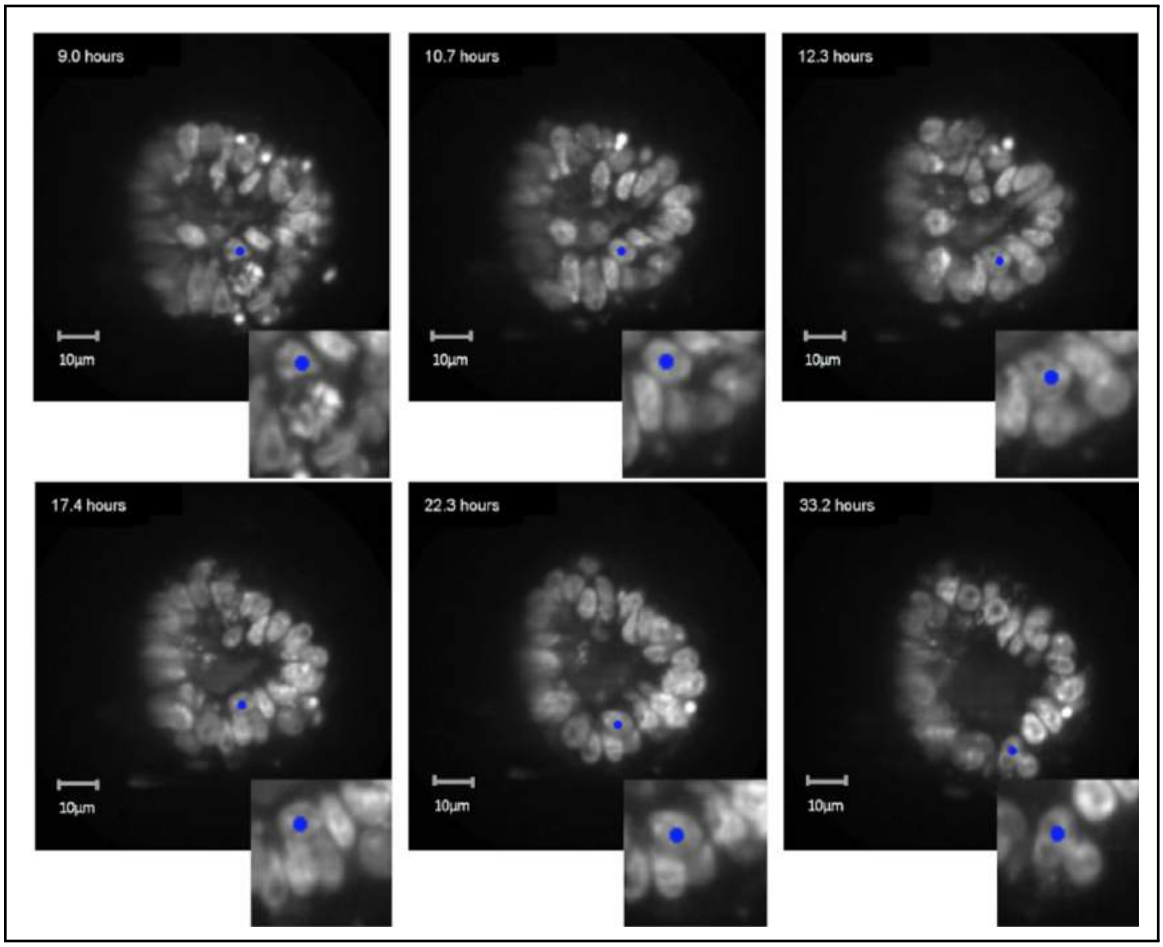


Figure 3.19: Tumour regression recorded at the In-Vi SPIM (Luxendo) with manual tracking . Blue tracking dot marks a cell that exhibited movement to the surviving rim during apoptosis of the core. Time stamps on the upper left corner of panel images show times since doxycycline withdrawal from the 3D culture media. H2B-mCherry fluorescence in cells shown in grey. Scalebar, 10 μm.

Once the sample mounting and imaging conditions were optimized, we proceeded with the objective of recording tumour organoid regression. Video stills in the Figure 3.19 show that imaging this process from the point of de-induction of oncogenes to almost 72 hours OFF DOX, is indeed possible. The inlay images in Figure 3.19 show the same single cell marked with a blue dot in the entire time frame. The cell looks to have journeyed from the center of the tumour organoid at 9 hours OFF DOX to integrating into the regressed rim at 33 hours OFF DOX. Clearly these movies of tumour regression already show some interesting phenomenon exhibited by cells during regression — cell movement to the basement membrane and repolarization into the epithelial monolayer as a method of survival.

But these anecdotal observations could not be reproduced simply because the datasets are too convoluted for the current image analysis methods described in detail in the previous project of this thesis (Figure 2.10). Inability to segment these structures into single cells in 3D impedes tracking their movement, divisions and apoptosis. Figure 3.20a shows the raw images (left panels) and resultant pixel probability maps (right panels) for two images of varying complexity. The lower panel images are from a tumour acinus and the upper panel images are from a never induced acinus – both recorded at the InVi SPIM during 96-hour image acquisitions. Clearly the algorithm cannot correctly classify pixels in the chaotic tumour acinus 2D frame efficiently.

To circumvent this technological deficit, we propose a lentiviral delivery system to stochastically mark single cells within an organoid before induction and de-induction of oncogenes (Figure 3.20b). Introducing, via lentiviral delivery, a reporter fluorophore (H2B-GFP) into some of the acinar cells before induction with doxycycline will allow us to selectively image them in a separate channel without the chaotic microenvironment. This imaging and image analysis modality are similar to the one successfully used for the stochastic model mentioned in the previous project. Two channel imaging with randomly marked tumour cells should therefore be able to simplify segmentation, allowing for tracking of some cells, if not all, during tumour regression. This imaging and data analysis are planned for my bridging postdoctoral period after PhD defence.

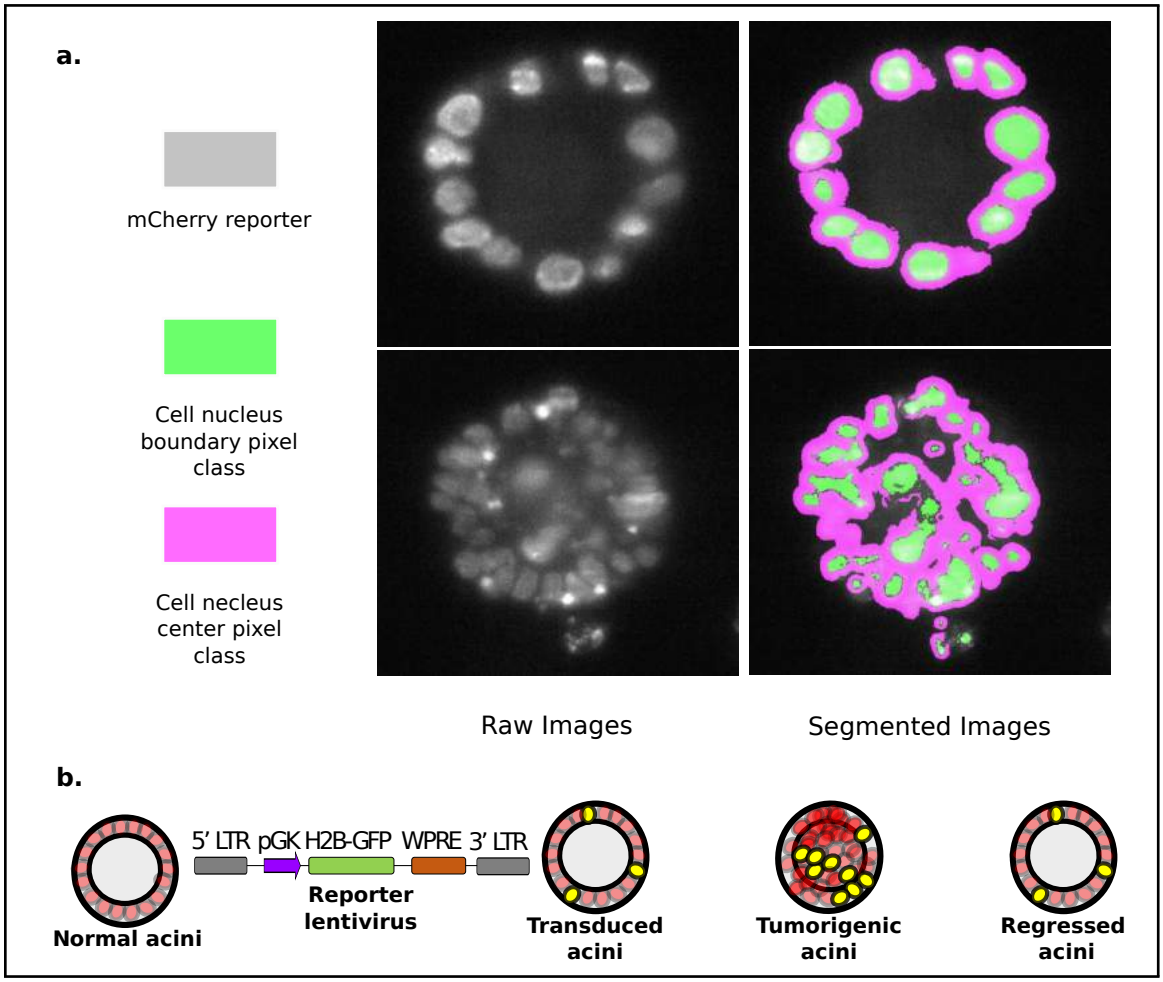


Figure 3.20: (a)Example segmentation using the CATS tool (Fiji) that uses a machine learning algorithm to classify image pixels via feature identification. Left panels show raw image planes(grey- H2B mCherry) and right panels show segmentation results (pink- nucleus boundary pixel class; green- nucleus center pixel class). (b) Schematic showing lentiviral marking strategy to reduce complexity of images recorded and simplify segmentation.

3.4 Discussion

The rationale for targeting breast cancer residual disease is clear from the clinical situation today. Upon treatment with targeted therapeutics, designed to specifically inhibit oncogenes, 30-40% of patients who receive satisfactory remission with targeted therapy, present with relapse tumors. This tumor recurrence is frequently attributed to residual disease: a reservoir of tumour cells that evade primary therapy and persist in a dormant state for long periods. This reservoir of cells represents a black box in terms of availability of patient material. This has resulted in our limited understanding of the mechanisms of residual disease survival and the molecular players that facilitate therapy evasion.

To address the need of molecular targets to combat this disease reservoir, we modelled an *in vitro* substrate of this disease using cells from an inducible mouse model. As described in the results section, this reconfirmed the observation [52] that the residual population of cells resulting from ideal targeted therapy (oncogene de-induction) were indeed not transcriptomically similar to their oncogene nascent (never induced) counterparts. RNA Seq analysis at the identified hallmark timepoints along the time course of establishment of a residual disease substrate further informed us of the demarcation between oncogene regulated pathways (Expected Up and Expected down clusters in Figure 3.7) and novel pathways triggered by oncogenic shock (Table 3.2). These pathways could be targeted to induce synthetic lethality in residual disease cells and represent a gene signature evolution that could be interfered with at multiple points during and after tumour regression (Objective 1).

Further, we performed in-depth analysis of the immune proficient mouse model during tumour regression, establishing time lines for tumor growth and regression in the animals based on IHC staining. We used these time lines to select time points for microarray analysis. Analysis of the transcriptomic data during mouse tumor regression, corroborated the finding that novel pathways are activated during tumour regression, independent of the oncogene and distinct from normal cells(Objective 2).

The subset of signalling pathways that could be validated both in the mouse model and its reductionist *in vitro* counterpart were evaluated for potentially synthetic lethal targets. Careful analysis of the molecular players in the pathways identified warranted the selection of Bcl6, c-Jun and the molecular players in the NF-κB cascade, as potential synthetic lethal targets(Objective 3). Below is a bullet list of biological findings that support the selection of each of these targets.

BCL6

- BCL6 rests at the centre of the DNA Damage Response pathway shown to be dysregulated during residual disease establishment in our model system. BCL6 has been shown to reduce the cells sensitivity to DNA damage [177] due to its direct repression of the ATR gene, which prevents DNA-damage responding ATR targets from being expressed. BCL6 also directly represses the CHEK1 gene, which is a critical mediator of the ATR-dependent DNA damage-signaling pathway[178]. BCL6 can also directly repress the TP53 and CDKN1A genes, thus severely impairing the function of cellular DNA damage checkpoints at multiple levels [179][180]. All these genes have also been downregulated significantly during tumour regression in the *in vitro* 3D culture model.

- BCL6 is a transcriptional repressor protein that confers B-cells the ability to tolerate rapid proliferation and simultaneous genetic recombination[181]. This ability conferred in residual cells could allow therapy evasion.

- In germinal centers (GC) (transient and dynamic cellular compartments that form within secondary lymphoid organs), BCL6 plays a big role in B-cell maturation to generate antibodies against specific antigens [182]. It does so by allowing rearrangement and mutation of the immunoglobulin loci [183]. This points to a phenotype of physiological genomic instability induced by BCL6 that could lead to cancer. Naturally, genetic lesions that cause constitutive expression of BCL6 in GCs are commonly associated with diffuse large B-cell lymphomas (DLBCL) [184] [185].

- BCL6 is expressed in most breast cancer cells lines and that its genetic locus is amplified in approximately 50% of breast tumour samples. Studies targeting BCL6 have proven its importance for breast cancer [175].

- Higher BCl-6 expression reduces relapse free survival in HER2+ patients (OnlineKM plotter [176]).

c-JUN

- c-Jun in combination with c-Fos, forms the AP-1 early response transcription factor. It has been implicated in many cancers as a potent oncogene. Naturally, seeing this gene upregulated in the residual cells (validated even on a protein level both *in vitro* and *in vivo*) motivated us to test its synthetic lethality.

- c-Jun has been shown to mediate an anti-apoptotic function in response to TNFα. TNFα is upregulated in our models during tumour regression and we hypothesized that c-Jun overexpression in some cells was protecting them from apoptosis [186].

- In breast cancer modelling with HER2 over-expression (analogous to the mouse model we employed with out rat Neu oncogene), the knocking out of of c-Jun, reduced cellular migration, invasion, and mammosphere formation of ErbB2-induced mammary tumours[187]. In this study, researchers have shown that secreted inflammatory cytokines like CCL5 – induced by ErbB2 expression — were dependent upon endogenous c-JUN expression and that CCL5 rescued the c-Jun-deficient breast tumour cellular invasion phenotype. Since CCL5 was also upregulated in our model of tumour regression, both *in vitro* and *in vivo*, we postulated an intertwined signalling between these two targets.

- A study analyzing gene expression data derived from serial tumour samples of patients with breast cancer who received Neo-Adjuvant Chemotherapy (NAC) in the I-SPY 1 TRIAL showed that the key transcriptional regulator AP-1, JUN and FOSB were among the 32 upregulated genes in residual tumours after chemotherapy. This points to a human relevance of c-JUN in residual disease [188].

Molecular players in the Nf-κB cascade

As mentioned above, the TNFα upregulation during the process of tumour regression observed in our model, propelled interest in the cascades activated by this cytokine. TNFα induced NF-κB activity seemed like a viable downstream progression – especially since all the NF-κB subunits were also upregulated in our transcriptomic data. We proposed that this TNFα induced NF-κB activity was causing the upregulation of anti-apoptotic mediators like Bir2, Birc3 and BCl2l2 (all of which are enriched in the residual cell transcriptome). These observations warranted targeting of the NF-κB activation cascade. We tried to do this using IKK2 inhibitors (Table 3.3). IKK2 is a multiprotein kinase complex that is responsible for the TNFα induced phosphorylation of IκB(Inhibitor of κB) which then dislocates and allows the NF-κB transcription factor to translocate to the nucleus and induce expression of certain genes. Another interesting target in this cascade was the inflammatory cytokine CCL5.

CCL5

- CCL5 has been shown to play an important role in many facets of tumour progression, such as invasion, metastasis, neoangiogenesis, and immune cell infiltration [189].

- In TNBC (triple negative breast cancer) patients , CCL5 expression has been shown to be enriched in larger tumours remaining after neoadjuvant chemotherapy indicating it's role in residual disease survival [190].

- In prostate cancer, the role of CCL5 in proliferation and metastasis was shown to be mediated through F-actin polymerization and Maraviroc, a CCR5 antagonist was successfully used to block it's effects [191] [192]
- In gene expression data-sets from breast cancer patients treated with neoadjuvant targeted or chemotherapy [193] [194]), CCL5 expression is elevated in residual tumour cells that survive therapy [195].
- Another group at Duke University recently published data with the same inducible mouse model used in this project. They also observe that oncogene withdrawal led to hyperactivation of the NF-κB pathway and downstream upregulation of CCL5. Confirming our hypothesis, they prove that this program was mediated by autocrine TNFα and dependent upon IKK/NF-κB signaling [195](also shown recently in lung cancer [196]). They implicated CCL5 in macrophage recruitment, collagen deposition and tumour recurrence.

The targets detailed above were tested using the screening pipeline (see Figure 3.13, Table 3.3) and residual disease cells that were treated with target inhibitors were assayed for viability and cell death. The experimental readouts for all inhibitors indicate that although the inhibitor severely reduces residual disease viability at higher doses, this correlate is far more refractory than normal never induced mammary organoids. These results indicate that the *in vitro* testing assay might not be the optimal readout for assessing synthetic lethality. 90% of the tumour mass apoptoses upon removal of doxycycline from induced organoids and the cell viability/death assay read outs might not be sensitive enough to record the additional 1% death caused by synthetic lethality of the proposed targets. Therefore, alternate validation methods will be required to validate the "synthetic lethality" of these targets. An *in vivo* validation in mice with a "Relapse free survival" read out might be an ideal method of target validation, given that prevention of relapse in inhibitor treated mice might be considered cure because of the eradication of residual cells with repopulating potential.

In an effort to identify targets in a more biased approach – observing regressing cells and identifying their vulnerabilities in a spatio-termporal dimension – we set up live cell imaging using light sheet microscopy for regressing tumour organoids (Objective 4).

As described earlier, the inner core of a tumour organoid *in vitro* undergoes rapid apoptosis upon removal of doxycycline from the media and we see repolarized epithelial rims within 72 hours of oncogene inactivation. So far, we do not know, how cells rearrange during this process to form a repolarized, surviving rim after 72 hours. Is there any trafficking of survivor cells from the core to the edge to form part of the re-polarized rim? Or do the inner cells apoptose simply because they do not have connection to the basement

membrane to latch onto, activate signaling pathways and survive? Clarification of these spatio-temporal dynamic mechanisms is hypothesized to point to possible interference approaches. For example, blocking of integrin signaling could be synthetic lethal in case we observe a major involvement of basement membrane signaling for survival in the residual cells.

Big-data analysis work-flows allow for visualization of tumour regression in 3D but current machine learning algorithms cannot segment these regressing tumour data-sets to allow for single cell tracking. To circumvent this hurdle, lentiviral reporter delivery is currently being adopted to allow for simplified segmentation and tracking (Figure 3.20).

3.5 Conclusions

Tumor recurrence is a major cause of death in women with breast cancer. Tumor recurrence or relapse is attributed to a subset of cells that somehow circumvent targeted therapeutics and lay dormant before resulting in refractory relapse tumours. These residual cells are a black box of patient material and consequently their molecular mechanisms are yet to be fully understood.

Using an inducible mouse model of breast cancer, we modelled residual disease using organoid cultures and characterized them during tumor regression using IF staining. We further explored this dynamic process of tumor regression and residual disease establishment on a transcriptomic level, using RNA sequencing at hallmark time points. Identification of hallmark time points during tumor regression coupled with kinetic transcriptome data further expanded our understanding of the pathways activated when a tumor organoid is no longer expressing oncogenes and is undergoing massive apoptosis. In addition to oncogene dependant signalling pathways, our sequencing data revealed a transcriptomic signature unique to residual tumor cells and provided the possibility to explore associated molecular pathways that could be responsible for their survival. Once equipped with this information that we derived from clean, reductionist 3D culture organoids, we aimed to validate them in the immune proficient conditions of the mouse model.

We performed an in depth characterization of the regressing mammary glands of mice using IHC staining and the transcriptomic analysis performed along the course of *in vivo* mouse tumor regression was overlaid onto the *in vitro* regression transcriptome. Molecular targets identified in the organoid cultures and validated in the mouse mammary glands, such as c-Jun, BCL6 and players of the Nf-κB cascade, were then evaluated for their synthetic lethality *in vitro*, although more robust, sensitive and relevant methods of assessing synthetic lethality need to be established.

In a more biased approach, to unravel the mechanisms of residual disease establishment in our *in vitro* tumour organoids, we set up long-term live cell imaging of regressing organoids using light sheet microscopy. Imaging movies that we recorded show some interesting anecdotal characteristics exhibited by cells during regression — cell movement to the basement membrane and repolarization into the epithelial monolayer as a method of survival. Athough our efforts in unravelling the spatio-temporal dynamics of tumor regression are currently impeded by image analysis limitations, we intend to circumvent them as detailed in the project above.

CHAPTER 4

MATERIALS AND METHODS

Contributions:

This thesis contains experiments performed by me in conjunction with members of the Jechlinger Group and analyzed with the help of various collaborators at EMBL, Heidelberg.

I cultured mammary organoids for all experiments, genotyped mice, performed regular checks of the mouse colony, harvested in vivo samples and performed the various analyses using bright-field, confocal and light-sheet microscopy. I also performed all the immunohistochemistry stains, immunofluorescence stains, Real-Time qPCR analysis, RNA extractions and image analysis workflow implementations described in this thesis.

Ksenija Radic (Jechlinger Group, Cell Biology and Biophysics Unit, EMBL Heidelberg) contributed significantly to the in vitro culture experiments including doxycycline titration, IF staining and collection and extraction of RNA.

Lucas Chaible (Jechlinger Group, Cell Biology and Biophysics Unit, EMBL Heidelberg) set up the stochastic tumorigenesis system including cloning, organoid transduction and lentivirus production experiments and contributed to imaging experiments.

Sabine Reither (Advanced Light Microscopy Facility, EMBL, Heidelberg) was consulted during imaging experiments and helped set up light sheet imaging.

Christian Tischer (Center for Bioimage Analysis, EMBL, Heidelberg) created the tools for SPIM data processing and segmentation. He also consulted on Imaris software usage and contributed to figure design and manuscript writing.

Jean-Karim Heriche performed computational feature analysis for single cells of the stochastic tumor organoids.

Katharina Zirngibl (Patil group, Structural and Computational Biology Unit, EMBL Heidelberg) performed RNA-sequencing data analysis (differential expression analysis, clustering analysis, correlation analysis for gene expression profiles and manual curation of clusters and gene targets).

Matt Rogon (Centre for Biomolecular Network Analysis, EMBL Heidelberg) performed the Gene Enrichment analysis and Network analysis for identification of synthetic lethal targets.

Ksenija Radic, Savannah Jackson and Marta Garcia Montero (all from the Jechlinger Group, Cell Biology and Biophysics Unit, EMBL Heidelberg) helped maintaining the mouse colony and harvesting samples. Marta Garcia Montero also provided technical assistance with histology.

Sylwia Gawrzak (Jechlinger Group, Cell Biology and Biophysics Unit, EMBL Heidelberg) helped design and perform experiments for synthetic lethal inhibitor testing using in 3D vitro cultures

Yuanyuan Chen (Sotillo Group, DKFZ, Heidelberg) helped perform IHC experiments and subsequent imaging.

RNA sequencing and RNA microarray-chip runs were performed at the Genomic Core Facility at EMBL.

Dr. Martin Jechlinger supervised the project, designed experiments and contributed to manuscript writing.

Publications:

1. Tracking the cells of tumor origin in breast organoids by light sheet microscopy

Alladin, A and Chaible, L and Reither, S and Löschinger, M and Wachsmuth, M and Hériché, JK and Tischer, C and Jechlinger, M

bioRxiv, 2019, 617837; DOI: https://doi.org/10.1101/617837.

2. Towards a holistic and mechanistic understanding of tumourigenesis via genetically engineered mouse models

Ashna Alladin and Martin Jechlinger

Current Opinion in Systems Biology, 2017; Volume 6, Pages 74-79, 2452-3100. DOI: 10.1016/j.coisb.2017.10.004

3. Metabolic shifts in residual breast cancer drive tumor recurrence

Kristina M. Havas, Vladislava Milchevskaya, Ksenija Radic, **Ashna Alladin**, Eleni Kafkia, Marta Garcia, Jens Stolte, Bernd Klaus, Nicole Rotmensz, Toby J. Gibson, Barbara Burwinkel, Andreas Schneeweiss, Giancarlo Pruneri, Kiran R. Patil, Rocio Sotillo, Martin Jechlinger

J Clin Invest. 2017; 127(6):2091-2105. DOI:10.1172/JCI89914.

4.1 Animal models

The mouse strains TetO-MYC/ MMTV-rtTA [87] and TetO-Neu/ MMTV-rtTA [48], that have been described previously, were bred in order to establish the tri-transgenic strain TetO-MYC/TetO-Neu/ MMTV-rtTA or bi-transgenic strain TetO-MYC/TetO-Neu. Reporter H2B-mCherry was crossed into the B and T lines using a R26-H2B-mCherry line [197] (RIKEN, CDB0239K). The mice were all bred in FVB background. Animals were kept on a 12-hour light/12-hour dark cycle, with constant ambient temperature (23±1°C) and humidity (60±8%), with pellet food and water ad libitum. Food pellets containing doxycycline (doxycycline hyclate 625 mg/kg) were obtained from Envigo Teklad. Breeding and maintenance of mouse colony was done in LAR (Laboratory Animal Resources) facility of EMBL Heidelberg, under veterinarian supervision and in accordance to the guidelines of the European Commission, revised Directive 2010/63/EU and AVMA Guidelines 2007.

4.1.1 Genotyping

After weaning, the LAR facility at EMBL, Heidelberg provided tail/ear cuts from all the mice and their genotypes were then determined by PCR on genomic DNA. Genomic DNA was extracted by digesting the tail/ear piece in 75 μl of digestion buffer (NaOH 25 mM, EDTA 0.2 mM) in an Eppendorf tube at 98°C for 1 hour at 1000 rpm shaking. Then, the solution was neutralized by addition of 75 μl Tris-HCl (40 mM, pH 5.5). The tube was then centrifuged at 4000 rpm for 3 min. 2 μl of the supernatant was used in the PCR reactions. Primer sequences for the transgenes used are as follows: Tet-O Neu (Forward: GACTCTCTCTCCTGCGAAGAATGG and Reverse: CCTCACATTGCCAAAAGACGG); Tet-O MYC (Forward: TAGTGAACCGTCAGATCGCCTG and Reverse: TTTGATGAAGGTCTCGTCGTCC); MMTV-rtTA (Forward: GTGAAGTGGGTCCGCGTACAG and Reverse: GTACTCGTCAATTCCAAGGGCATCG). Agarose gel-electrophoresis was used for the detection of PCR products (MYC 630 bp, Neu 386 bp, rtTA 380 bp), which was done on the 1.5 % agarose (Sigma, Cat. #A9539-500G) gel with Ethidium bromide solution in a final concentration of 0.5 μg/ml (Sigma, Cat. # E1510-10ML). The products were visualized using Quantum-Capt1 documentation system and instrument (Peqlab).

4.1.2 *In vitro* experimental design

Female virgin mice of the desired genotype and between 8-10 weeks old were sacrificed by asphyxiation and dissected with a small cut at the level of the pubis, followed by a median longitudinal cut superior to the chin. The skin was then dissected and turned on one side exposing the the subcutaneous, superficial lymph nodes and mammary glands. All ten mammary glands were harvested, digested and singularized for establishing organoid cultures

4.1.3 *In vivo* experimental design and harvest

In vivo tumour formation and regression was achieved by regulating doxycycline in the animal diet for TetO-MYC/TetO-Neu/ MMTV-rtTA animals. For induction of tumours, food pellets supplemented with doxycycline (625 mg/kg) was given to the mice once they reached 8 weeks of age. Animals were monitored weekly for tumor development and overall health. Full blown tumors developed in the period of 4-8 weeks in triple transgenic mice. When the tumour burden was too large (diameter of the largest nodule was more than 2 cm), animals were given food without doxycycline which resulted in the fast tumour regression to a non-palpable state. After 9 weeks of doxycycline withdrawal from the animal diet the tumours were considered fully regressed. Wild-type (or non-inducible) siblings were treated in the same way alternating food containing doxycycline. They represented age-matched controls and were harvested at the same time and in the same manner. Harvesting triple transgenic animals at the point of high tumor burden or along the regression phase (after doxycycline withdrawl) was done in an organized fashion. Before harvesting, the mice were evaluated for the phase of estrous cycle they were in, by taking vaginal smears according to the modified protocol in [198]: 20-40 μl of 1x Phosphate Buffered Saline (PBS) was used to obtain vaginal lavage. The smears were dried on the slide at room temperature and further fixed in 10 % formalin, washed in 1x PBS, stained with Crystal Violet solution (Sigma, Cat. # V5265) washed in tap water and visualized using Leica Application Suite X and Leica DFC7000 T microscope (Leica Microsystems). If the mice were found to be in Diesterous phase of the esterous cycle, harvest was postponed by 1-3 days. If the esterous cycle phase was evaluated to be ideal for harvest, the mouse was first sacrificed by asphyxiation. Then the external appearance was recorded by photographing. The mouse was dissected open and all affected mammary glands were removed and the tumour material was partitioned for paraffin embedding (for histology staining), OCT embedding (for cryo-histology) and tissue freezing (in liquid nitrogen to preserve DNA, RNA and protein).

4.2 3D Cell culture

4.2.1 Matrigel culture

Mammary glands harvested from mice (see above), were digested in order to prepare a single cell solution. For this, the tissue was divided in four loosely capped 50 ml falcons, each supplemented with 5 ml serum-free media (DMEM/F12 supplemented with 25mM HEPES and 1% Pen Strep (100 U/ml Penicillin; 100 μg/ml Streptomycin; ThermoFisher Cat. # 15140122)), 750 U of Collagenase Type 3 (Worthington Biochemical Corp, Cat. # LS004183), 20 μg of Liberase (Roche Cat. # 5401020001) and incubated overnight (not more than 16 hours) at 37°C and 5%CO_2. The glands were then mechanically disrupted using a 5 ml pipette, and washed in PBS before being pelleted at 1000 rpm for 5 minutes. The interphase between upper fat layer and cell pellet was removed and 5 ml of 0.25% Trypsin-EDTA (Invitrogen, Cat.# 25200-056) was added and incubated for 45 minutes at 37°C and 5%CO_2. The enzymatic reaction was then neutralized using 40 ml of serum supplemented media (DMEM/F12 with 25mM HEPES, 1% Pen Strep and 10% FBS Tetracycline Free certified (Biowest Cat. # S181T)). The cells were pelleted again, resuspended in Mammary Epithelial Cell Basal Medium (PromoCell Cat. # C-21210) and seeded in collagen coated plates (Corning Cat. # 354400) overnight at 37°C and 5% CO_2. This allows for epithelial cells to adhere to the surface of the plates while the other cell types float on top in the media and can be easily removed by vacuum suction. The epithelial cells were detached from the collagen coated plates by incubating them with 0.25% Trypsin-EDTA for 5-7 minutes at 37°C and 5%CO_2, following inactivation with serum supplemented media. The single cell solution was pelleted, resuspended in MEBM and counted. For seeding 3D culture gels, typically 10,000-12,000 cells we mixed rapidly on ice with the prepared Matrigel-collagen mixture – 4:2:1 proportions of Matrigel Matrix basement Membrane growth factor reduced phenol red free (Corning Cat. # 356231), Cultrex 3D Collagen I rat tail (TEMA Ricerca, Cat. # 3447-020-01) and cell suspension diluted in PBS(1x) respectively. Matrigel-collagen-cell mixture droplets 100 μl in volume , were then dispensed into flat bottom wells (Corning CellBIND 12 Well Clear Multiple Well Plates, Cat. # 3336) and incubated for 30-40 minutes at 37°C and 5%CO_2, until the matrigel solidified. The gels were supplemented with 1.5 ml MEBM and allowed to grow at 37°C and 5%CO_2.

For induction of oncogenes in the cells of the organoids, doxycycline (Sigma Cat. # D9891) was supplemented in the media. During oncogene de-induction of the organoids the media supplemented with doxycycline was removed from the plate wells and the gels were washed thrice for 10 mins, using serum supplemented media, PBS(1x) and serum free media, in that order.

For RNA seq sample collection experiments, from the start of experiment until collection, media was changed every alternate day at precisely the same time. An extra media change was performed after 24 hours following the de-induction of oncogenes by removing

doxycycline.

4.2.2 Organoid transduction

For transduction of organoids either with Inducer-reporter (pLenti-rtTA-GFP) lentiviral particles or Reporter (pLenti-GFP) lentiviral particles transduction, Matrigel cultures were established as mentioned above. After 3 days of growth, the gels were mechanically disrupted and placed in a 15 ml falcon. Two disrupted gels were placed in one 15 ml falcon with 2ml of MEBM supplemented with 25U of Collagenase type I and 5 μg of Liberase. Following incubation in this solution for 2 hours at 37°C and 5%CO_2, when the matrigel was totally digested, the organoids were washed 3 times with 15 ml of serum supplemented media and once with 15 ml of serum free media, and pelleted at 1000 rpm for 5 minutes. We then supplemented the organoid pellet (from two original gels) in 10 μl of MEBM and added 6 x 10^5 lentivirus particles to the solution. We then mixed this solution with 90 μl matrigel and plated it in 35 mm dishes (Greiner Bio-One Cat. # 627160) and placed in incubator for 30-40 minutes until the matrigel solidified. The gels were supplemented with 3 ml MEBM and incubated for 2 days at 37°C and 5%CO_2 in order to allow for organoid recovery and lentiviral gene expression.

For induction of oncogenes in the cells of the organoids, doxycycline (Sigma Cat. # D9891) was supplemented in the media. 800 ng/ml of doxycycline was used to induce tri-transgenic organoids and 600 ng/ml was used for the bi-transgenic organoids. qPCR analysis was used to standardize the doxycycline dosage for bi-transgenic organoids.

4.2.3 Target validation pipeline

For testing synthetic lethality of identified targets, target inhibitors were supplemented in the media during tumour regression of organoids grown in 96 well plates (Falcon® 96-well Black/Clear Flat Bottom TC-treated Imaging Microplate with Lid, Corning, Cat. # 353219). Brightfield imaging was performed to follow the organoid regression visually and cell assays were implemented at end points to ascertain cell viability and cell death.

Mouse mammary glands were harvested and processed into single cell suspension as detailed above. Matrigel-cell suspension mixture was prepared on ice as usual, consisting of Matrigel:CellSuspension ratio in 2:1 part. The mixture contained 10,000 cells per 100 μl. Small 5 μl gels were seeded into the wells of a 96 well plate from this mixture, resulting in 500 cells dispensed into each well. The gels were incubated for 30-40 minutes at 37°C and 5%CO_2, until the matrigel solidified. Using a multi-channel pipette, 200 μl of MEBM was supplemented into each well of the 96 well plate and the gels were allowed to grow at 37°C and 5%CO_2. Wells on the corners were not used to prevent edge-effect of assay readouts and other anomalies in culture systems frequently experienced.

Media was replenished every alternate day to support growth of single cells into organoids. After 4-5 days of normal growth, the media was supplemented with 200 ng/ml

of doxycycline. Induction of oncogenes with doxycycline was carried out for 5 days via usual media replenishment. After 5 days of maintaining the 3D cultures on doxycycline, the gels were washed thrice for 10 mins, using serum supplemented media, PBS(1x) and serum free media, in that order. MEBM with increasing concentrations of the inhibitor of choice was added along with the necessary controls. Five technical replicates were analysed per condition. Normal uninduced cells, grown alongside induced tumours, were used as internal controls to test for the toxicity of the inhibitor being tested at all concentrations. Gels were supplemented with media and supplements every alternate day as usual. Seven days after de-induction and addition of inhibitors, the experiment had reached it endpoint and was processed for cell viability and cell death assays.

Inhibitors tested via this assay included: BCL6 Inhibitors: 79-6 (Calbiochem, Cat. # 197345), FX-1 (MedKoo, Cat. # 407501) c-Jun Inhibitors: Tanshinone IIA -AP-1 inhibitor (Enzo, Cat. # BML-GR336-0005), SP 600125 (Tocris, Cat. # 1496/10). Nf-kB Pathway Inhibitors: CCL5 Inhibitor- Maraviroc (Sigma, Cat. # PZ0002), IKK-2 Inhibitor IV - CAS 507475-17-4(Merck, Cat. # 401481), IKK-2 Inhibitor VIII - CAS 406208-42-2 (Merck, Cat. # 401487)

ScanR imaging and data representation

Inhibitor testing experiments performed in all 96-well plates were imaged on the ScanR (High Content Screening Station, Olympus) immediately after de-induction and at "Day0" of inhibitor treatment and then again after seven days at "Day 7" of inhibitor treatment. 4x objective was used to image the 3D cultures in 4 quadrant fields of view per well. 21 images were recorded at z-step intervals of 100 μm in each of the four fields of view to cover the entire three-dimensional range of organoids occurring in the gel. The z stacks in each field of view were processed to show a "maximum projection" image and all 4 maximum projections per well were "stitched" together to get an overview of the whole gel drop in each well. Custom-made codes, including the macros created by the Advance Digital Microscopy Core Facility at IRB Barcelona used for image analysis, are available upon request from the authors of Gawrzak, et al., 2018 [199].

Cell Titre Glow and CellTox assay multiplexing

At the end point of each inhibitor validation experiment, i.e., after seven days of oncogene de-induction and Day 7 of inhibitor treatment, the cultures were assayed for cell viability and cell death. After removal of 100 μl of media from each well, 20 μl of the working solution of CellTox^{TM} Green Cytotoxicity Assay (Promega, Cat. # G8741) was added to the media in each well and the whole plate was gently shaken at room temperature for 1 hour. The CellTox^{TM} Green Cytotoxicity Assay is based on detection of biomarkers released into the medium upon cell death and membrane perforation. The substrate in the assay binds to the cytotoxic indicator markers and produces a green fluorescence signal that is

proportional to overall cell death in each matrigel culture drop. The fluorescence from the whole 96-well plate was recoded at the Chemical Core Facility at EMBL Heidelberg, using a EnVision^{TM} Multilabel Plate Reader (PerkinElmer).

After recording the cell death, 50 μl of CellTiter-Glo® Luminescent Cell Viability Assay (Promega, Cat. # G7570) was added to the media in each well and the whole plate was gently shaken at room temperature for 1 hour. CellTiter-Glo® Luminescent Cell Viability Assay is a method of determining the number of viable cells in culture based on quantitation of the ATP present, an indicator of metabolically active cells. The assay results in cell lysis and generation of a luminescent signal proportional to the amount of ATP present. The amount of ATP is directly proportional to the number of cells present in culture. It generates a "glow-type" luminescent signal that was measured similarly, using a EnVision^{TM} Multilabel Plate Reader (PerkinElmer), at the Chemical Core Facility at EMBL Heidelberg, in RLU (relative luminescence units).

Raw data from the plate reader was analysed to ascertain the inhibitor's effect on the never induced and regressed organoids. This was done by constructing a dose-response curve. A dose-response curve describes the relationship between increasing the dose (or concentration) of the drug (inhibitor) and the change in response that results from this increase in concentration. The concentration ranges spanned by the inhibitors are indicated in Table 3.3. The dose in these curves is usually represented using a semi-logarithmic plot. On a semi-logarithmic plot, the amount of drug(inhibitor) is plotted (on the X axis) as the log of drug concentration and response is plotted (on the Y axis)using a linear scale. Dose response curve analysis and calculation of IC_{50} values for each drug and assay was performed using the GraphPad 6 (Prism) statistical analysis software.

4.3 Immunofluorescence staining

Matrigel cultures required for immunofluorescence analysis were grown as described in Section 4.2 on either Nunc™ Lab-Tek™ II chambers (Thermo Cat. # 155382) or on glass slides placed in 12-well tissue culture plates (Corning CellBIND 12 Well Clear Multiple Well Plates, Cat. # 3336). At pre-defined timepoints, the gels were transferred to deactivated clear glass screw neck vials (Waters Cat.# 186000989DV) and fixed using 4% PFA for 2-3 mins, following three 10 min washes with PBS (1x) and one 20 min wash with IF buffer (containing NaCl, Na2HPO4, NaN3, BSA, TritonX-100, Tween-20; pH 7.4). The gels were blocked with Blocking Buffer (1x IF buffer with 10 % goat serum (Jackson Immuno Research Cat.# 005-000-121)) for 2 hours at room temperature. The gels were incubated with primary antibodies (diluted in Blocking Buffer) overnight at 4°C and washed 3 times the next day in 1x IF buffer for 20 mins per wash. Glass vials were then covered in aluminum foil and the gels were incubated in blocking buffer with secondary Alexa antibodies and DAPI (4', 6'-diamino-2-phenylindole) (ThermoFisher Cat. # 62248, ddilution 1:1000) for 1 hour. Following incubation with secondary Alexas and DAPI, the gels were washed briefly in PBS(1x) for 5 mins. The gels were then transferred to a Nunc™ Lab-Tek™ II chamber and mounted in anti-fading mounting medium VECTASHIELD® Mounting Medium with DAPI (Vecto Cat. # H1500-10). Stained gels were imaged on Leica SP5 confocal microscope using 63x water lens and the LAS AF imaging software.

The following primary antibodies and dilutions were used for Matrigel cultures: c-MYC (Cell Signaling Technologies, Cat. # D84C12, dilution 1:900), alpha6-integrin (Millipore Cat. # MAB1378, dilution 1:80), ZO1 (Life Technologies Cat. # 61-7300, dilution 1:500), GM-130 (BD Biosciences Cat. # 610823, diluted 1:100), E-cadherin (Invitrogen Cat. # 3-1900, diluted 1:200) , Cleaved Caspase3 (Cell Signalling, Cat. # 9661S, diluted 1:100), c-Jun (Cell Signalling (60A8), Cat. # 9165, diluted 1:100) and BCL6 (Atlas antibodies, Cat. # HPA004899, diluted 1:50). Anti-rabbit, anti-mouse, and anti-rat antibodies were purchased coupled with Alexa Fluor dyes (Invitrogen, Cat.# A21247, A11034, A11036) for secondary staining.

4.4 Immunohistochemistry staining

Harvested mammary gland tissue and tumor tissue during progression and regression was fixed in 10 % formalin solution (neutral buffered; Sigma, Cat.# HT501128-4L) overnight on a shaker to allow optimal penetration of fixing agent. Subsequent to embedding in paraffin via standardized procedures, the embedded tissue was sectioned at 5 μm thickness with a microtome (Leica, Cat.# RM 2135) and placed on slides (Superfrost , Cat.# VWR630-0954). Tissue sections were incubated at 37°CC for 16 hours before storage and further staining. Staining by Hematoxylin QS (Vector, Cat.#H-3404) and Eosin 1 % Aqueous (RA Lamb, Cat.# LAMB/100-D) was done following the standard protocol and samples were analyzed at microscope (Leica, Cat.# LMD 7000).

Cleaved Caspase3 (Cell Signalling, Cat. # 9661S, diluted 1:100), c-Jun (Cell Signalling (60A8), Cat. # 9165, diluted 1:350), BCL6 (Atlas antibodies, Cat. # HPA004899, diluted 1:750) and Ki67 (Vector, Cat.# VP-K451, diluted 1:100) antibody staining was done on FFPE tissue following the standard IHC protocol: deparaffinization and rehydration of the samples, followed by antigen retrieval using citric acid-based Antigen unmasking solution (Vector, Cat.# H-3300) for 30 min in a steamer and inactivation of endogenous hydrogen peroxidase activity with 10 % H2O2 solution (Sigma, Cat.# H1009) for 10-15 mins. Blocking was done using 10 % Normal goat serum (Jackson Immuno Research, Cat.# 00500121) in 1x Phosphate Saline Buffer (PBS) for 2 hours, following incubation with primary antibody (diluted in blocking buffer) at 4°CC overnight. Next day, the slides were washed 3 times for 5 mins each using PBS(1x). Incubation with biotinylated antibody from Vectastain Elite ABC HRP Kit (Peroxidase, Rabbit IgG; Vector Laboratories, Cat.# PK-6101) was done for 30 min, followed by washing and incubation with Horse Radish Peroxidase (HRP) conjugated antibody. Detection of antibody localization was done using DAB Peroxidase (HRP) Substrate Kit (Vector, Cat.# SK-4100). Counterstaining was done using Hematoxylin QS (Vector, Cat.#H-3404), after which the sections were dehydrated, mounted with DPX Mountant for histology (Sigma, Cat.# 06522) and analyzed using a wide-field microscope (Leica, Cat.# LMD 7000) equipped with Leica CD310 digital camera and LASV3.7 (Leica) software.

4.5 qPCR analysis

The qPCR technique was performed following the MIQE guidelines, where the total RNA was isolated from the mammary gland organoids using RNA PureLink Mini Kit (ThermoFisher Cat. # 12183018A) and 2.5ug was reverse transcribed to cDNA using SuperScript VILO cDNA Synthesis Kit (ThermoFisher Cat. # 11754050). Using Primer3 software we designed specific primers for DNA intercalating fluorescent dye approach for the transgenes Neu (Forward: CGTTTTGTGGTCATCCAGAACG and Reverse: CTTCAGCGTCTACCAGGTCACC) and c-MYC (Forward: GCGACTCTGAGGAGGAACAAGA and Reverse: CCAGCAGAAGGTGATCCAGACT). As endogenous controls, mCherry (Forward: GAGGCTGAAGCTGAAGGAC and Reverse: GATGGTGTAGTCCTCGTTGTG) and Pum1 (Forward: AATGTGTGGCCGGATCTTGT and Reverse: CCCACAGTGCCTTATACACCA) were used. Primer efficiency was verified and established between 95% and 105% Each sample was analyzed in duplicate and non-template controls were used in each qPCR run. Analyses were carried out using a StepOne device (Applied Biosystems, USA). Analysis of relative gene expression data was performed according to the 2ΔΔCq method and the results were expressed as fold change of ΔΔCq values obtained from the reference T800 organoids.

4.6 Lentivirus cloning and production

The lentivirus design is based on pWPXL backbone, which was a gift from Didier Trono (Addgene Cat. #12257). The coding region from the original plasmid was excised using ClaI and NdeI in order to insert a new multiple cloning site (MCS). The pGK promoter was PCR amplified from pLVPT-GDNF-rtTR-KRAB-2SM2, which was a gift from Patrick Aebischer Didier Trono (Addgene Cat.#11647) and cloned using XhoI and EcoRI restriction sites. For the plasmid pLenti-rtTA-GFP the synthetic region rtTA-p2A-H2B-GFP was cloned downstream of the pGK promoter using EcoRI and NheI sites. The plasmid pLenti-Null-GFP is derived from the pLenti-rtTA-GFP by removing the rtTA sequence, using the restriction sites EcoRI and BamHI, and retaining H2B-GFP in the coding region. For production of lentivirus particles, we seeded 1.6 x 10^7 HEK-293T cells (Lenti-X - Clontech Cat. # 632180) in 500cm^2 square dishes (Corning Cat. # 431110). After 24 hours, the cells were supplemented with media containing 25uM of chloroquine diphosphate (Sigma-Aldrich Cat. # C6628). After a 5-hour incubation, using 360 μg of polyethyleneimine (4 μg for each μg of plasmid), we transfect the cells with a mixture of endotoxin free plasmids: 20 μg pCMV-VSV-G (Addgene Cat. #8454); 30 μg psPAX2 (Addgene Cat.#12260); 40 μg transfer plasmids pLenti-rtTA-GFP or pLenti-Null-GFP. We harvested the media 48 hours, 72 hours and 96 hours post transfection. Concentration of the lentivirus from the collected media was performed using an ultracentrifuge (Beckman Sw32 rotor) at 25,000 rpm for 2 hours at 4°C. The lentivirus pellet was resuspended in 1000 μl of HBBS buffer, aliquoted and stored at -80°C. The lentivirus titer was measured using FACS analyses as described by Kutner and colleagues [200].

4.7 Selective Plane Illumination Microscopy

4.7.1 Sample holder preparation and mounting

Imaging was performed on the InVi SPIM (Inverted light-sheet microscope from Luxendo Light-Sheet, Bruker Corporation). Sample mounting for the InVi SPIM is suitable for 3D matrigel cultures that are used to grow and transduce mammary organoids (see above). The sample holder is made of medical grade plastic (PEEK). A 25 μm thin membrane (FEP; Luxendo) with a refractive index matching that of water is glued to the upper surface of a groove in the sample holder with a biocompatible silicone glue (Silpuran 4200; Wacker), forming a trough with transparent bottom. Matrigel cultures were carefully cut with a scalpel into rectangular slivers and transferred onto the FEP membrane's trough. Once the gel sliver was aligned in place, 20-30 μl of fresh matrigel drops were poured onto the gel sliver in the sample holder until there was a thin layer of liquid matrigel on top of the gel sliver. The setup was incubated for 20 minutes at 37°C in a 5% CO_2 incubator to allow the matrigel layer on top to solidify. Once the gel was solidified, 600-800 μl of MEBM supplemented with/without doxycycline was added to the sample holder's FEP sheet trough. Preferably, freshly mounted sample gels were allowed to settle overnight in the incubator to prevent any gel drift during imaging, when the holder is placed into the imaging chamber of the microscope. The imaging chamber acts as an incubator with environmental control and it has a reservoir for immersion medium, which is filled with water so that both objective lenses and the bottom of the sample holder are below the water surface.

4.7.2 Imaging configurations and conditions

The InVi SPIM is equipped with a Nikon CFI 10x/0.3NA water immersion lens for illumination and a Nikon CFI-75 25x/1.1NA water immersion lens for detection. For excitation of GFP and mCherry, 488 nm and 594 nm laser lines were used, respectively, while emission was selected using a 497-554 nm band pass filter and a 610 nm long pass filter, respectively. 3D image stacks were acquired with a light-sheet thickness of 4 μm, a final magnification of 62.5x, resulting in 104 nm pixel size. The In-Vi SPIM environmental control was set to 37°C, 5% CO_2 and 95% humidity. A series of optimization experiments, involving different laser powers, exposure times and z-step sizes yielded laser powers of 13 μW for 488 nm and 36 μW for 594 nm, 100 millisecond exposure time per frame and 1μm z-spacing between frames to be optimal for long term imaging (96-120 hours) without photo-bleaching or photo-toxic effects on growth. Images were recorded as 2D planes ranging from 100-500 in number, depending on the organoid size. Each 3D stack of planes was recorded in 2 channels - mCherry (all cells) and GFP (transduced cells). Depending on the duration of the time lapse imaging, 450-600 image stacks (equivalent to 72-96 hours) were recorded per organoid at 10-minute intervals.

4.7.3 Big-image data analsysis workflow

Big Data Processor, a Fiji plugin for lazy loading of big image data, was used to visualize the images in 2D slicing mode, crop stacks in x, y, z, and t, bin images (3 x 3 x 1 in x, y, z), perform chromatic shift correction between channels and convert .h5 files from the InVi SPIM into an Imaris compatible multi-resolution file format (.ims) for further analysis [90]. The oncogenic cells (H2B-GFP channel) displayed heterogeneous morphologies as well as varying intensity textures, making it difficult to segment them using conventional thresholding approaches. We thus used a trainable segmentation approach to convert the raw intensity values into pixel probability maps, using the Fiji plugin called CATS [91]. Using the H2B-GFP channel images as input, we trained three pixel classes: background, nucleus center and nucleus boundary. For training we drew about 20(background), 120(nucleus center), 100(nucleus boundary) labels distributed across the different time-frames of the movie. After feature computation and training of a Random Forest classifier the whole dataset was processed on EMBL's high performance computer cluster. The segmentation of one data set -typically 100 timepoints- is distributed across few hundred jobs, each job using 32 GB RAM, 16 cores, and running for about 30 minutes. The nucleus center probability maps were then exported from CATS and added as an additional channel to the converted intensity data.

The data were then loaded into Imaris [89] for 3D visualization and further processing. Using the Imaris' Surfaces function, we segmented the nucleus center probability maps into objects. To do so, probability maps were manually thresholded, using a surface smoothening parameter of 0.3 μm; the minimum quality parameter for seed points was set to 0.1, and object splitting was applied for objects larger than 5.5 μm. Objects with volumes less than 20 μm3 were excluded. Next, all objects were tracked over time using Imaris' Lineage tracking algorithm with a maximum distance between objects in subsequent time-points limited to 10 μm and a maximum gap size between identification of the object in a particular track limited to 10 time points. Most of the errors in the object segmentation were false merges, where two cells were segmented as one (Figure 2.11. This kind of error is frequently not sustained in the previous or following time-points and the maximum gap size parameter of the tracking algorithm thus frequently provides correct tracks nonetheless. The resulting lineage trees of proliferating tumour cells within the organoid were corrected manually within Imaris, e.g., excluding apoptotic cells and auto-fluorescent debris. Center of mass coordinates of each cell were measured and exported from Imaris for subsequent feature analysis.

4.7.4 Computational feature analysis

Observations suggest that tumors in organoids originate from clusters of oncogene-expressing cells produced by independent transduction events. To identify these clusters, we computed the pairwise Euclidean distances between all oncogene-expressing cells in an organoid

at the start of the experiment and applied hierarchical clustering with complete linkage. Clusters were identified automatically by cutting the branches of the trees using the dynamic tree cut algorithm [92]. This defined a cluster as a group of oncogene-expressing cells that are closer to each other than to other oncogene-expressing cells of the same organoid. Note that a cluster can be composed of a single cell if this cell is comparatively isolated from other transduced cells. For each cluster we identified the following features as possibly linked to tumor formation:
(1) number of cells in the organoid
(2) cell density expressed as the ratio of number of cells to organoid surface area computed by assuming the organoid is a sphere with diameter equal to the distance between the two most distant cells
(3) number of oncogene-expressing cells in the organoid
(4) number of cells (including both oncogene-expressing and normal cells) in the cluster volume defined as the sphere centered at the center of mass of the cluster with diameter equal to the distance between the two farthest oncogene-expressing cells of the cluster
(5) number of oncogene-expressing cells in the cluster
(6) average pairwise distance between all cells in the cluster volume
(7) average pairwise distance between oncogene-expressing cells in the cluster
(8) fraction of oncogene-expressing cells in the cluster volume
(9) number of contacts between oncogene-expressing cells in the cluster. Two cells are presumed in contact if they are less than the average cell diameter + 2 standard deviation apart.

Oncogene-expressing cells were tracked over time and a cluster was associated with a tumor outcome if any of its cells lead to tumor formation. To identify which features were linked to this outcome, we took an information-theoretic approach to model selection. We fitted a logistic regression model for all possible linear combinations of features and selected the best model based on the Akaike information criterion (with correction for small sample sizes) [93]. This model included only three features: number of oncogene-expressing cells in the cluster, number of oncogene-expressing cells in the organoid and number of cells in the cluster of which only the first (number of oncogene-expressing cells in the cluster) contributed significantly to tumor formation with an odds ratio of 9.1. Computing relative variable importance across all models also indicated that the number of oncogene-expressing cells in a cluster is the most important feature.

4.8 RNA Seq data analysis

4.8.1 RNA extraction

For RNA sequencing analysis of organoids in 3D culture, 2 gels of 100μl volumes were pooled during harvest. At the predetermined timepoint, media from the wells were removed and gels were digested using 900 μl of mirVana lysis buffer, and subsequently extracted using mirVana miRNA Isolation Kit (Ambion, Cat. # AM1560). RNA quality and concentration was assessed using the Bioanalyzer (Agilent 2100, Cat. # G2939BA) at the Genomics Core facility, EMBL, Heidelberg. RNA was sequenced in the Genomics Core on Illumina NextSeq 500 platform, using NGS HiSeq protocol – 75bp read length, 30x coverage.

4.8.2 RNA Seq data analysis

After assessing the quality of the raw RNA sequencing reads by FastQC version 0.11.3 [201], adapter trimming using cutadapt version 1.9.1 [202] with default options providing the standard Illumina TrueSeq Index adapters was done. FaQCs version 1.34 [203] was used for subsequent quality trimming and filtering, applying following parameters: -q 20 -min L 25 -n 5 -discard 1. This resulted in 34,1 to 52 million total reads per sample. Reads were further aligned to the reference genome (mouse-GRCm38.p4) using Tophat2 version 2.0.10 [204] with the following parameter: -G -T -x 20 -M - -microexon-search - - nocoverage-search - -no-novel-juncs-mate-std-dev 100-r 50 -min-segment-intron 20 -I 30 -a6. For differential expression analysis, only reads with unique mappings were considered. The count HTSeq python library version 0.6.1.pl. script [205] with default options was used to obtain gene level count tables. All reads mapped in total to 19500-20800 genes and the statistical analysis was done with Bioconductor package DESeq2 version 1.12.4 [206]. Normalization was done based on size-factor to control for batch effects and inter-sample variability. Package defaults were used for dispersion estimation and differential expression analysis, which include multiple testing correction, independent filtering and Cook's cutoff [207] for outlier detection. For comparison of gene expression between residual versus normal cells, DESeq2 raw p-values were used as input to fdrtool version 1.2.15 [208] to compute q-values. Genes with q-value0,05 were considered as significantly differentially expressed. For comparison of gene expression between tumor versus normal cells, Bonferroni correction was used for multiple testing correction and genes with padj 0,1 were considered as significantly differentially expressed. R V.3.3.1. (R Development Core Team) was used for conducting biostatistical analyses. For performing dimensionality reduction with Principal Component analysis (PCA) and hierarchical clustering rlog DESeq2 transformed transcript counts were utilized. For the calculation of the ellipses on the PCA plots the "sat ellipse" function from the R package ggplot2 was used.

4.8.3 Clustering and Correlation analysis

Clustering analysis using GPClust [209], yielded over 2000 genes into 75 clusters with FDR (false discovery rate) of 1%. Individual analysis of gene trajectories in interesting-trajectory clusters was used to classify interesting trajectories into 5 broad manually curated clusters:
I. Sustained targets – upregulated consistently during regression up to MRD state
II. Early responders – unregulated early during regression before apoptosis decisions are made (up to 8 hours after de-induction)
III. Intermediate responders – unregulated after apoptosis decisions are made to potentially salvage residual cells/aid in tissue rearrangement to increase survival (12-24 hours after de-induction)
IV. Late targets – unregulated during re-polarization of epithelial cells into residual rim (36 hours after de-induction onwards)
V. Shoulder trajectories – Upregulated in tumors and continued sustained expression during regression (at least up to 34 hours after de-induction).

Correlation analysis performed by Katharina Zirngibl was used to enrich the aforementioned clusters from the whole dataset of differentially regulated genes based on gene expression trajectories.

4.8.4 Gene enrichment analysis

Gene enrichment analysis was performed on the 1796 identified targets from the 5 manually curated clusters using Cytoscape software by Matt Rogon. Analysis of individual and multi-cluster gene ontology analysis yielded a wealth of interesting pathways and gene families that were differentially regulated. The following Gene Ontology terms were manually picked for further network mapping and druggable target analysis on Cytoscape:

I. GO:0002819 Regulation of adaptive immune response

II. GO:0098742 Cell-cell adhesion via plasma-membrane adhesion molecules

III. GO:0002009 Morphogenesis of an epithelium

IV. GO:0001213 Formation of the cornified envelope

V. GO:0031663 Lipopolysaccharide-mediated signaling pathway

VI. GO:0017017 MAP kinase tyrosine/serine/threonine phosphatase activity

VII. GO:0000097 Downregulation of ERBB4 signaling

VIII. GO:0000618 Oxidative Stress Induced Senescence

IX. GO:0008593 Regulation of Notch signaling pathway

X. GO:0000571 FAS pathway and Stress induction of HSP regulation

4.9 Microarray data analysis

4.9.1 RNA extraction

TetO-MYC/ TetO-Neu/ MMTV-rtTA adult female mice induced to have tumors were organ-harvested during regression phase after removal of doxycycline from the animal diet. During harvest 20-30 mg tissue pieces were snap frozen in liquid nitrogen for RNA extraction. The snap frozen tissue pieces were ground in a mortar and pestel (maintained at -50°C by placing it in a trough full of dry ice). The frozen tissue powder was then transferred to an Eppendorf tube with 900μl of mirVana lysis buffer. Total RNA was subsequently extracted using mirVana miRNA Isolation Kit (Ambion, Cat. # AM1560). RNA quality and concentration were assessed using the Bioanalyzer (Agilent 2100, Cat. # G2939BA) at the Genomics Core facility, EMBL, Heidelberg. RNA extracted from the mammary gland tissues at hallmark time points during regression were were run on Affymetrix MoGene-2 microarray chips.

4.9.2 Microarray data analysis

Microarray data files was analysed using the Transcriptome Analysis Console (TAC) 4.0 software. Principal Component Analysis (PCA) plotting and differentially regulated pathways compared to the control samples were curated manually. Manual overlay of gene enrichment analysis for both *in vivo* and *d*atasets yielded the following interesting subgroup of pathways:
(a) Focal Adhesion PI3K- AKT-mTor Sinalling Pathway
(b) MAPK Signalling (specifically: c-Jun Kinase/Stress-activated Pathway)
(c) TNF-alpha and NfKb signalling pathway
(d) Tgf-Beta Signalling pathway
(e) DNA damage response (ATM dependent)
(f) Apoptosis pathways.
Key molecular players from these pathways were selected for validation, including c-Jun, BCl-6, CCL5 and c-Rel.

CHAPTER 5

SUPPLEMENTARY MATERIAL

Organoid	Movie Code	Tumor outcome
Organoid-1	0308S1	Yes
Organoid-2	0308S2	Yes
Organoid-3	0308S3	No
Organoid-4	0308S4	Yes
Organoid-5	0308S8	Yes
Organoid-6	2607S0	Yes
Organoid-7	2607S2	No
Organoid-8	2607S6	No
Organoid-9	2607S11	Yes
Organoid-10	3107S0	No
Organoid-11	3107S1	No
Organoid-12	3107S2	Yes
Organoid-13	3107S3	No
Organoid-14	31074	No
Organoid-15	3107S5	No
Organoid-16	2111S0	Yes
Organoid-17	2111S2	No
Organoid-18	2111S4	Yes
Organoid-19	2111S6	Yes
Organoid-20	2111S8	Yes

Table 5.1: List of transduced organoids analyzed for feature analysis.

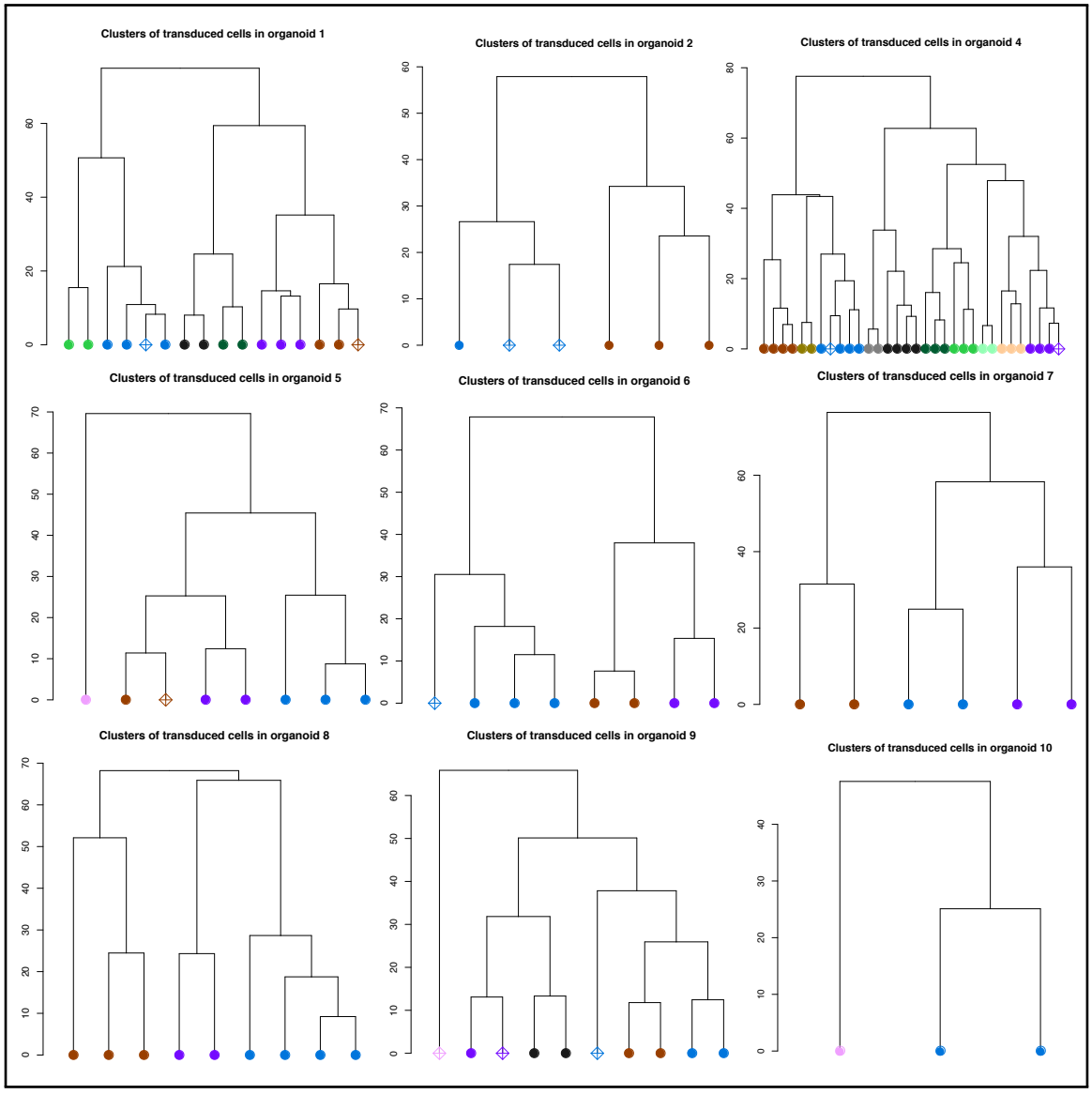


Figure 5.1: Hierarchical clustering of transduced cells in bi-transgenic organoids undergoing stochastic tumourigenesis. Organoids are name coded according to Table 5.1.

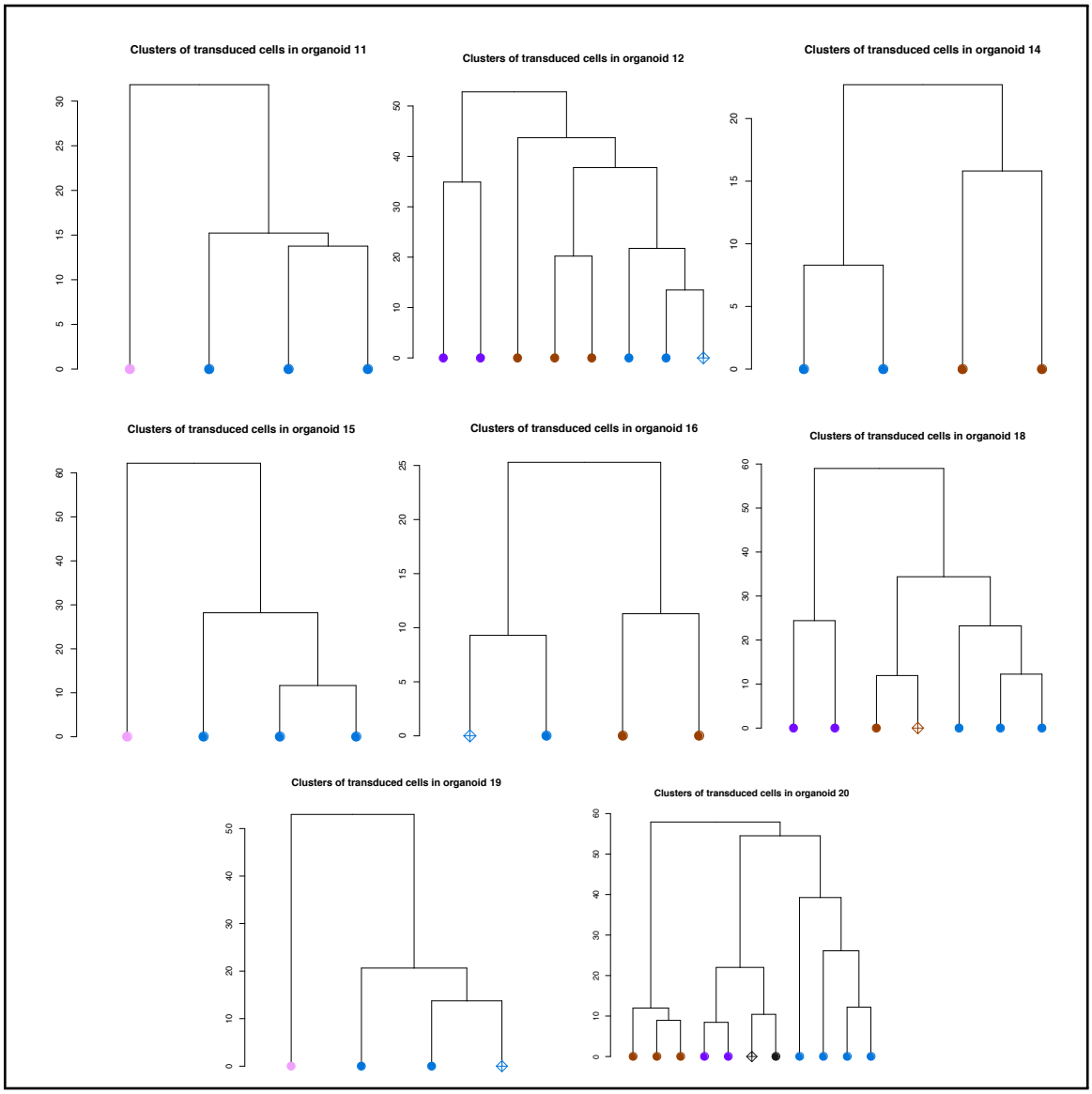


Figure 5.2: Hierarchical clustering of transduced cells in bi-transgenic organoids undergoing stochastic tumourigenesis. Organoids are name coded according to Table 5.1.

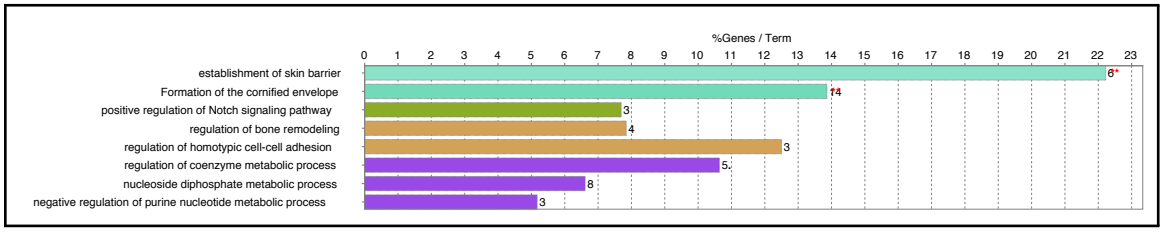


Figure 5.3: Gene Ontology Enrichment analysis of genes with 'Target' trajectories.

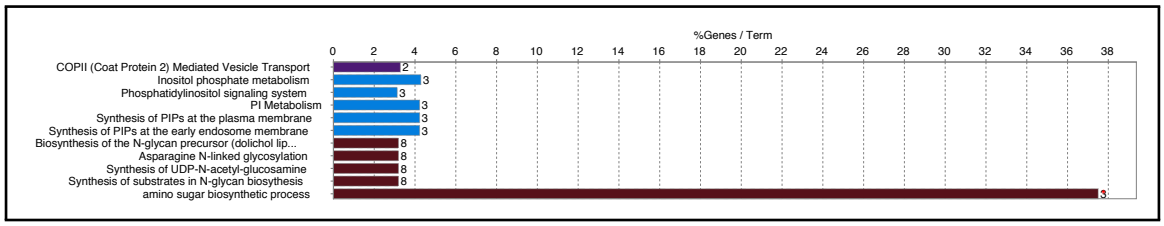


Figure 5.4: Gene Ontology Enrichment analysis of genes with 'Early Responder' trajectories.

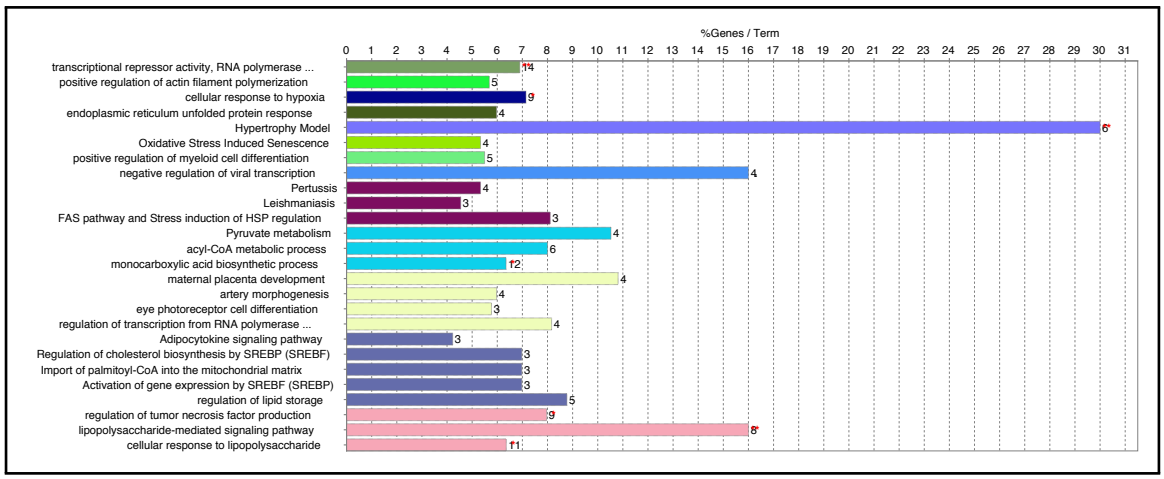


Figure 5.5: Gene Ontology Enrichment analysis of genes with 'Intermediate Responder' trajectories.

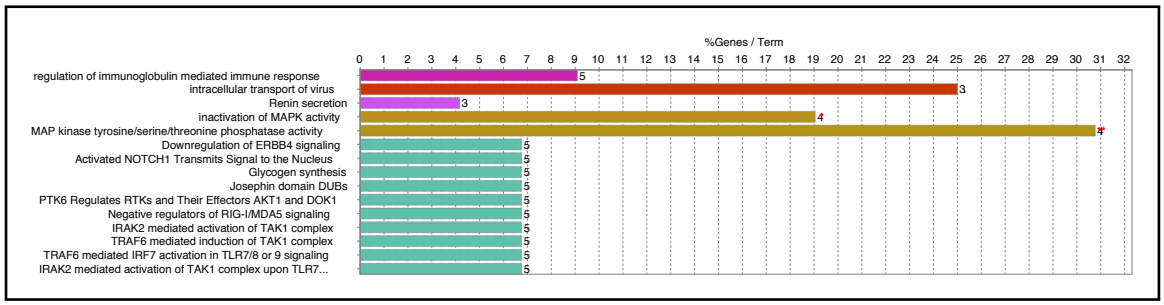


Figure 5.6: Gene Ontology Enrichment analysis of genes with 'Late Responder' trajectories.

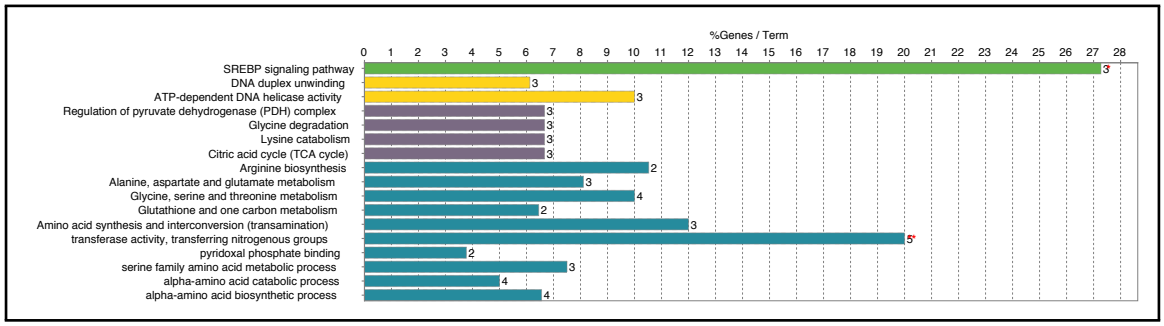


Figure 5.7: Gene Ontology Enrichment analysis of genes with 'Shoulder' trajectories.

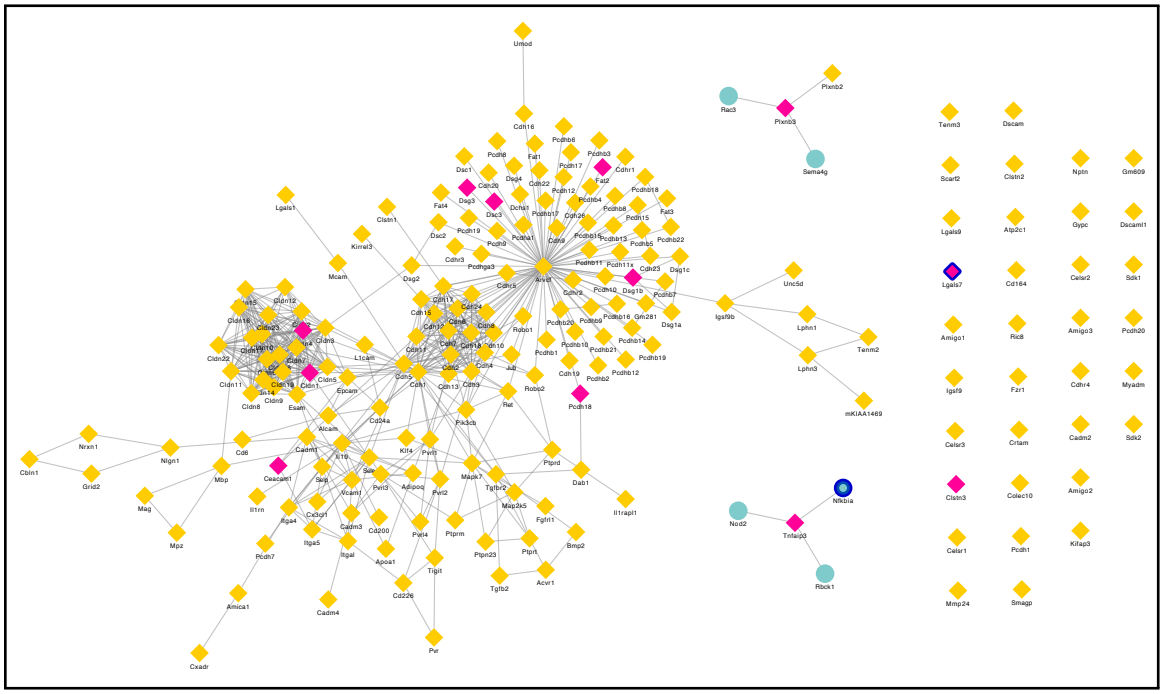


Figure 5.8: Gene network for the 'Cell-cell adhesion via plasma-membrane adhesion molecules' GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

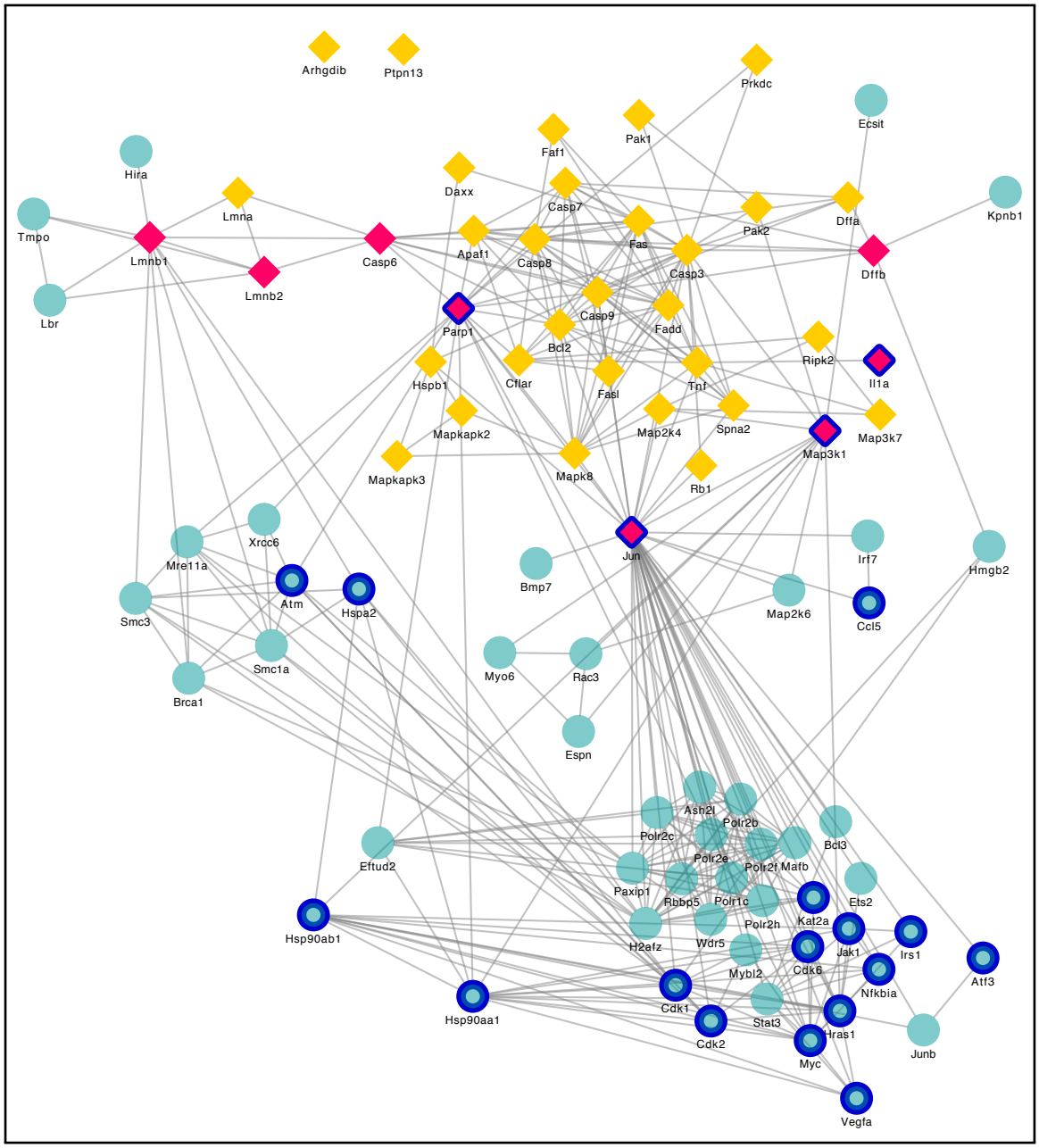


Figure 5.9: Gene network for the 'FAS pathway and Stress induction of HSP regulation' GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

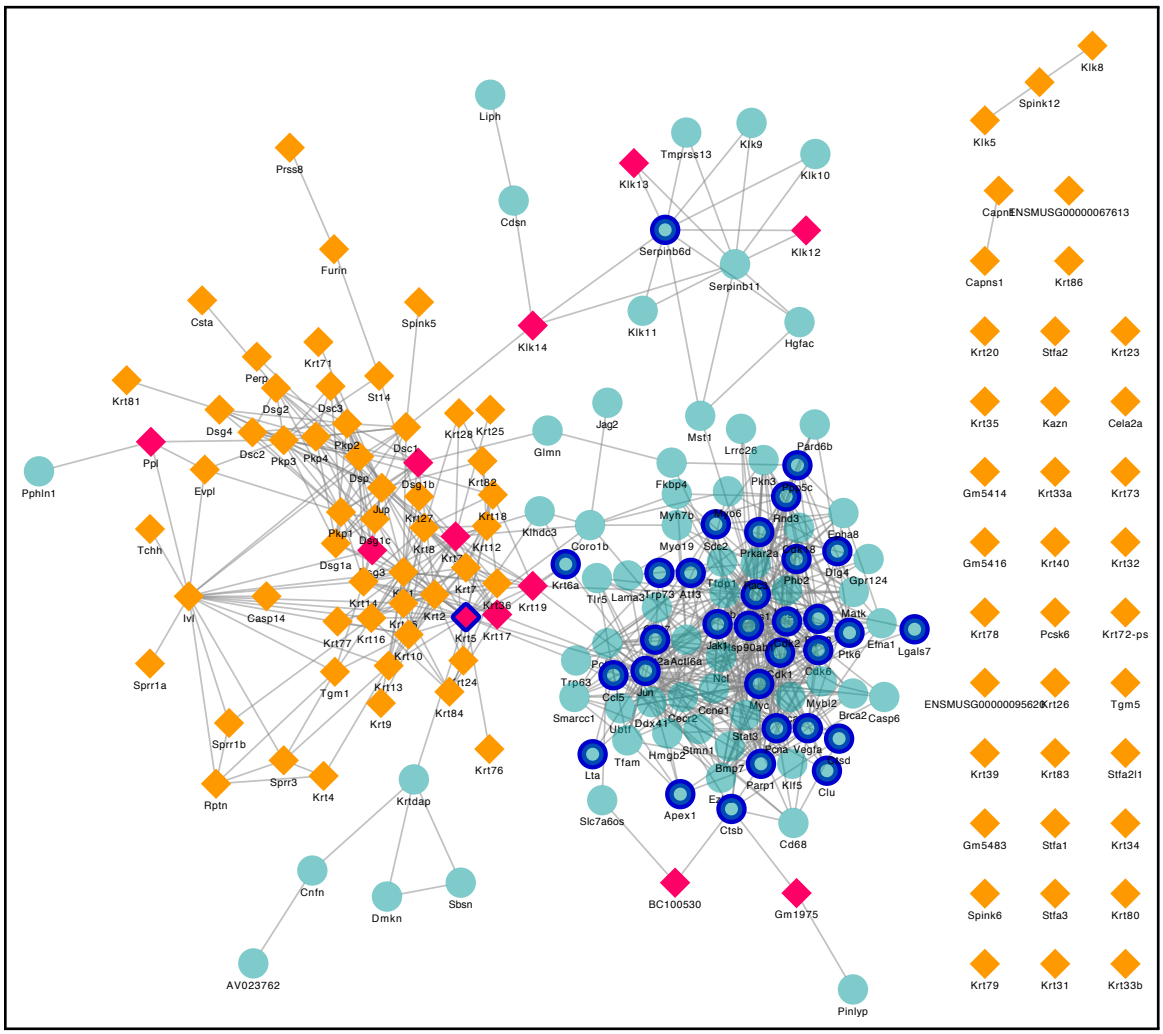


Figure 5.10: Gene network for the 'Formation of the cornified envelope'GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

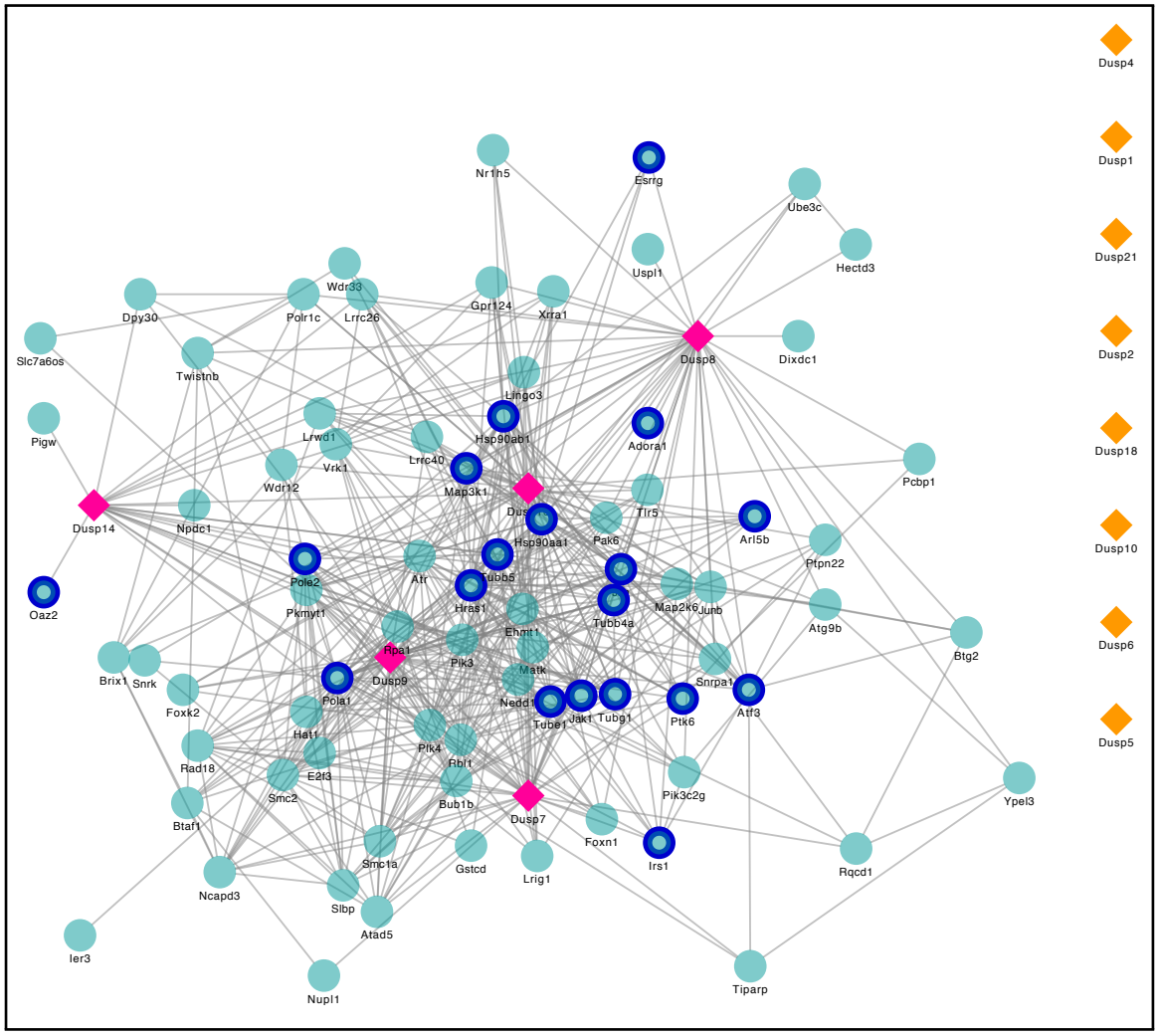


Figure 5.11: Gene network for the 'MAP kinase tyrosine/serine/threonine phosphatase activity' GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

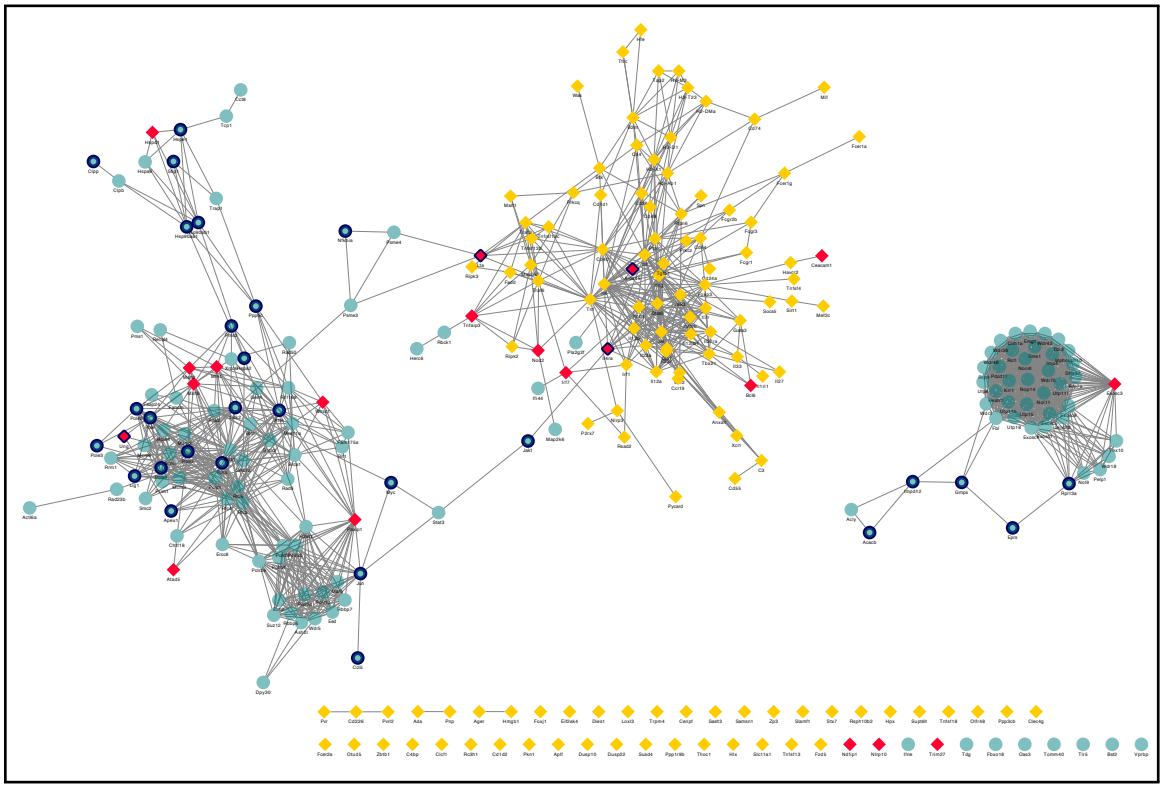


Figure 5.12: Gene network for the 'Regulation of adaptive immune response' GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

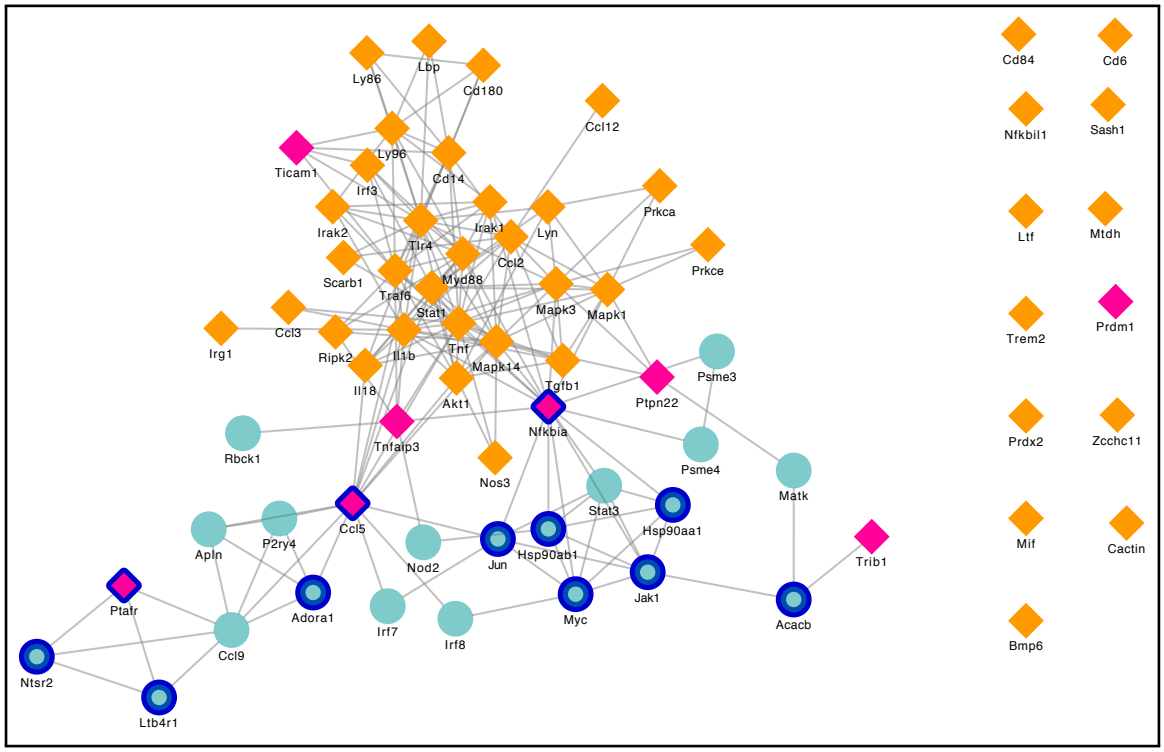


Figure 5.13: Gene network for the 'Lipopolysaccharide-mediated signaling pathway' GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

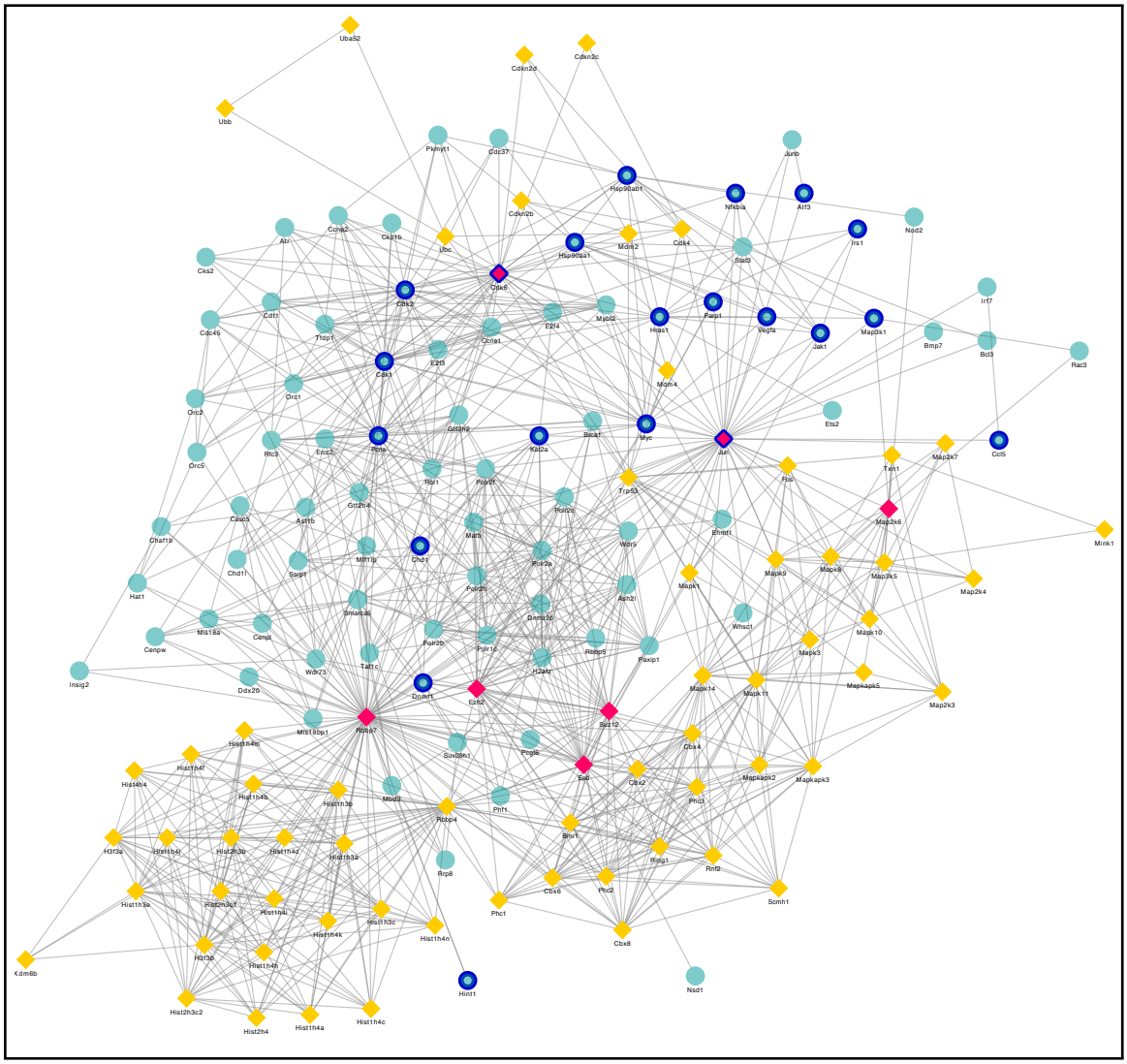


Figure 5.14: Gene network for the 'Oxidative Stress Induced Senescence'GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

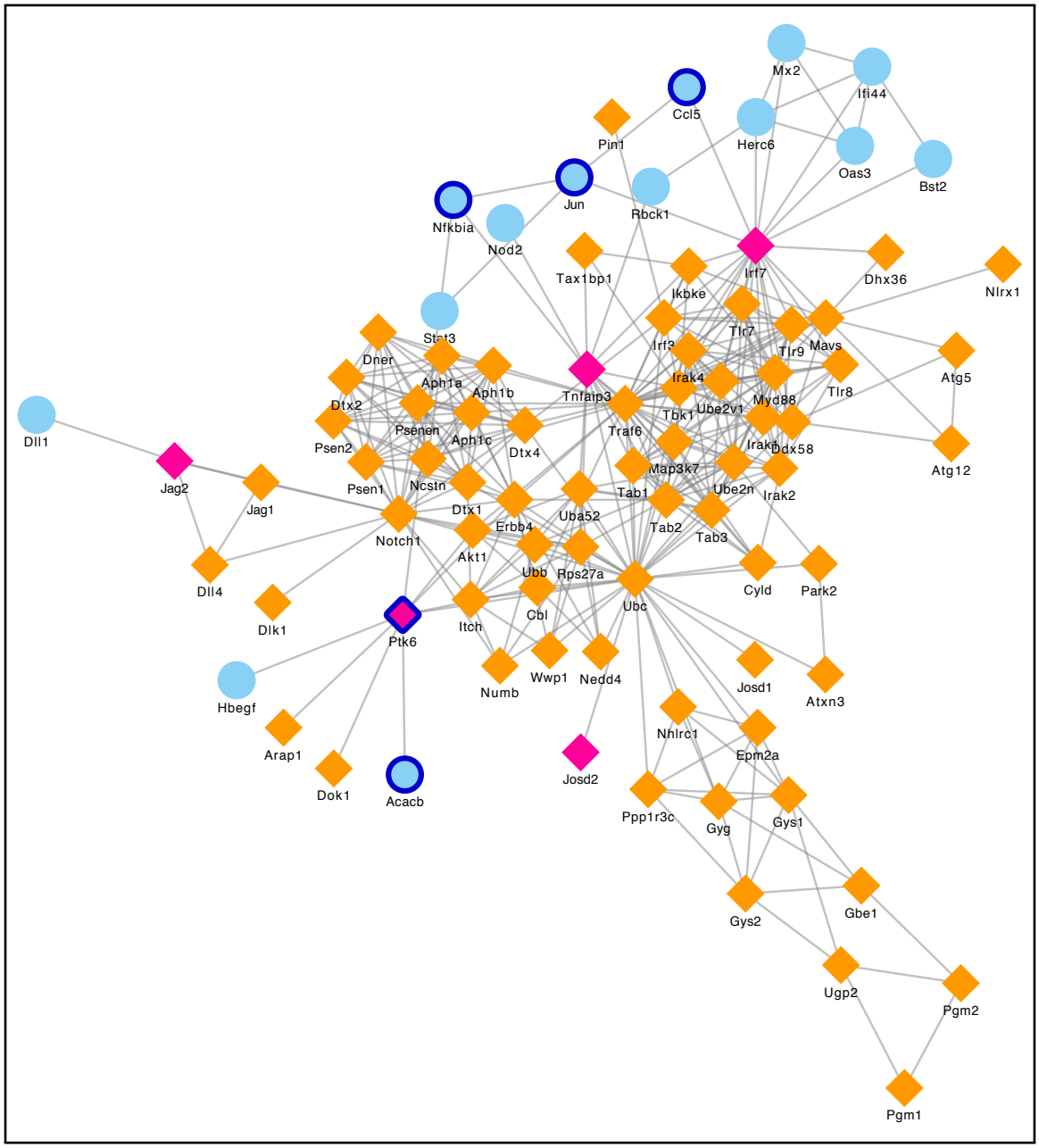


Figure 5.15: Gene network for 'Downregulation of ERBB4 signaling' GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

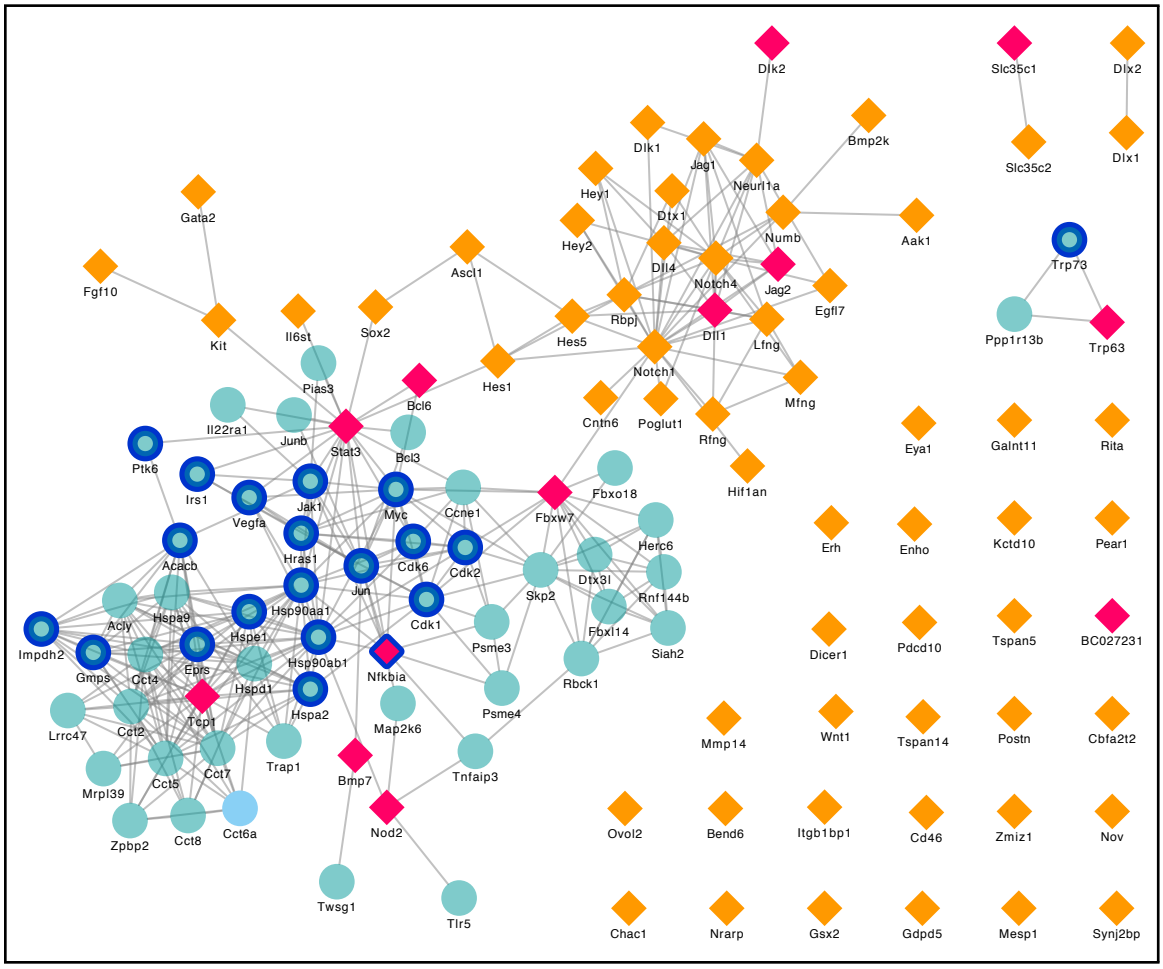


Figure 5.16: Gene network for the 'Regulation of Notch signaling pathway' GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

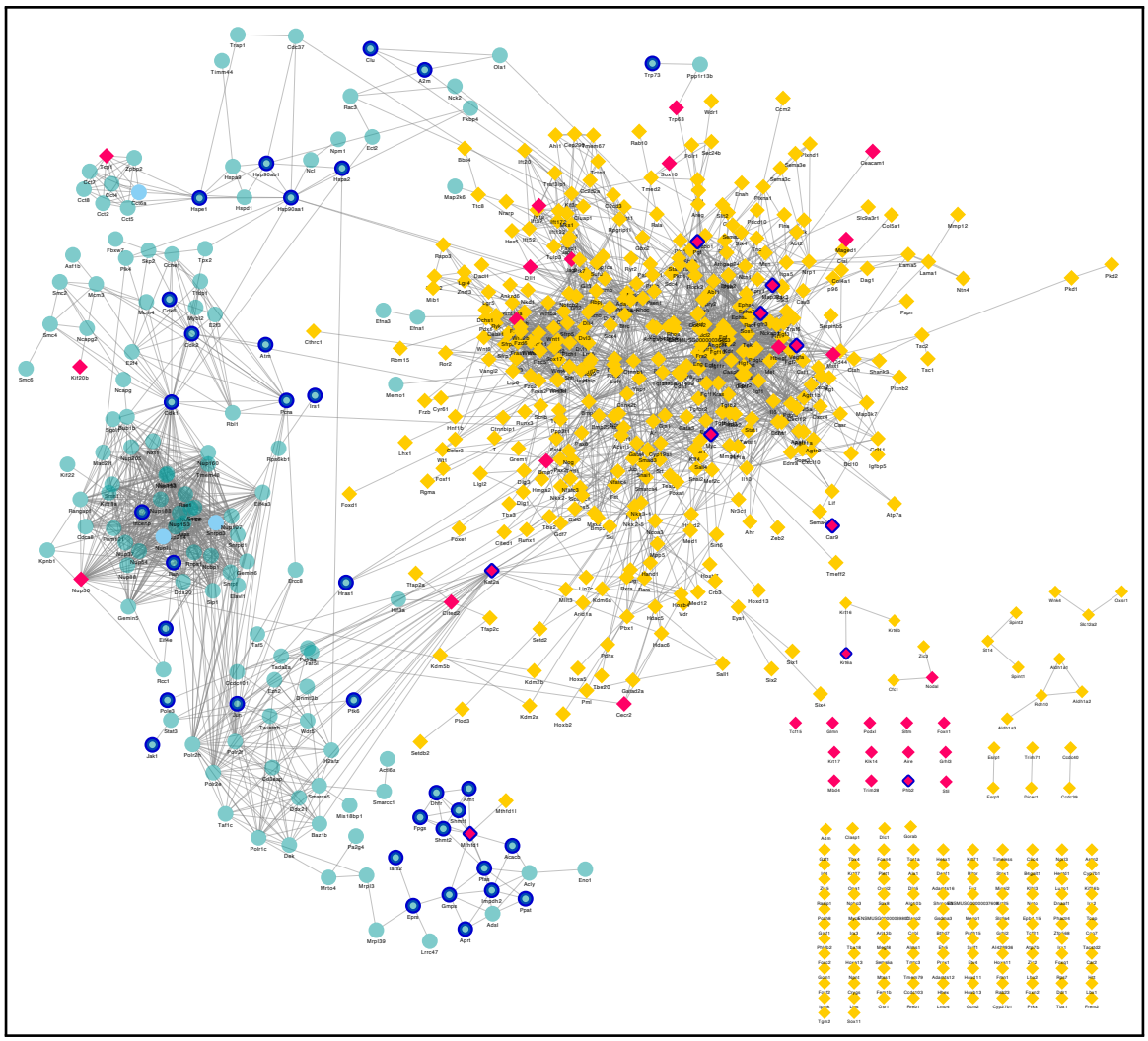


Figure 5.17: Gene network for the 'Morphogenesis of an epithelium' GO Term. Yellow diamonds represent all genes in the network. Blue circles represent the first neighbours of all genes based on protein-protein interactions. Pink diamonds represent the deferentially regulated genes in the network during the process of tumor regression in the organoid cultures. Dark blue borders on genes indicate that the gene is druggable with known inhibitors.

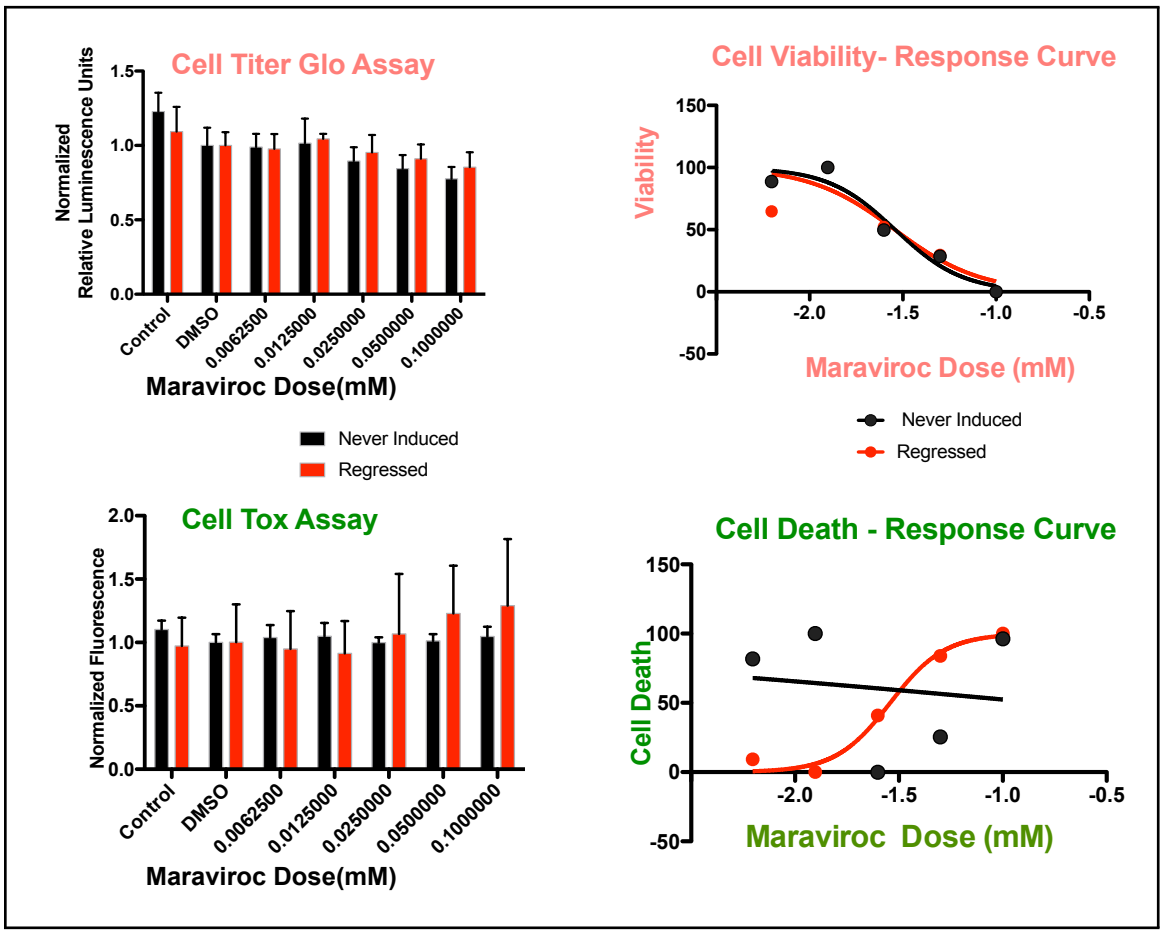


Figure 5.18: Cell Viability and Cell Death readouts from synthetic lethality testing of the CXCR5 antagonist Maraviroc. Upper left graph shows the readout from the CellTiter-Glo® Luminescent Assay. Cell viability readouts for the 5 doses of Maraviroc tested and normalized against the control are shown. Derived from this graph is the upper right graph depicting the Cell Viability response curve. IC_{50} values for Maraviroc in the Cell Viability response curve are 0.0287 mM for Never Induced and 0.0286 mM for Regressed cultures. Lower left graph shows the readout from the CellTox^{TM} Green Fluorescence Assay. Cell death readouts for the 5 doses of Maraviroc tested and normalized against the control are shown. Derived from this graph is the lower right graph depicting the Cell Death response curve. IC_{50} values for Maraviroc in the Cell Death response curve are 0.1 mM for Never Induced and 0.028 mM for Regressed cultures.

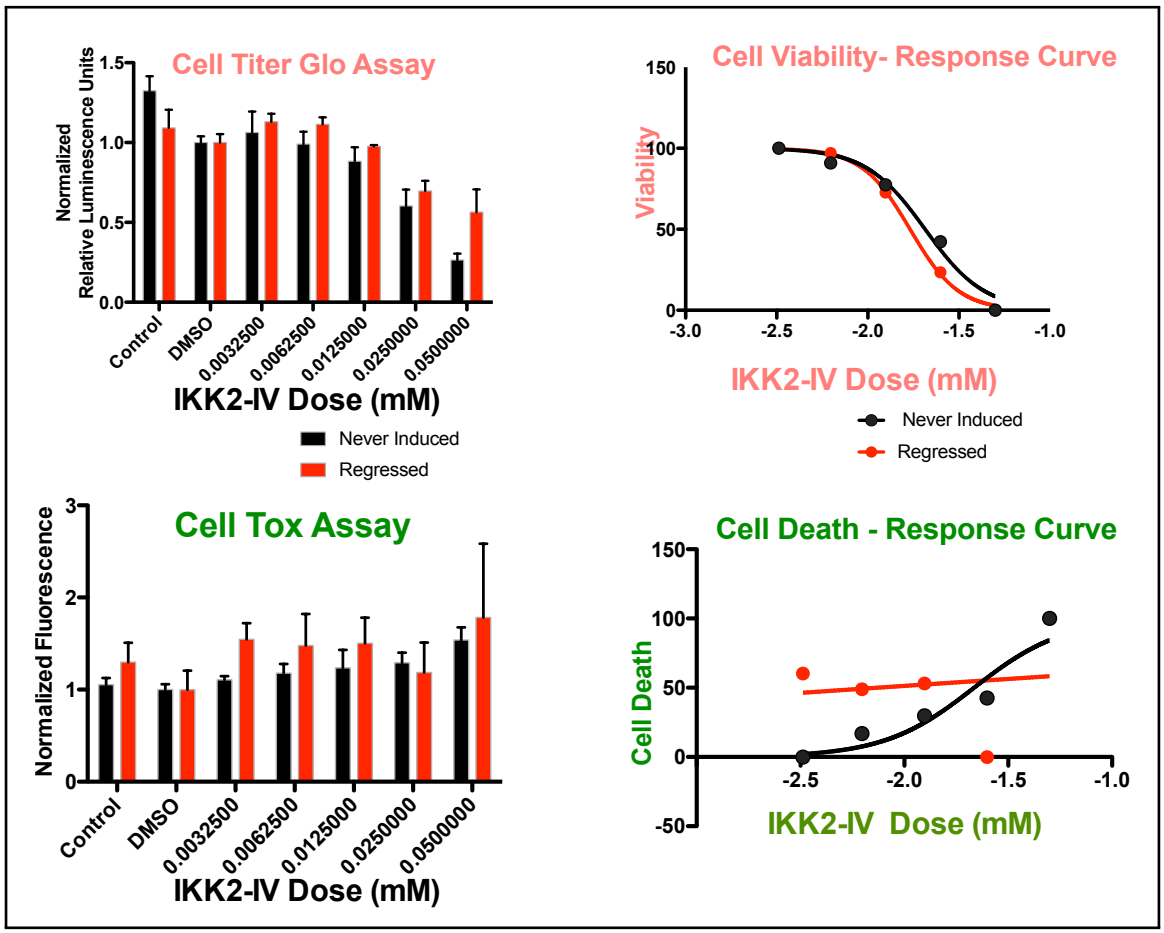


Figure 5.19: Cell Viability and Cell Death readouts from synthetic lethality testing of the IKK2 Inhibitor IV. Upper left graph shows the readout from the CellTiter-Glo® Luminescent Assay. Cell viability readouts for the 5 doses of IKK2 Inhibitor IV tested and normalized against the control are shown. Derived from this graph is the upper right graph depicting the Cell Viability response curve. IC_{50} values for IKK2 Inhibitor IV in the Cell Viability response curve are 0.02 mM for Never Induced and 0.017 mM for Regressed cultures. Lower left graph shows the readout from the CellTox^{TM} Green Fluorescence Assay. Cell death readouts for the 5 doses of IKK2 Inhibitor IV tested and normalized against the control are shown. Derived from this graph is the lower right graph depicting the Cell Death response curve. IC_{50} values for IKK2 Inhibitor IV in the Cell Death response curve are 0.02 mM for Never Induced and 0.00073 mM for Regressed cultures.

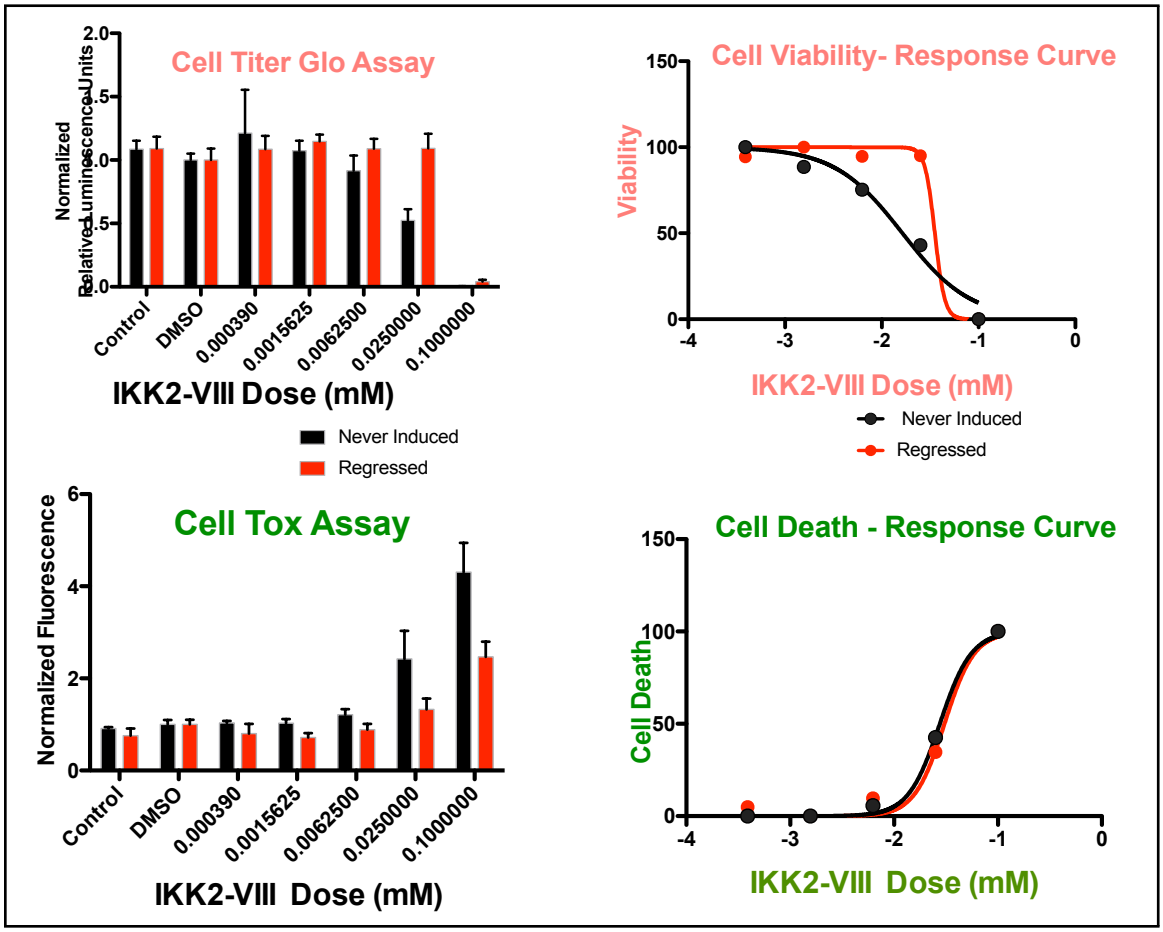


Figure 5.20: Cell Viability and Cell Death readouts from synthetic lethality testing of the IKK2 Inhibitor VIII. Upper left graph shows the readout from the CellTiter-Glo® Luminescent Assay. Cell viability readouts for the 5 doses of IKK2 Inhibitor VIII tested and normalized against the control are shown. Derived from this graph is the upper right graph depicting the Cell Viability response curve. IC_{50} values for IKK2 Inhibitor VIII in the Cell Viability response curve are 0.016 mM for Never Induced and 0.035 mM for Regressed cultures. Lower left graph shows the readout from the CellTox^{TM} Green Fluorescence Assay. Cell death readouts for the 5 doses of IKK2 Inhibitor VIII tested and normalized against the control are shown. Derived from this graph is the lower right graph depicting the Cell Death response curve. IC_{50} values for IKK2 Inhibitor VIII in the Cell Death response curve are 0.027 mM for Never Induced and 0.03055 mM for Regressed cultures.

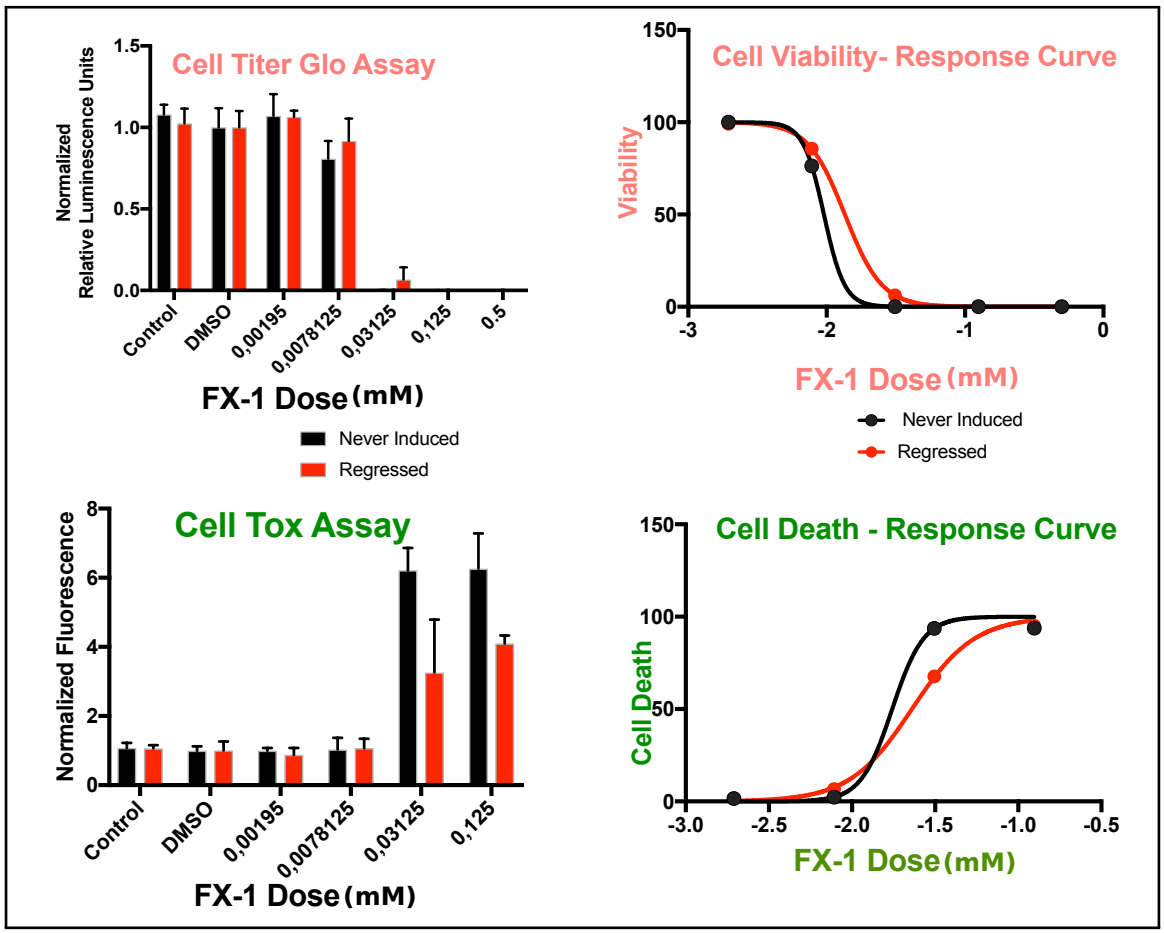


Figure 5.21: Cell Viability and Cell Death readouts from synthetic lethality testing of FX-1. Upper left graph shows the readout from the CellTiter-Glo® Luminescent Assay. Cell viability readouts for the 5 doses of FX-1 tested and normalized against the control are shown. Derived from this graph is the upper right graph depicting the Cell Viability response curve. IC_{50} values for FX-1 in the Cell Viability response curve are 0.009 mM for Never Induced and 0.013 mM for Regressed cultures. Lower left graph shows the readout from the CellTox^{TM} Green Fluorescence Assay. Cell death readouts for the 5 doses of FX-1 tested and normalized against the control are shown. Derived from this graph is the lower right graph depicting the Cell Death response curve. IC_{50} values for FX-1 in the Cell Death response curve are 0.0174 mM for Never Induced and 0.022 mM for Regressed cultures.

CHAPTER 6

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] MD Melissa Conrad Stöppler. *Anatomical features of the breast.* URL: https://www.medicinenet.com/breast_anatomy/article.htm#what_should_i_know_about_breast.

[2] *Medical Illustrations by Patrick Lynch, generated for multimedia teaching projects by the Yale University School of Medicine, Center for Advanced Instructional Media, 1987-2000.* URL: http://patricklynch.net.

[3] Cleaveland Clinic. *Breast Anatomy.* URL: https://my.clevelandclinic.org/health/articles/8330-breast-anatomy.

[4] Jonine D Figueroa et al. "Terminal duct lobular unit involution of the normal breast: implications for breast cancer etiology". In: *JNCI: Journal of the National Cancer Institute* 106.10 (2014).

[5] Thorarinn Gudjonsson et al. "Myoepithelial cells: their origin and function in breast morphogenesis and neoplasia". In: *Journal of mammary gland biology and neoplasia* 10.3 (2005), pp. 261–272.

[6] Kandice Tanner. "Regulation of the basement membrane by epithelia generated forces". In: *Physical biology* 9.6 (2012), p. 065003.

[7] David L Page et al. "Lobular neoplasia of the breast: higher risk for subsequent invasive cancer predicted by more extensive disease". In: *Human pathology* 22.12 (1991), pp. 1232–1239.

[8] Inc Genomic Health. *Diagnosing Breast Cancer.* URL: https://www.mybreastcancertreatment.org/en-US/PersonalizeYourTreatment/YourDiagnosis.

[9] *Cancer Research UK.* URL: https://www.cancerresearchuk.org/about-cancer/breast-cancer/stages-types-grades/types/lobular-carcinoma-in-situ-lcis.

[10] A Spitale et al. “Breast cancer classification according to immunohistochemical markers: clinicopathologic features and short-term survival analysis in a population-based study from the South of Switzerland”. In: *Annals of oncology* 20.4 (2008), pp. 628–635.

[11] World Health Organization (WHO) - International Agency for Research on Cancer. *Cancer Fact Sheets - Breast, GLOBOCAN 2018*. URL: https://gco.iarc.fr/today/fact-sheets-cancers.

[12] Xiaofeng Dai et al. “Breast cancer intrinsic subtype classification, clinical use and future trends”. In: *American journal of cancer research* 5.10 (2015), p. 2929.

[13] Britta Weigelt, Frederick L Baehner, and Jorge S Reis-Filho. “The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade”. In: *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland* 220.2 (2010), pp. 263–280.

[14] Takayuki Iwamoto and Lajos Pusztai. “Predicting prognosis of breast cancer with gene signatures: are we lost in a sea of data?” In: *Genome medicine* 2.11 (2010), p. 81.

[15] Maggie CU Cheang et al. “Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer”. In: *JNCI: Journal of the National Cancer Institute* 101.10 (2009), pp. 736–750.

[16] Carlos S Vallejos et al. “Breast cancer classification according to immunohistochemistry markers: subtypes and association with clinicopathologic variables in a peruvian hospital database”. In: *Clinical breast cancer* 10.4 (2010), pp. 294–300.

[17] Jorge S Reis-Filho et al. “Molecular profiling: moving away from tumor philately”. In: *Science translational medicine* 2.47 (2010), 47ps43–47ps43.

[18] Christos Sotiriou and Lajos Pusztai. “Gene-expression signatures in breast cancer”. In: *New England Journal of Medicine* 360.8 (2009), pp. 790–800.

[19] Therese Sørlie et al. “Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications”. In: *Proceedings of the National Academy of Sciences* 98.19 (2001), pp. 10869–10874.

[20] Marcel Smid et al. “Subtypes of breast cancer show preferential site of relapse”. In: *Cancer research* 68.9 (2008), pp. 3108–3114.

[21] Charles M Perou et al. “Molecular portraits of human breast tumours”. In: *nature* 406.6797 (2000), p. 747.

[22] Therese Sorlie et al. “Repeated observation of breast tumor subtypes in independent gene expression data sets.” In: *Proceedings of the National Academy of Sciences of the United States of America* 100.14 (2003), pp. 8418–8423.

[23] Christos Sotiriou et al. “Breast cancer classification and prognosis based on gene expression profiles from a population-based study”. In: *Proceedings of the National Academy of Sciences* 100.18 (2003), pp. 10393–10398.

[24] Dalia M Abd El-Rehim et al. “High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses”. In: *International journal of cancer* 116.3 (2005), pp. 340–350.

[25] Katie M O’Brien et al. “Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study”. In: *Clinical Cancer Research* 16.24 (2010), pp. 6100–6110.

[26] Xiaofeng Dai, Ana Chen, and Zhonghu Bai. “Integrative investigation on breast cancer in ER, PR and HER2-defined subgroups using mRNA and miRNA expression profiling”. In: *Scientific reports* 4 (2014), p. 6566.

[27] Mark M Moasser. “The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis”. In: *Oncogene* 26.45 (2007), p. 6469.

[28] Dennis J Slamon et al. “Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer”. In: *Science* 244.4905 (1989), pp. 707–712.

[29] I Bernard Weinstein and Andrew Joe. “Oncogene addiction”. In: *Cancer research* 68.9 (2008), pp. 3077–3080.

[30] Jeffrey Settleman. “Oncogene addiction”. In: *Current biology* 22.2 (2012), R43–R44.

[31] JC Singh, K Jhaveri, and FJ Esteva. “HER2-positive advanced breast cancer: optimizing patient outcomes and opportunities for drug development”. In: *British journal of cancer* 111.10 (2014), p. 1888.

[32] Diogo Mendes et al. “The benefit of HER2-targeted therapies on overall survival of patients with metastatic HER2-positive breast cancer–a systematic review”. In: *Breast Cancer Research* 17.1 (2015), p. 140.

[33] Yoichi Nagata et al. “PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients”. In: *Cancer cell* 6.2 (2004), pp. 117–127.

[34] James D Brenton et al. “Molecular classification and molecular forecasting of breast cancer: ready for clinical application?” In: *Journal of clinical oncology* 23.29 (2005), pp. 7350–7360.

[35] Trever G Bivona and Robert C Doebele. “A framework for understanding and targeting residual disease in oncogene-driven solid cancers”. In: *Nature medicine* 22.5 (2016), p. 472.

[36] Bernhard Polzer and Christoph A Klein. "Metastasis awakening: the challenges of targeting minimal residual cancer". In: *Nature medicine* 19.3 (2013), p. 274.

[37] Kelly Kersten et al. "Genetically engineered mouse models in oncology research and cancer medicine". In: *EMBO molecular medicine* 9.2 (2017), pp. 137–153.

[38] Katerina Politi and William Pao. "How genetically engineered mouse tumor models provide insights into human cancers". In: *Journal of Clinical Oncology* 29.16 (2011), p. 2273.

[39] Stefanie Urlinger et al. "Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity". In: *Proceedings of the National Academy of Sciences* 97.14 (2000), pp. 7963–7968.

[40] Ashna Alladin and Martin Jechlinger. "Towards a holistic and mechanistic understanding of tumourigenesis via genetically engineered mouse models". In: *Current Opinion in Systems Biology* 6 (2017), pp. 74–79.

[41] MH Barcellos-Hoff et al. "Functional differentiation and alveolar morphogenesis of primary mammary cultures on reconstituted basement membrane". In: *Development* 105.2 (1989), pp. 223–235.

[42] Karen L Schmeichel and Mina J Bissell. "Modeling tissue-specific signaling and organ function in three dimensions". In: *Journal of cell science* 116.12 (2003), pp. 2377–2388.

[43] Andrew D Rhim, Martin Jechlinger, and Anil K Rustgi. "Culturing mouse tumor cells". In: *Cold Spring Harbor Protocols* 2015.6 (2015), pdb–top069989.

[44] Madeline A Lancaster and Juergen A Knoblich. "Organogenesis in a dish: modeling development and disease using organoid technologies". In: *Science* 345.6194 (2014), p. 1247125.

[45] Lincoln D Nadauld et al. "Metastatic tumor evolution and organoid modeling implicate TGFBR2 as a cancer driver in diffuse gastric cancer". In: *Genome biology* 15.8 (2014), p. 428.

[46] Xingnan Li et al. "Oncogenic transformation of diverse gastrointestinal tissues in primary organoid culture". In: *Nature medicine* 20.7 (2014), p. 769.

[47] Celina M D'Cruz et al. "c-MYC induces mammary tumorigenesis by means of a preferred pathway involving spontaneous Kras2 mutations". In: *Nature medicine* 7.2 (2001), p. 235.

[48] Susan E Moody et al. "Conditional activation of Neu in the mammary epithelium of transgenic mice results in reversible pulmonary metastasis". In: *Cancer cell* 2.6 (2002), pp. 451–461.

[49] Carlos L Arteaga et al. "Treatment of HER2-positive breast cancer: current status and future perspectives". In: *Nature reviews Clinical oncology* 9.1 (2012), p. 16.

[50] Jinhua Xu, Yinghua Chen, and Olufunmilayo I Olopade. “MYC and breast cancer”. In: *Genes & cancer* 1.6 (2010), pp. 629–640.

[51] R Nair et al. “c-Myc and Her2 cooperate to drive a stem-like phenotype with poor prognosis in breast cancer”. In: *Oncogene* 33.30 (2014), p. 3992.

[52] Kristina M Havas et al. “Metabolic shifts in residual breast cancer drive tumor recurrence”. In: *The Journal of clinical investigation* 127.6 (2017), pp. 2091–2105.

[53] Nowell and Peter C. “Tumor progression: a brief historical perspective, Seminars in Cancer Biology”. In: *DeepDyve* (Aug. 2002).

[54] Doris P Tabassum and Kornelia Polyak. “Tumorigenesis: it takes a village”. In: *Nature Reviews Cancer* 15.8 (2015), p. 473.

[55] Savas D Soysal, Alexandar Tzankov, and Simone E Muenst. “Role of the tumor microenvironment in breast cancer”. In: *Pathobiology* 82.3-4 (2015), pp. 142–152.

[56] Maria Aparecida Azevedo Koike Folgueira et al. “Markers of breast cancer stromal fibroblasts in the primary tumour site associated with lymph node metastasis: a systematic review including our case series”. In: *Bioscience reports* 33.6 (2013), e00085.

[57] Shinichi Tsutsui et al. “Macrophage infiltration and its prognostic implications in breast cancer: the relationship with VEGF expression and microvessel density”. In: *Oncology reports* 14.2 (2005), pp. 425–431.

[58] Sandra D Bohling and Kimberly H Allison. “Immunosuppressive regulatory T cells are associated with aggressive breast cancer phenotypes: a potential therapeutic target”. In: *Modern Pathology* 21.12 (2008), p. 1527.

[59] Pengfei Lu, Valerie M Weaver, and Zena Werb. “The extracellular matrix: a dynamic niche in cancer progression”. In: *J Cell Biol* 196.4 (2012), pp. 395–406.

[60] Minhui Lee and Valeri Vasioukhin. “Cell polarity and cancer–cell and tissue polarity as a non-canonical tumor suppressor”. In: *Journal of cell science* 121.8 (2008), pp. 1141–1150.

[61] David Bilder. “Epithelial polarity and proliferation control: links from the Drosophila neoplastic tumor suppressors”. In: *Genes & development* 18.16 (2004), pp. 1909–1925.

[62] S Kuphal et al. “Expression of Hugl-1 is strongly reduced in malignant melanoma”. In: *Oncogene* 25.1 (2006), p. 103.

[63] Paola Massimi et al. “HPV E6 specifically targets different cellular pools of its PDZ domain-containing tumour suppressor substrates for proteasome-mediated degradation”. In: *Oncogene* 23.49 (2004), p. 8033.

[64] S Nakagawa et al. “Analysis of the expression and localisation of a LAP protein, human scribble, in the normal and neoplastic epithelium of uterine cervix”. In: *British Journal of Cancer* 90.1 (2004), p. 194.

[65] Daniela Gardiol et al. “Human discs large and scrib are localized at the same regions in colon mucosa and changes in their expression patterns are correlated with loss of tissue architecture during malignant progression”. In: *International journal of cancer* 119.6 (2006), pp. 1285–1290.

[66] Carl C Schimanski et al. “Reduced expression of Hugl-1, the human homologue of Drosophila tumour suppressor gene lgl, contributes to progression of colorectal cancer”. In: *Oncogene* 24.19 (2005), p. 3100.

[67] Ana Laura Cavatorta et al. “Differential expression of the human homologue of drosophila discs large oncosuppressor in histologic samples from human papillomavirus–associated lesions as a marker for progression to malignancy”. In: *International journal of cancer* 111.3 (2004), pp. 373–380.

[68] P Massimi et al. “HPV E6 degradation of p53 and PDZ containing substrates in an E6AP null background”. In: *Oncogene* 27.12 (2008), p. 1800.

[69] Miranda Thomas et al. “HPV E6 and MAGUK protein interactions: determination of the molecular basis for specific protein recognition and degradation”. In: *Oncogene* 20.39 (2001), p. 5431.

[70] Huafeng Zhang et al. “Mitochondrial autophagy is an HIF-1-dependent adaptive metabolic response to hypoxia”. In: *Journal of Biological Chemistry* 283.16 (2008), pp. 10892–10903.

[71] Nancy E Hynes and Heidi A Lane. “ERBB receptors and cancer: the complexity of targeted inhibitors”. In: *Nature Reviews Cancer* 5.5 (2005), p. 341.

[72] Bryan Linggi and Graham Carpenter. “ErbB receptors: new insights on mechanisms and biology”. In: *Trends in cell biology* 16.12 (2006), pp. 649–656.

[73] Victoria Aranda et al. “Par6–aPKC uncouples ErbB2 induced disruption of polarized epithelial organization from proliferation control”. In: *Nature cell biology* 8.11 (2006), p. 1235.

[74] WG Kaelin. “The von Hippel-Lindau tumor suppressor protein: roles in cancer and oxygen sensing”. In: *Cold Spring Harbor Symposia on Quantitative Biology*. Vol. 70. Cold Spring Harbor Laboratory Press. 2005, pp. 159–166.

[75] William G Kaelin. “The von Hippel-Lindau tumor suppressor protein and clear cell renal carcinoma”. In: *Clinical Cancer Research* 13.2 (2007), 680s–684s.

[76] Heiwa Okuda et al. “The von Hippel-Lindau tumor suppressor protein mediates ubiquitination of activated atypical protein kinase C”. In: *Journal of Biological Chemistry* 276.47 (2001), pp. 43611–43617.

[77] Bernhard Schermer et al. “The von Hippel-Lindau tumor suppressor protein controls ciliogenesis by orienting microtubule growth”. In: *J Cell Biol* 175.4 (2006), pp. 547–554.

[78] Derrick J Rossi and Irving L Weissman. “Pten, tumorigenesis, and stem cell self-renewal”. In: *Cell* 125.2 (2006), pp. 229–231.

[79] Fernando Martin-Belmonte et al. “PTEN-mediated apical segregation of phosphoinositides controls epithelial morphogenesis through Cdc42”. In: *Cell* 128.2 (2007), pp. 383–397.

[80] Hans Clevers. “Modeling development and disease with organoids”. In: *Cell* 165.7 (2016), pp. 1586–1597.

[81] Douglas Hanahan and Robert A Weinberg. “Hallmarks of cancer: the next generation”. In: *cell* 144.5 (2011), pp. 646–674.

[82] Jan Krieger. *Selective Plane Illumination Microscopy*.

[83] Carla S Verissimo et al. “Targeting mutant RAS in patient-derived colorectal cancer organoids by combinatorial drug screening”. In: *Elife* 5 (2016), e18489.

[84] Marie Held et al. “Ex vivo live cell tracking in kidney organoids using light sheet fluorescence microscopy”. In: *PloS one* 13.7 (2018), e0199918.

[85] Johanna F Dekkers et al. “Characterizing responses to CFTR-modulating drugs using rectal organoids derived from subjects with cystic fibrosis”. In: *Science Translational Medicine* 8.344 (2016), 344ra84–344ra84.

[86] Atze T Das, Liliane Tenenbaum, and Ben Berkhout. “Tet-On systems for doxycycline-inducible gene expression”. In: *Current gene therapy* 16.3 (2016), pp. 156–167.

[87] Celina M D’Cruz et al. “c-MYC induces mammary tumorigenesis by means of a preferred pathway involving spontaneous Kras2 mutations”. In: *Nature medicine* 7.2 (2001), p. 235.

[88] Johannes Schindelin et al. “Fiji: an open-source platform for biological-image analysis”. In: *Nature methods* 9.7 (2012), p. 676.

[89] *Imaris x64, v., Bitplane AG*. URL: http://bitplane.com.

[90] Christian Tischer and Rainer Pepperkok. *BigDataProcessor: Fiji plugin for big image data inspection and processing*. Feb. 2019. DOI: 10.5281/zenodo.2575681. URL: https://doi.org/10.5281/zenodo.2575681.

[91] Christian Tischer and Rainer Pepperkok. *CATS: Fiji plugin for context aware trainable segmentation for big image data*. Mar. 2019. DOI: 10.5281/zenodo.2600293. URL: https://doi.org/10.5281/zenodo.2600293.

[92] Peter Langfelder, Bin Zhang, and Steve Horvath. “Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R”. In: *Bioinformatics* 24.5 (2007), pp. 719–720.

[93] Vincent Calcagno, Claire de Mazancourt, et al. “glmulti: an R package for easy automated model selection with (generalized) linear models”. In: *Journal of statistical software* 34.12 (2010), pp. 1–29.

[94] Nobuyoshi Kosaka et al. “Competitive interactions of cancer cells and normal cells via secretory microRNAs”. In: *Journal of Biological Chemistry* 287.2 (2012), pp. 1397–1405.

[95] Mikhail Binnewies et al. “Understanding the tumor immune microenvironment (TIME) for effective therapy”. In: *Nature medicine* 24.5 (2018), p. 541.

[96] Yoichiro Tamori and Wu-Min Deng. “Cell competition and its implications for development and cancer”. In: *Journal of genetics and genomics* 38.10 (2011), pp. 483–495.

[97] Marina Simian and Mina J Bissell. “Organoids: a historical perspective of thinking in three dimensions”. In: *J Cell Biol* 216.1 (2017), pp. 31–40.

[98] Luke Martin McCaffrey et al. “Loss of the Par3 polarity protein promotes breast tumorigenesis and metastasis”. In: *Cancer cell* 22.5 (2012), pp. 601–614.

[99] Johanna I Partanen, Topi A Tervonen, and Juha Klefström. “Breaking the epithelial polarity barrier in cancer: the strange case of LKB1/PAR-4”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 368.1629 (2013), p. 20130111.

[100] Bin Xue et al. “Loss of Par3 promotes breast cancer metastasis by compromising cell–cell cohesion”. In: *Nature cell biology* 15.2 (2013), p. 189.

[101] Ian G Macara and Luke McCaffrey. “Cell polarity in morphogenesis and metastasis”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 368.1629 (2013), p. 20130012.

[102] Charles D Blanke et al. “Phase III randomized, intergroup trial assessing imatinib mesylate at two dose levels in patients with unresectable or metastatic gastrointestinal stromal tumors expressing the kit receptor tyrosine kinase: S0033”. In: *J Clin Oncol* 26.4 (2008), pp. 626–632.

[103] Makoto Maemondo et al. “Gefitinib or chemotherapy for non–small-cell lung cancer with mutated EGFR”. In: *New England Journal of Medicine* 362.25 (2010), pp. 2380–2388.

[104] Jean-Charles Soria et al. “First-line ceritinib versus platinum-based chemotherapy in advanced ALK-rearranged non-small-cell lung cancer (ASCEND-4): a randomised, open-label, phase 3 study”. In: *The Lancet* 389.10072 (2017), pp. 917–929.

[105] Julien Mazières et al. “Crizotinib therapy for advanced lung adenocarcinoma and a ROS1 rearrangement: results from the EUROS1 cohort”. In: *Journal of clinical oncology* 33.9 (2015), pp. 992–999.

[106] Georgina V Long et al. “Combined BRAF and MEK inhibition versus BRAF inhibition alone in melanoma”. In: *New England Journal of Medicine* 371.20 (2014), pp. 1877–1888.

[107] Paolo A Ascierto et al. “Cobimetinib combined with vemurafenib in advanced BRAFV600-mutant melanoma (coBRIM): updated efficacy results from a randomised, double-blind, phase 3 trial”. In: *The Lancet Oncology* 17.9 (2016), pp. 1248–1260.

[108] Christopher S Hourigan and Judith E Karp. “Minimal residual disease in acute myeloid leukaemia”. In: *Nature reviews Clinical oncology* 10.8 (2013), p. 460.

[109] Stephen Adler et al. “Minimum lesion detectability as a measure of PET system performance”. In: *EJNMMI physics* 4.1 (2017), p. 13.

[110] Marlise R Luskin et al. “Targeting minimal residual disease: a path to cure?” In: *Nature Reviews Cancer* 18.4 (2018), p. 255.

[111] Pierre Mordant et al. “Which metastasis management allows long-term survival of synchronous solitary M1b non-small cell lung cancer?” In: *European Journal of Cardio-Thoracic Surgery* 41.3 (2012), pp. 617–622.

[112] Janneke E Jaspers et al. “Loss of 53BP1 causes PARP inhibitor resistance in Brca1-mutated mouse mammary tumors”. In: *Cancer discovery* 3.1 (2013), pp. 68–81.

[113] William Pao et al. “Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain”. In: *PLoS medicine* 2.3 (2005), e73.

[114] Matthew J Niederst and Jeffrey A Engelman. “Bypass mechanisms of resistance to receptor tyrosine kinase inhibition in lung cancer”. In: *Sci. Signal.* 6.294 (2013), re6–re6.

[115] Barbara Norquist et al. “Secondary somatic mutations restoring BRCA1/2 predict chemotherapy resistance in hereditary ovarian carcinomas”. In: *Journal of clinical oncology* 29.22 (2011), p. 3008.

[116] Wataru Sakai et al. “Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers”. In: *Nature* 451.7182 (2008), p. 1116.

[117] Christopher J Lord, Andrew NJ Tutt, and Alan Ashworth. “Synthetic lethality and cancer therapy: lessons learned from the development of PARP inhibitors”. In: *Annual review of medicine* 66 (2015), pp. 455–470.

[118] Nicholas C Turner and Andrew NJ Tutt. *Platinum chemotherapy for BRCA1-related breast cancer: do we need more evidence?* 2012.

[119] Marieke A Vollebergh et al. “Genomic patterns resembling BRCA1-and BRCA2-mutated breast cancers predict benefit of intensified carboplatin-based chemotherapy”. In: *Breast Cancer Research* 16.3 (2014), R47.

[120] Ann-Marie Patch et al. “Whole–genome characterization of chemoresistant ovarian cancer”. In: *Nature* 521.7553 (2015), p. 489.

[121] Sohvi Blatter and Sven Rottenberg. “Minimal residual disease in cancer therapy–Small things make all the difference”. In: *Drug Resistance Updates* 21 (2015), pp. 1–10.

[122] Tannishtha Reya et al. “Stem cells, cancer, and cancer stem cells”. In: *nature* 414.6859 (2001), p. 105.

[123] Piet Borst. “Cancer drug pan-resistance: pumps, cancer stem cells, quiescence, epithelial to mesenchymal transition, blocked cell death pathways, persisters or what?” In: *Open biology* 2.5 (2012), p. 120066.

[124] Russell L Blaylock. “Cancer microenvironment, inflammation and cancer stem cells: A hypothesis for a paradigm change and new targets in cancer control”. In: *Surgical neurology international* 6 (2015).

[125] John J Kim and Ian F Tannock. “Repopulation of cancer cells during therapy: an important cause of treatment failure”. In: *Nature Reviews Cancer* 5.7 (2005), p. 516.

[126] Antonina V Kurtova et al. “Blocking PGE 2-induced tumour repopulation abrogates bladder cancer chemoresistance”. In: *Nature* 517.7533 (2015), p. 209.

[127] Andrea Viale et al. “Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function”. In: *Nature* 514.7524 (2014), p. 628.

[128] Daniel J Klionsky. “Autophagy revisited: a conversation with Christian de Duve”. In: *Autophagy* 4.6 (2008).

[129] Takatsune Shimizu et al. “IGF2 preserves osteosarcoma cell survival by creating an autophagic state of dormancy that protects cells against chemotherapeutic stress”. In: *Cancer research* 74.22 (2014), pp. 6531–6541.

[130] Benny Zhitomirsky and Yehuda G Assaraf. “Lysosomes as mediators of drug resistance in cancer”. In: *Drug Resistance Updates* 24 (2016), pp. 23–33.

[131] Jonathan A Ewald et al. “Therapy-induced senescence in cancer”. In: *JNCI: Journal of the National Cancer Institute* 102.20 (2010), pp. 1536–1546.

[132] Shawn M Cross et al. “A p53-dependent mouse spindle checkpoint”. In: *Science* 267.5202 (1995), pp. 1353–1356.

[133] Qin Wang et al. “Polyploidy road to therapy-induced cellular senescence and escape”. In: *International journal of cancer* 132.7 (2013), pp. 1505–1515.

[134] Shikhar Sharma, Theresa K Kelly, and Peter A Jones. “Epigenetics in cancer”. In: *Carcinogenesis* 31.1 (2010), pp. 27–36.

[135] Olivier Trédan et al. “Drug resistance and the solid tumor microenvironment”. In: *Journal of the National Cancer Institute* 99.19 (2007), pp. 1441–1454.

[136] Ravid Straussman et al. “Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion”. In: *Nature* 487.7408 (2012), p. 500.

[137] Timothy R Wilson et al. “Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors”. In: *Nature* 487.7408 (2012), p. 505.

[138] Robert D Schreiber, Lloyd J Old, and Mark J Smyth. “Cancer immunoediting: integrating immunity’s roles in cancer suppression and promotion”. In: *Science* 331.6024 (2011), pp. 1565–1570.

[139] OJ Finn. “Immuno-oncology: understanding the function and dysfunction of the immune system in cancer”. In: *Annals of oncology* 23.suppl_8 (2012), pp. viii6–viii9.

[140] Kazuhiro Kakimi et al. “CTLs regulate tumor growth via cytostatic effects rather than cytotoxicity: a few T cells can influence the growth of many times more tumor cells”. In: *Oncoimmunology* 4.3 (2015), e970464.

[141] Ian F Tannock and Daniela Rotin. “Acid pH in tumors and its potential for therapeutic exploitation”. In: *Cancer research* 49.16 (1989), pp. 4373–4384.

[142] Leo E Gerweck, Shashirekha Vijayappa, and Sergey Kozin. “Tumor pH controls the in vivo efficacy of weak acid and base chemotherapeutics”. In: *Molecular cancer therapeutics* 5.5 (2006), pp. 1275–1279.

[143] Trevor T Price et al. “Dormant breast cancer micrometastases reside in specific bone marrow niches that regulate their transit to and from bone”. In: *Science translational medicine* 8.340 (2016), 340ra73–340ra73.

[144] Julio A Aguirre-Ghiso. “Models, mechanisms and clinical evidence for cancer dormancy”. In: *Nature Reviews Cancer* 7.11 (2007), p. 834.

[145] Xiao-lei Gao et al. “Cancer cell dormancy: mechanisms and implications of cancer recurrence and metastasis”. In: *OncoTargets and therapy* 10 (2017), p. 5219.

[146] Filippo G Giancotti. “Mechanisms governing metastatic dormancy and reactivation”. In: *Cell* 155.4 (2013), pp. 750–764.

[147] Mary Osisami and Evan Keller. “Mechanisms of metastatic tumor dormancy”. In: *Journal of clinical medicine* 2.3 (2013), pp. 136–150.

[148] Ignacio A Rodriguez-Brenes and Dominik Wodarz. “Preventing clonal evolutionary processes in cancer: Insights from mathematical models”. In: *Proceedings of the National Academy of Sciences* 112.29 (2015), pp. 8843–8850.

[149] Nicholas McGranahan and Charles Swanton. “Clonal heterogeneity and tumor evolution: past, present, and the future”. In: *Cell* 168.4 (2017), pp. 613–628.

[150] Norman Wolmark et al. “Postoperative adjuvant chemotherapy or BCG for colon cancer: results from NSABP protocol C-011”. In: *JNCI: Journal of the National Cancer Institute* 80.1 (1988), pp. 30–36.

[151] Norman Wolmark et al. “The benefit of leucovorin-modulated fluorouracil as postoperative adjuvant therapy for primary colon cancer: results from National Surgical Adjuvant Breast and Bowel Project protocol C-03.” In: *Journal of Clinical Oncology* 11.10 (1993), pp. 1879–1887.

[152] Thierry André et al. “Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial”. In: *J clin oncol* 27.19 (2009), pp. 3109–3116.

[153] Early Breast Cancer Trialists’ Collaborative Group et al. “Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials”. In: *The Lancet* 379.9814 (2012), pp. 432–444.

[154] Nabeel Pervaiz et al. “A systematic meta-analysis of randomized controlled trials of adjuvant chemotherapy for localized resectable soft-tissue sarcoma”. In: *Cancer: Interdisciplinary International Journal of the American Cancer Society* 113.3 (2008), pp. 573–581.

[155] Zhu Yu et al. “Adjuvant endocrine monotherapy for postmenopausal early breast cancer patients with hormone-receptor positive: a systemic review and network meta-analysis”. In: *Breast Cancer* 25.1 (2018), pp. 8–16.

[156] Timothy Winton et al. “Vinorelbine plus cisplatin vs. observation in resected non–small-cell lung cancer”. In: *New England Journal of Medicine* 352.25 (2005), pp. 2589–2597.

[157] EMIL FREI et al. “The effectiveness of combinations of antileukemic agents in inducing and maintaining remission in children with acute leukemia”. In: *Blood* 26.5 (1965), pp. 642–656.

[158] Calvin B Bridges. “The origin of variations in sexual and sex-limited characters”. In: *The American Naturalist* 56.642 (1922), pp. 51–63.

[159] Nigel J O’Neil, Melanie L Bailey, and Philip Hieter. “Synthetic lethality and cancer”. In: *Nature Reviews Genetics* 18.10 (2017), p. 613.

[160] Greg Gibson. “Decanalization and the origin of complex disease”. In: *Nature Reviews Genetics* 10.2 (2009), p. 134.

[161] Marc Kirschner and John Gerhart. “Evolvability”. In: *Proceedings of the National Academy of Sciences* 95.15 (1998), pp. 8420–8427.

[162] Joanna Masel and Meredith V Trotter. “Robustness and evolvability”. In: *Trends in Genetics* 26.9 (2010), pp. 406–414.

[163] John L Hartman, Barbara Garvik, and Lee Hartwell. "Principles for the buffering of genetic variation". In: *Science* 291.5506 (2001), pp. 1001–1004.

[164] Angelique W Whitehurst et al. "Synthetic lethal screen identification of chemosensitizer loci in cancer cells". In: *Nature* 446.7137 (2007), p. 815.

[165] Suzanne L Rutherford and Susan Lindquist. "Hsp90 as a capacitor for morphological evolution". In: *Nature* 396.6709 (1998), p. 336.

[166] Ben Lehner et al. "Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways". In: *Nature genetics* 38.8 (2006), p. 896.

[167] Sebastian MB Nijman. "Synthetic lethality: general principles, utility and detection using genetic screens in human cells". In: *FEBS letters* 585.1 (2011), pp. 1–6.

[168] Leland H Hartwell et al. "Integrating genetic approaches into the discovery of anticancer drugs". In: *Science* 278.5340 (1997), pp. 1064–1068.

[169] I Bernard Weinstein. "Addiction to oncogenes–the Achilles heal of cancer". In: *Science* 297.5578 (2002), pp. 63–64.

[170] Nicole L Solimini, Ji Luo, and Stephen J Elledge. "Non-oncogene addiction and the stress phenotype of cancer cells". In: *Cell* 130.6 (2007), pp. 986–988.

[171] Helen E Bryant et al. "Specific killing of BRCA2-deficient tumours with inhibitors of poly (ADP-ribose) polymerase". In: *Nature* 434.7035 (2005), p. 913.

[172] Peter C Fong et al. "Inhibition of poly (ADP-ribose) polymerase in tumors from BRCA mutation carriers". In: *New England Journal of Medicine* 361.2 (2009), pp. 123–134.

[173] Michael Karin, Zheng-gang Liu, and Ebrahim Zandi. "AP-1 function and regulation". In: *Current opinion in cell biology* 9.2 (1997), pp. 240–246.

[174] Cun-Yu Wang et al. "NF-κB induces expression of the Bcl-2 homologue A1/Bfl-1 to preferentially suppress chemotherapy-induced apoptosis". In: *Molecular and cellular biology* 19.9 (1999), pp. 5923–5929.

[175] Sarah R Walker et al. "The transcriptional modulator BCL6 as a molecular target for breast cancer therapy". In: *Oncogene* 34.9 (2015), p. 1073.

[176] Balazs Győrffy et al. "An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients". In: *Breast cancer research and treatment* 123.3 (2010), pp. 725–731.

[177] Stella Maris Ranuncolo et al. "Bcl-6 mediates the germinal center B cell phenotype and lymphomagenesis through transcriptional repression of the DNA-damage sensor ATR". In: *Nature immunology* 8.7 (2007), p. 705.

[178] Stella M Ranuncolo, Jose M Polo, and Ari Melnick. "BCL6 represses CHEK1 and suppresses DNA damage pathways in normal and malignant B-cells". In: *Blood cells, molecules, and diseases* 41.1 (2008), pp. 95–99.

[179] Ryan T Phan and Riccardo Dalla-Favera. "The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells". In: *Nature* 432.7017 (2004), p. 635.

[180] Ryan T Phan et al. "BCL6 interacts with the transcription factor Miz-1 to suppress the cyclin-dependent kinase inhibitor p21 and cell cycle arrest in germinal center B cells". In: *Nature immunology* 6.10 (2005), p. 1054.

[181] Renee D Paulsen and Karlene A Cimprich. "The ATR pathway: fine-tuning the fork". In: *DNA repair* 6.7 (2007), pp. 953–966.

[182] Deborah L Hardie et al. "Quantitative analysis of molecules which distinguish functional compartments within germinal centers". In: *European journal of immunology* 23.5 (1993), pp. 997–1004.

[183] Laura Pasqualucci et al. "Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas". In: *Nature* 412.6844 (2001), p. 341.

[184] Giorgio Cattoretti et al. "Deregulated BCL6 expression recapitulates the pathogenesis of human diffuse large B cell lymphomas in mice". In: *Cancer cell* 7.5 (2005), pp. 445–455.

[185] Beverly W Baron et al. "The human BCL6 transgene promotes the development of lymphomas in the mouse". In: *Proceedings of the National Academy of Sciences* 101.39 (2004), pp. 14198–14203.

[186] Pramod C Rath and Bharat B Aggarwal. "TNF-induced signaling in apoptosis". In: *Journal of clinical immunology* 19.6 (1999), pp. 350–364.

[187] Xuanmao Jiao et al. "c-Jun induces mammary epithelial cellular invasion and breast cancer stem cell expansion". In: *Journal of Biological Chemistry* 285.11 (2010), pp. 8218–8226.

[188] Mark Jesus M Magbanua et al. "Serial expression analysis of breast tumors during neoadjuvant chemotherapy reveals changes in cell cycle and immune pathways associated with recurrence and response". In: *Breast Cancer Research* 17.1 (2015), p. 73.

[189] Donatella Aldinucci and Alfonso Colombatti. "The inflammatory chemokine CCL5 and cancer progression". In: *Mediators of inflammation* 2014 (2014).

[190] Jhajaira M Araujo et al. "Effect of CCL5 expression in the recruitment of immune cells in triple negative breast cancer". In: *Scientific reports* 8.1 (2018), p. 4899.

[191] Marco Velasco-Velázquez et al. "CCR5 antagonist blocks metastasis of basal breast cancer cells". In: *Cancer research* 72.15 (2012), pp. 3839–3850.

[192] Santosh Kumar Singh et al. “CCR5/CCL5 axis interaction promotes migratory and invasiveness of pancreatic cancer cells”. In: *Scientific reports* 8.1 (2018), p. 1323.

[193] Chad J Creighton et al. “Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features”. In: *Proceedings of the National Academy of Sciences* 106.33 (2009), pp. 13820–13825.

[194] Michael Untch et al. “Pathologic complete response after neoadjuvant chemotherapy plus trastuzumab predicts favorable survival in human epidermal growth factor receptor 2–overexpressing breast cancer: results from the TECHNO trial of the AGO and GBG study groups”. In: *Journal of Clinical Oncology* 29.25 (2011), pp. 3351–3357.

[195] Andrea Walens et al. “CCL5 promotes breast cancer recurrence through macrophage recruitment in residual tumors”. In: *eLife* 8 (2019), e43653.

[196] Ke Gong et al. “TNF-driven adaptive response mediates resistance to EGFR inhibition in lung cancer”. In: *The Journal of clinical investigation* 128.6 (2018), pp. 2500–2518.

[197] Takaya Abe and Toshihiko Fujimori. “Reporter mouse lines for fluorescence imaging”. In: *Development, growth & differentiation* 55.4 (2013), pp. 390–405.

[198] Ashleigh C McLean et al. “Performing vaginal lavage, crystal violet staining, and vaginal cytological evaluation for mouse estrous cycle staging identification”. In: *JoVE (Journal of Visualized Experiments)* 67 (2012), e4389.

[199] Sylwia Gawrzak et al. “MSK1 regulates luminal cell differentiation and metastatic dormancy in ER+ breast cancer”. In: *Nature cell biology* 20.2 (2018), p. 211.

[200] Robert H Kutner, Xian-Yang Zhang, and Jakob Reiser. “Production, concentration and titration of pseudotyped HIV-1-based lentiviral vectors”. In: *Nature protocols* 4.4 (2009), p. 495.

[201] Simon Andrews et al. *FastQC: a quality control tool for high throughput sequence data*. 2010.

[202] Marcel Martin. “Cutadapt removes adapter sequences from high-throughput sequencing reads”. In: *EMBnet. journal* 17.1 (2011), pp. 10–12.

[203] Chien-Chi Lo and Patrick SG Chain. “Rapid evaluation and quality control of next generation sequencing data with FaQCs”. In: *BMC bioinformatics* 15.1 (2014), p. 366.

[204] Daehwan Kim et al. “TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions”. In: *Genome biology* 14.4 (2013), R36.

[205] Simon Anders, Paul Theodor Pyl, and Wolfgang Huber. "HTSeq—a Python framework to work with high-throughput sequencing data". In: *Bioinformatics* 31.2 (2015), pp. 166–169.

[206] Michael I Love, Wolfgang Huber, and Simon Anders. "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". In: *Genome biology* 15.12 (2014), p. 550.

[207] Simon Anders and Wolfgang Huber. "Differential expression analysis for sequence count data". In: *Genome biology* 11.10 (2010), R106.

[208] Korbinian Strimmer. "fdrtool: a versatile R package for estimating local and tail area-based false discovery rates". In: *Bioinformatics* 24.12 (2008), pp. 1461–1462.

[209] Magdalena Strauss et al. "GPseudoClust: deconvolution of shared pseudo-trajectories at single-cell resolution". In: *BioRxiv* (2019), p. 567115.

CHAPTER 7

LISTS OF ABBREVIATIONS, FIGURES AND TABLES

Akt	AKT Serine/Threonine Kinase 1
ALK	Anaplastic Lymphoma Kinase
AP-1	Activator Protein 1
aPKC	Atypical Protein kinase C
ATM	ATM serine/threonine kinase
ATR	ATR serine/threonine kinase
Bcl-2	B-cell lymphoma 2
BCL2L1	Bcl-2-like protein 1
BCL6	B-cell lymphoma 6
BDP	Big Data Processor
BIRC2	Baculoviral IAP Repreat Containing 2
BRAF	Serine/threonine-protein kinase B-Raf
BRCA1	Breast Cancer 1 susceptibility protein
CASP3	Caspase-3
CATS	Context Aware Trainable Segmentation
CCL5	Chemokine (C-C motif) ligand 5
Cdc42	Cell division control protein 42 homolog
CDH1	Cadherin-1
CDK1	Cyclin-dependent kinase 1
CHEK1	Checkpoint kinase 1
CDKN2A	cyclin-dependent kinase Inhibitor 2A
CML	Chronic myelogenous leukemia
CR	Complete Remission
CSC	Cancer stem cell
CT	Computed tomography
CTL	Cytotoxic T lymphocyte
CXCR4	C-X-C chemokine receptor type 4

CXCR5 C-X-C chemokine receptor type 5

DAB 3,3-Diaminobenzidine

DAPI 4,6-diamidino-2-phenylindole

DCIS Ductal carcinoma in situ

DLBCL Diffuse large B-cell lymphoma

Dlg Discs large homolog 1

DMEM Dulbecco's modified Eagle's medium

DNA Deoxyribonucleic acid

DOX Doxycycline

E6 Viral oncogene E6

EDTA Ethylenediaminetetraacetic acid

EGFR Epidermal growth factor receptor

EMT Epithelial–mesenchymal transition

ER Estrogen receptor

FACS Fluorescence-activated cell sorting

FAS CD95 or Apoptosis antigen 1

FDR False Discovery Rate

FEP Fluorinated ethylene propylene

FFPE Formalin-fixed paraffin-embedded

FosB FBJ murine osteosarcoma viral oncogene homolog B

FVB Friend leukemia virus B mouse strain

GC Germinal Centers

GEMM Genetically engineered mouse model

GFP Green fluorescent protein

GM130 Golgin subfamily A member 2

GO Gene Ontology

GRB7 Growth factor receptor-bound protein 7

H2B	Histone H2B
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER2	Human epidermal growth factor receptor 2
HIF1	Hypoxia-inducible factor 1
HPV	Human papillomavirus
Hsp	Heat shock protein
IC_{50}	Half maximal inhibitory concentration
IF	Immunofluorescence
IHC	Immunohistochemistry
IKK2/4	Inhibitor of nuclear factor kappa-B kinase 2/4
ITGA6	Alpha-6 integrin
KDM5A	Lysine-specific demethylase 5A
Ki67	Proliferation marker Ki67
c-Kit	Mast/stem cell growth factor receptor
LASSO	Least absolute shrinkage and selection operator
LCIS	Lobular carcinoma in situ
Lgl	Lethal 2 giant larvae protein homolog 1
LKB1	Liver kinase B1
lncRNA	Long non-coding RNA
MAPK	Mitogen-activated protein kinase
MF	Molecular Function
miRNA	Micro RNA
MMTV	Mouse mammary tumor virus
MRD	Minimal residual disease
MRI	Magnetic resonance imaging
mTOR	mammalian target of rapamycin
Myc	V-myc avian myelocytomatosis viral oncogene homolog

NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NGS	Next-generation sequencing
NSCLC	Non-small-cell lung carcinoma
OCT	Optimal cutting temperature compound
PARP	Poly (ADP-ribose) polymerase
PBS	Phosphate-buffered saline
PCA	Principal component analysis
PDAC	Pancreatic ductal adenocarcinoma
PET	Positron emission tomography
PFA	Paraformaldehyde
PGAP3	Post-GPI Attachment To Proteins 3
PGE2	Prostaglandin E2
PI3K	Phosphoinositide 3-kinase
PLC	Phospholipase C
PR	Progesterone receptor
PtdIns3P	Phosphatidylinositol 3-phosphate
PTEN	Phosphatase and tensin homolog
qPCR	Quantitative polymerase chain reaction
RAS	Ras (Rat sarcoma) oncoprotein
c-Rel	REL proto-oncogene, NF-kB subunit
RLU	Relative Luminescence Units
RNA	Ribonucleic acid
ROS	Reactive Oxygen Species
rtTA	Reverse tetracycline-controlled transactivator
Scrib	Scribbled homolog (Drosophila)
SEM	Standard error of the mean
SPIM	Selective Plane Illumination Microscopy

TAC	Transcriptome Analysis Console
TDLU	Terminal duct lobular unit
Tet	Ten-eleven translocation
TGFβ	Transforming growth factor beta
TIS	Therapy Induced Senescence
TKI	Tyrosine kinase inhibitor
TNBC	Triple-negative breast cancer
TNFα	Tumor necrosis factor alpha
TP53	Tumor protein p53
VHL	von Hippel–Lindau tumor suppressor
WPRE	Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element
ZO-1	Zonula occludens-1

LIST OF FIGURES

1.1 Breast anatomy . . . 14
1.2 Ductal and Lobular carcinomas of the breast . . . 15
1.3 Molecular approaches to HER2 targeted therapy. . . . 18

2.1 Principles of Selective Plane Illumination Microscopy . . . 26
2.2 Schematic showing the adaptation of the tissue wide tumourigesis model in 3D culture to allow stochastic transformation . . . 29
2.3 qPCR expression analysis of oncogene expression for establishing optimal doxycycline concentrations. . . . 30
2.4 Oncogene co-localization analysis in tissue wide and stochastic breast tumour organoids . . . 31
2.5 Sample mounting procedure for light sheet microscopy with the InVi SPIM 32
2.6 Time lapse imaging of stochastically transduced organoids . . . 34
2.7 Characterization of stochastic tumourigeneis *in vitro* . . . 35
2.8 Pre-processing of big data images acquired at the light sheet microscope. . . 36
2.9 Nuclear segmentation of tumourigenic cells within a stochastically transformed organoid . . . 38
2.10 Schematic representation of the big-image data analysis pipeline developed to analyze the light sheet microscopy images. . . . 39
2.11 Assesment of nuclear segmentation accuracy on Imaris . . . 40
2.12 Respresentaive tracking of single transformed cells in breast organoids . . . 41
2.13 Feature analysis of stochastically transformed cells within organoids . . . 43
2.14 Best logistic regression model for all possible linear combinations of features based on the Akaike information criterion . . . 44
2.15 Alternate logistic regression model for all possible combinations of features using LASSO regularization . . . 45

2.16 Representative stochastic tumourigenesis shown in organoids at the start and end of light sheet imaging . . . 46

3.1 Exploiting synthetic lethality for anti-cancer therapy . . . 60
3.2 Characterization of the *in vitro* culture system to model minimal residual disease. . . . 63
3.3 Characterization of the *in vitro* cultures during tumour organoid regression using immunofluorescence staining. . . . 64
3.4 Hallmark timepoints of *in vitro* tumour progression and regression at which RNA sequncing was performed . . . 66
3.5 PCA analysis plot for RNA Sequencing data from *in vitro* cultures . . . 69
3.6 Clustering analysis of deferentially regulated genes along the RNA Sequencing kinetic . . . 70
3.7 Example gene cluster trajectories from clustering analysis . . . 71
3.8 Manual clustering and gene list curation using trajectory and correlation analysis . . . 73
3.9 GO terms enriched in potential synthetic lethal target trajectory clusters . 74
3.10 Characterization of the *in vivo* modelling of minimal residual disease . . . 75
3.11 PCA analysis plot for micro array datasets from *in vivo* harvests during tumour regression in mice . . . 77
3.12 Transcriptomic levels and protein validation of *c-jun* as a target for synthetic lethality. . . . 80
3.13 Experimental design for the *in vitro* target validation . . . 81
3.14 Transcriptomic levels and protein validation of *Bcl6* as a target for synthetic lethality. . . . 84
3.15 Schematic representation of the molecular players in the NF-κB cascade that are upregulated *in vitro* breast cancer regression. . . . 85
3.16 Cell Viability and Cell Death readouts from synthetic lethality testing of inhibitor Tanshinone IIA. . . . 86
3.17 Cell Viability and Cell Death readouts from synthetic lethality testing of inhibitor JNK Inhibitor SP600125. . . . 87
3.18 Optimizing image acquisition parameters at the InVi SPIM . . . 88
3.19 Tumour regression recorded at the In-Vi SPIM with manual tracking . . . 90
3.20 Segmentation of nuclei during tumour regression. . . . 92

5.1 Hierarchical clustering of transduced cells in bi-transgenic organoids undergoing stochastic tumourigenesis. . . . 121
5.2 Hierarchical clustering of transduced cells in bi-transgenic organoids undergoing stochastic tumourigenesis. . . . 122
5.3 Gene Ontology Enrichment analysis of genes with 'Target' trajectories. . . . 122

5.4 Gene Ontology Enrichment analysis of genes with 'Early Responder' trajectories. . . . 123
5.5 Gene Ontology Enrichment analysis of genes with 'Intermediate Responder' trajectories. . . . 123
5.6 Gene Ontology Enrichment analysis of genes with 'Late Responder' trajectories. . . . 123
5.7 Gene Ontology Enrichment analysis of genes with 'Shoulder' trajectories. . . . 123
5.8 Network mapping analysis for GO:0098742 (Cell-cell adhesion via plasma-membrane adhesion molecules) 124
5.9 Network mapping analysis for GO:0000571 (FAS pathway and Stress induction of HSP regulation) 125
5.10 Network mapping analysis for GO:0001213 (Formation of the cornified envelope) 126
5.11 Network mapping analysis for GO:0017017 (MAP kinase tyrosine/serine/threonine phosphatase activity) 127
5.12 Network mapping analysis for GO:0002819 (Regulation of adaptive immune response) 128
5.13 Network mapping analysis for GO:0031663 (Lipopolysaccharide-mediated signaling pathway) 129
5.14 Network mapping analysis for GO:0000618 (Oxidative Stress Induced Senescence) 130
5.15 Network mapping analysis for GO:0000097 (Downregulation of ERBB4 signaling) 131
5.16 Network mapping analysis for GO:0008593 (Regulation of Notch signaling pathway) 132
5.17 Network mapping analysis for GO:0002009 (Morphogenesis of an epithelium) 133
5.18 Cell Viability and Cell Death readouts from synthetic lethality testing of inhibitor Maraviroc. . . . 134
5.19 Cell Viability and Cell Death readouts from synthetic lethality testing of inhibitor IKK2 Inhibitor IV. . . . 135
5.20 Cell Viability and Cell Death readouts from synthetic lethality testing of inhibitor IKK2 Inhibitor VIII. . . . 136
5.21 Cell Viability and Cell Death readouts from synthetic lethality testing of inhibitor FX-1. . . . 137

LIST OF TABLES

1.1 Overview of the molecular sub-types of breast cancer 17

3.1 Overview of gene numbers in selected manual clusters 72
3.2 GO Terms enriched in the five potentially synthetic lethal clusters manually curated from RNA Seq analysis 72
3.3 Inhibitor list tested for synthetic lethality in regressing mammary matrigel cultures 82

5.1 List of transduced organoids analyzed for feature analysis. 120