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

M.Sc. Ksenija Radić Shechter
born in Novi Pazar, Serbia

Oral examination: February 11th, 2019

Exploring the metabolic landscape and resulting vulnerabilities of minimal residual disease in breast cancer

Referees: Dr. Carsten Schultz

Prof. Dr. Britta Brügger

to Alex, Mom and Dad

ACKNOWLEDGEMENTS

This thesis owes its existence to the wide-open door of EMBL to the young scientists. That is why I would like, first and foremost, to express my gratitude to EMBL for giving me the great opportunity to continue my education at this prestigious international institute. Being here, among esteemed scientists from all over the world, motivated me strongly, enabled me new insights into the scientific challenges of today and helped me improve my knowledge and skills. In that sense, it was an honor and privilege to be part of this eminent scientific community. Also, expressions of deep gratitude refer to the Deans and employers of the Graduate Office (**Helke Hillebrand**, **Monika Lachner**, **Matija Grgurinovic** and **Meriam Bezohra**) and other professional, administrative and technical services of EMBL for professional support and precious assistance throughout my PhD studies.

Special and great gratitude to my supervisor **Martin Jechlinger** for accepting me in his group, giving me the chance to take a step in the field that has filled me with excitement, allowing me to take part in the project that sparked my scientific curiosity and interest and supporting me on this pursuit.

Furthermore, I am very grateful to my **Thesis Advisory Committee**: **Britta Brügger**, **Carsten Schultz**, **Kiran Raosaheb Patil**, **Theodore Alexandrov** and **Stephan Herzig** for their invested time, discussing all the important issues from different perspectives and giving me valuable advices and scientific inputs. I am very grateful to **Stefan Wiemann** for agreeing to join the Defense Committee.

This is a wonderful opportunity to mention and say a big thanks to all present and past members of **Jechlinger group**. My dear colleagues and friends, thank you for all the nice times during these years, as well as the support and help you have given me.

Honest gratitude to the lab colleagues in Heidelberg: **Ashna Alladin** – for many hours and efforts shared together in our numerous experiments, support and help whenever I needed it as well as for the nice moments outside EMBL with her and **Areeb Javed**; **Sylwia Gawrzak** – for valuable scientific discussion and advice, her friendly spirit, editorial support, critical feedback, help and encouragement, as well as fun times outside the lab; **Marta Garcia Montero** – for her help in working with mice and pleasant little surprises to fix my bad days; **Matthew Stephen Boucher** and **Savannah Jade Jackson** – for their help and willingness to support me whenever it was necessary, help with mouse work and for proofreading and doing language corrections in my thesis; **Lucas Chaible** – for his helpful tips and great friendship with him and his lovely **Angela Schwahofer de Carvalho** whose magical hands always fixed me after long and hard experiments and kept me in a good condition; **Federico Villa** – for helping with experiments during his short stay in the lab.

I also remember with pleasure my start in EMBL Monterotondo and all nice people I met there. I am indebted and greatly thankful to: **Kristina Havas Cavalletti** - for patience, good will and time spent in teaching me and shaping me as a scientist; although very short, time spent with her was valuable and enjoyable for me, giving me a good basis for the future work; **Tomoko Ishibashi** –

for a warm welcome and unselfishly sharing her knowledge on many practical skills she has mastered**; Giuseppe (Pino) Chiapparelli** - for being an excellent mouse caretaker and my Italian brother; **Maria Giubettini** – for being my great friend; and all the other dear colleagues and friends with whom I spent short but unforgettable time in Monterotondo.

I would like to express my appreciations to our collaborators from **Patil, Brügger, Alexandrov** group and **GSK-Cellzome**, who brought their expertise to this project and contributed to its progress: **Eleni Kafkia**, for all her help and advices in all the metabolic experiments, for being a great colleague, collaborator and a real team player; **Daniel Sevin** for his efficiency, responsiveness and help in obtaining data for untargeted metabolomics; **Andrew Palmer** for patiently teaching me imaging mass spectrometry; **Christian Lüchtenborg** for obtaining lipidomics datasets and **Katharina Zirngibl** for analysis of transcriptomics data. Special thanks to **EMBL Core Facilities** – **LAR** (all the animal technicians), **Gene Core** (Vladimir Benes, Ferris Jung), **MCF** (Prasad Phapale), **CBCF** (Kerstin Putzker), **ALMF and EMCF**.

I am very thankful to **Rocio Sottilo** and her lab at the DKFZ, especially to **Charles Yuanyuan Chen** for his immense patience and assistance with quantification during my long days of slide scanning and analyses; as well as the whole **Sottilo group** for giving me support and making this journey more enjoyable.

Many thanks to **Hanna Sladitschek**, **Matt Rogon** and other colleagues at EMBL who invested their time on constructive discussions and supported me with valuable advices. Thanks to **Dimitri Kromm** for translating the summary of this thesis to German language.

My sincere gratitude to my **former professors and colleagues at INGEB** from Sarajevo for my first steps in research and encouragement for greater challenges.

Finally, my deepest gratitude goes to **my beloved husband Alex** for his immense patience and support during my studies and thesis writing, valuable discussions, all the help and constructive suggestions especially in the graphical-technical design and, most importantly, for his big and warm heart and all that he means in my life.

I would not be here without all the love and support from the people whose lives are deeply bound with mine. Nothing would be the same without **my amazing Mom and Dad** who have been giving me unconditional love and support throughout all my life; **Keka**, **Dado**, **Grandma**, **uncle Slavce**, **Iva**, **Veronica**, **Ludmilla**, **Eva**, **Solomon** and **all the other members of my family** who believed in me and my choices, supported me; as well as **all my wonderful friends for life** everywhere in the world. Although now scattered across the globe and physically far away, you are all a big part of my life and the source of my strength.

Working for many long nights and dealing with though academic life would be impossible without a proper good metal music, especially Symphony-X and Neurotech Wolf´s "Symphonies".

SUMMARY

Despite overall improvements in breast cancer management, breast cancer continues to be a major health challenge worldwide, with high mortality rates. Following initially successful therapies, some cells are able to evade the treatment and remain in the body at undetectable levels (known as minimal residual disease) and could eventually lead to a lethal tumor recurrence. Understanding the nature of this elusive cell population is of a great importance in tumor eradication and relapse prevention. Nevertheless, the intricate mechanistic details about these processes and the nature of these cells remain poorly understood, mainly due to difficulties in obtaining patient material of MRD and being able to study them over the course of treatment.

To characterize MRD in breast cancer, we employed an inducible *TetO-MYC/TetO-Neu/MMTV-rtTA* mouse model and primary 3D cultures of mammary organoids that yield a correlate of MRD upon oncogene silencing. A combination of immunofluorescence, RNA sequencing, lipidomics and metabolomics revealed the unique nature of the residual cells. Despite the phenotypic similarity to the normal population, the residual cells exhibited a distinct transcriptional profile. This profile was also different when compared to the tumor, highlighting the unique properties of the residual cell population. In addition, lipid profiles of the residual cells were also distinguished from both normal and tumor populations.

Surprisingly, despite the inactivated oncogenes and seemingly normal phenotype of the residual structures, the residual cells bore a metabolic resemblance to tumor cells. They retained some tumor metabolic hallmarks, which persevered long after the oncogenes were silenced and tumors had regressed. This was demonstrated in our 3D cultures and verified *in vivo* on histological sections and fresh samples of the mammary glands following tumor regression. Enhanced glycolysis, the urea cycle and NOS2 activity were the most prominent features preserved in the residual cells. As shown in correlation with publicly available microarray datasets of patient samples following neoadjuvant treatment, these traits could be particularly important in MRD of the basal-like HER2 positive and HER2 negative breast cancer subtypes.

In conclusion, our findings suggest vast and profound effects of the changes happening over the course of tumor progression, which result in an altered metabolic network even in the absence of oncogene signaling, implying the existence of a "metabolic memory" probably imprinted through the changes in the epigenetic landscape. This phenomenon, particularly its potential driver(s), remains to be investigated. Furthermore, enhanced glycolysis, urea cycle and NOS2 activity in the residual cells compared to the normal, could provide an opportunity to interfere with MRD, offering the potential of preventing tumor recurrences.

ZUSAMMENFASSUNG

Trotz weitreichender Entwicklungen in der Brustkrebsbehandlung, stellt Brustkrebs, mit seinen hohen Todesfällen, nach wie vor eine wesentliche gesundheitliche Herausforderung auf globaler Ebene dar. Nach zunächst erfolgreicher Therapie schaffen es manche Zellen der Behandlung zu entgehen und in nicht detektierbarer Anzahl im Körper zu verweilen (minimale Resterkrankung (MRD) genannt). Zu einem späteren Zeitpunkt können sie zu einem Wiederauftreten des Tumors mit tödlichen Folgen führen. Daher ist es für die Tumorbeseitigung und zur Prävention der Tumorwiederkehr von größter Bedeutung das Verhalten dieser schwer detektierbaren Zell-Population zu verstehen. Gleichwohl sind die komplizierten Mechanismen dieser Prozesse sowie die Eigenschaften dieser Zellen nur unzureichend verstanden. Dies liegt hauptsächlich an Schwierigkeiten bei der Beschaffung von Patientengut sowie an den begrenzten Möglichkeiten dieses während des gesamten Behandlungsverlaufs zu untersuchen.

Um MRD im Falle von Brustkrebs zu charakterisieren, verwendeten wir ein induzierbares *TetO-MYC/TetO-Neu/MMTV-rtTA* Mausmodell sowie 3D-Primärzellkulturen von Brust-Organoiden welche nach Gen-Stilllegung eine Korrelation zu MRD bieten. Immunfluoreszenz in Kombination mit RNA-Sequenzierung, Lipidomik und Metabolomik offenbarte die einzigartige Natur der residualen Zellen. Trotz der phänotypischen Ähnlichkeit zur normalen Population, wiesen die residualen Zellen ein eindeutiges Transkriptionsprofil auf. Dieses Profil unterschied sich weiterhin von dem Tumor, was die einzigartigen Eigenschaften der residualen Population nochmals unterstreicht. Zusätzlich unterschieden sich die Lipid-Profile der residualen Zellen von denen der normalen und der Tumorzellen.

Erstaunlicherweise wiesen die Residualzellen trotz der inaktivierten Onkogene und der scheinbar normalen Phänotypen eine metabolische Ähnlichkeit zu Krebszellen auf. So behielten die Residualzellen einige metabolische Kennzeichen, die sie lange nach der Onkogenhemmung und Turmorregression bewahrten. Dies wurde in unseren 3D-Kulturen *in vivo* und anhand von histologischen Schnitten sowie frischen Proben von Milchdrüsen nach Tumorregression demonstriert. Eine gesteigerte Glykolyse, der Harnstoffzyklus sowie die NOS2-Aktivität waren die auffälligsten Merkmale, die in den Residualzellen erhalten blieben. Wie in Korrelation zu öffentlich zugänglichen Microarray-Datensätzen von Patientenproben nach neoadjuvanter Therapie gezeigt, könnten diese Merkmale besonders wichtig für die MRD der basal-ähnlichen HER2-positven und HER2-negativen Krebssubtypen sein.

Zusammenfassend weisen unsere Erkenntnisse auf breite und tiefgreifende Auswirkungen der Änderungen, die während der Tumorfortschreitung stattfinden, hin. Das Ergebnis ist ein abgewandeltes metabolisches Netzwerk, das auch bei Abwesenheit

onkogener Signale bestehen bleibt. Dies wiederum impliziert die Existenz eines „Metabolomischen Gedächtnisses“, welches vermutlich durch Änderung der epigenetischen Landschaft imprägniert wird. Dieses Phänomen, vor allem sein potenzieller Antrieb bzw. seine potenziellen Antriebe müssen noch untersucht werden. Weiter kann die gesteigerte Glykolyse, der Harnstoffzyklus und die NOS2-Aktivität der residualen Zellen möglicherweise genutzt werden, um in die MRD einzugreifen. Dies könnte die Möglichkeit bieten das erneute Auftreten des Tumors zu unterbinden.

CONTENTS

Table of Contents

CONTENTS

CONTENTS

Chapter 1

INTRODUCTION

INTRODUCTION

Cancer is a disease caused by the uncontrolled proliferation of abnormal cells, leading to the formation of a malignant mass that invades and destroys the surrounding healthy tissue (1). Even though, in its essence, the term "cancer" reflects a simple phenomenon, it encompasses a group of heterogeneous ailments with their own etiology, diverse features, specific risk factors and different courses of progression and treatment. This heterogeneity and immense complexity is what makes cancer a major health problem worldwide and a leading cause of death (2). According to the data from 2015, there have been 14 million new cases of cancer and 8,7 million of deaths (3) worldwide. Moreover, projections predict that the number of new cancer cases and cancer-related deaths will increase, due to increase in proportion of the aging population, and also increased exposure to environmental risk factors brought about by specific lifestyles. This includes unhealthy diet and physical inactivity which leads to a rise in obesity levels, alcohol and tobacco consumption, exposure to certain viruses such as human papillomavirus (HPV) and hepatitis B (HBV) and C (HCV) viruses (2).

Despite these concerning statistics, great steps forward have been made in understanding cancer in particular on a molecular level, leading to the recognition of many different types and subtypes. This has made it possible to identify key molecular events and players driving tumor progression in certain subgroups and therefore develop more specific drugs. However, due to differences in genetic makeup and great inter- and intratumor heterogeneity, efficacy of the drugs and patients' response to the treatment are not universal, imposing the need for better therapeutic options that are tailored for specific cases. Better understanding of tumor biology could eventually make this possible, along with the development of targeted therapies and pharmacogenomics (4).

Overall, growing knowledge about risk factors, etiology and molecular biology of certain tumor types, earlier detection of the disease and better therapeutic strategies have led to improvements in cancer management and a decrease in overall cancer death rates (5). This is particularly true for breast, lung, prostate and colorectal cancer (5). Nonetheless, challenges still remain, mainly imposed by resistance to therapy, tumor recurrence and metastatic disease which remains still incurable, and is a major cause of cancer-related deaths. Understanding the elusive and intricate processes behind these phenomena is of utmost importance as it gives hope that this frightening disease could be managed in a more effective way. This could be by giving more efficient treatments and combined therapies that would prevent relapses, be more selective and exert less side-effects on the healthy tissues, eventually leading to longer life of cancer patients.

1.1 BREAST CANCER

Breast cancer is a malignant tumor originating in the mammary glands.

A simplified depiction of the breast anatomy is given in Figure 1.

The female breast consists of:

- 15-20 lobes of glandular tissue divided into 20-40 smaller lobules (containing rounded acini of milk-producing cells);
- small milk ducts connected to the lobules, that gradually join into larger ducts and a duct system that opens at the nipple;
- surrounding adipose and connective tissue (6).

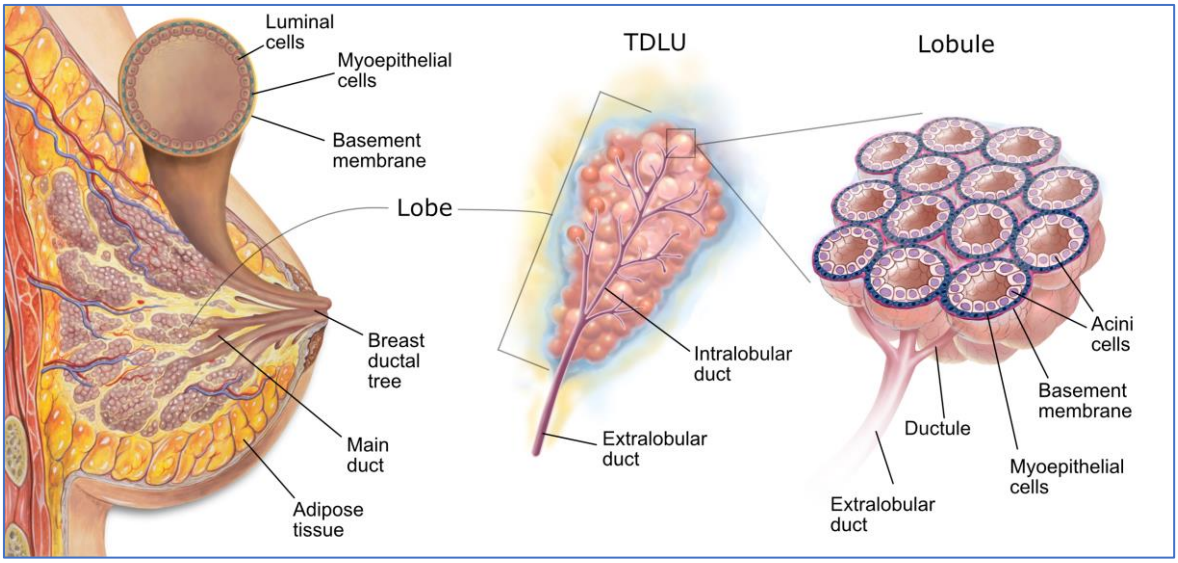

Figure 1: Female breast anatomy. Figure modified from (7) and (8), with permission. Abbreviations are explained in the List of Abbreviations.

The basic functional unit of the mammary gland is the terminal ductal lobular unit (TDLU), comprised of a lobule, intralobular and extralobular duct (9). Cuboidal or columnar, polarized luminal epithelial cells are localized in an innermost layer around the lumen of the healthy ducts and acini, surrounded by an outer layer of myoepithelial (basal) cells. These are divided from the surrounding stroma by a basement membrane, which serves as an important barrier and dynamic structure composed of extracellular matrix proteins (10,11).

Breast cancer arises mostly from abnormal epithelial cells in the lobules (lobular carcinoma) or ducts (ductal carcinoma). In rare cases it can also originate from the structures of breast stroma (12). Following an oncogenic insult (Figure 2A), cells divide uncontrollably, filling up the mammary gland lumen, leading to lobular carcinoma *in situ* (LCIS) or ductal carcinoma *in situ* (DCIS), which correspond to the earlier stages of cancer (Figure 2B). At these earlier stages, the basement membrane stays intact, as opposed to invasive and metastatic cancer (Figure 2C,D) in which DCIS or LCIS can progress further.

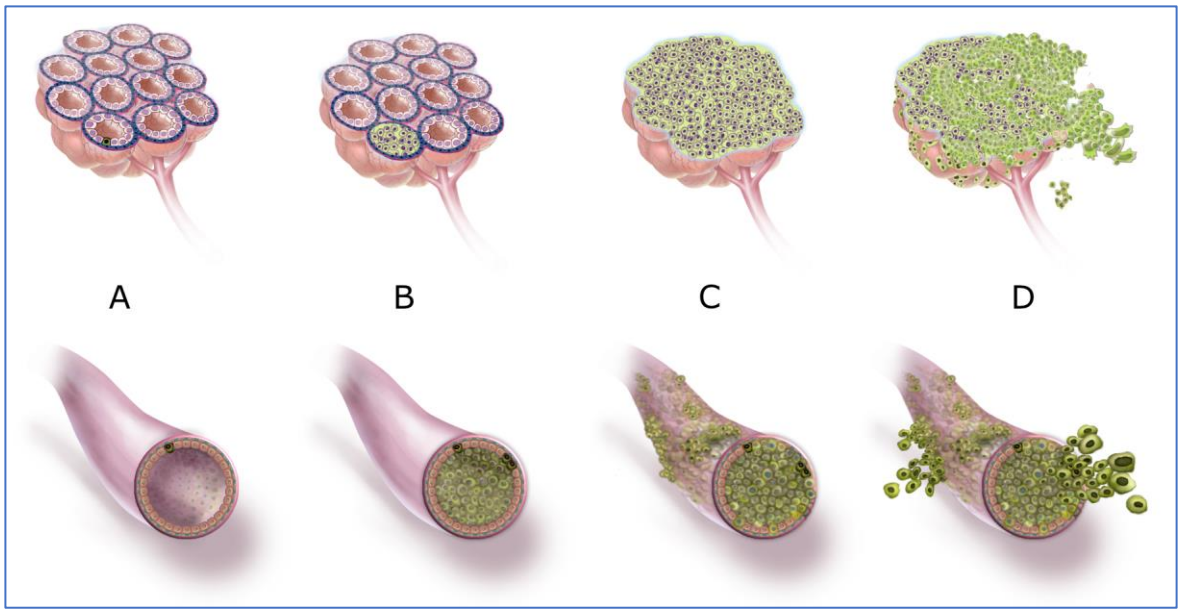

Figure 2: Tumor initiation and progression in breast lobules and ducts. A) Tumor initiation; B) LCIS and DCIS; C) Invasive lobular and ductal carcinoma; D) Metastatic lobular and ductal carcinoma. Figure modified from (8), with permission.

Breast cancer is the most frequent cancer for females worldwide, with approximately 2,1 million new cases per year (13). It is a heterogeneous group of diseases, with more than 20 different histological types (14) and four intrinsic molecular subtypes (15) (Table 1). Invasive ductal carcinoma (IDC) is the most common type of breast cancer, seen in 80% of cases (14). Although breast cancer can occur in the male population, it is quite rare, accounting for less than 1% of all breast cancer cases (16).

Advanced diagnostics tools have enabled earlier detection and classification in some cases. This has resulted in a better outlook for breast cancer patients, which has also been helped by improvements in managing these subtypes by identifying some of the key drivers in breast tumorigenesis and developing appropriate therapeutic options.

Knowledge about the implicated oncogenes and pathways paved the way to the development of targeted therapies that directly interfere with the oncoproteins of interest. Phenomenon known as "oncogene addiction" - where tumor maintenance is dependent on the action of a single or a few oncogenes (17), has allowed tailored therapies to be developed that inhibit the driver oncogene signaling.

Table 1: Four intrinsic molecular subtypes of breast cancer and their main general features. The table summarizes data reported by (18–22). Abbreviations are explained in the List of Abbreviations.

Molecular subtype	Receptor status	Ki-67 status	Histologic grade	Prognosis	Prevalence	10-year recurrence risk	Targeted therapy
Luminal A	ER+ and/or PR+, HER2-	low levels	low-grade	very good	~50-60 %	12,5 %	+
Luminal B	ER+ and/or PR+, HER2+/-	high levels	high-grade	poorer DFS	~40 %	23,4 %	+
Triple negative/basal-like	triple negative/ER, PR-, HER2+	high levels	high-grade	poor	~10-30 %	30%	+/-
HER2-enriched	ER-, PR-, HER2+	high levels	high-grade	poorer than luminal	~15-20 %	26,6 %	+

Certain features of different molecular subtypes in breast cancer have been exploited by developing drugs that target them specifically (Table 2). However, triple negative breast cancer (TNBC), a particularly aggressive subtype, remains a challenge, as it lacks the commonly targeted receptors. Even though epidermal growth factor receptor (EGFR) is often overexpressed in TNBC, existing therapies targeting EGFR were unsuccessful (23). Because of this, it is mainly treated by chemotherapy, and still yields poor outcomes. This leads to an urgent need to identify druggable targets such as poly ADP ribose polymerase (PARP) which proved successful in clinical trials (23).

With the exception of the TNBC subtype (where only a small subset of patients responds to currently available drugs against EGFR and PARP), targeted therapies show a great efficacy in clinical practices, usually leading to fast tumor regression. In breast cancer types dependent on estrogen, inhibition of tumor growth is achieved through application of selective estrogen receptor modulators (SERM) that obstruct estrogen action or by aromatase inhibitors (AI) that block the key enzyme in androgens biosynthesis, leading to reduction of estrogen levels (24). The Human epidermal growth factor receptor 2

(HER2) is a member of the HER family of transmembrane proteins that respond to extracellular ligands by dimerization and transphosphorylation of their intracellular domains, activating various signaling pathways in the cell (25). As these pathways are also involved in cell proliferation and survival, HER2 serves as a particularly strong target. Several therapies have been developed for HER2-enriched tumors (Figure 3) and have successfully led to dramatically improved prognoses, increasing the overall survival of Her2 positive patients (26). All these efforts led to improvements in breast cancer management, which is well reflected in the rise of the 5 year-relative survival rate, which currently reaches up to 89,7% (27). This is especially true for developed countries, such as the United States, where women who have survived breast cancer represent the majority of cancer survivors (28).

Table 2: Available targeted therapies in breast cancer. The table summarizes data reported by (23,24,29). Abbreviations are explained in the List of Abbreviations.

Breast cancer types according to receptor status	Percentage (in all breast cancer cases)	Target	Available targeted therapies
ER positive	75 %	ER	SERM (tamoxifen, raloxifene, lasofoxifene), AI (anastrozole, letrozole, exemestane)
HER-2 positive	25-30 %	HER2; HER and EGFR pathways	Recombinant monoclonal antibodies (trastuzumab, pertuzumab), dual tyrosine kinase inhibitor (lapatinib), conjugated monoclonal antibody TDM1
Triple negative	10-30 %	EGFR, PARP	EGFR TKI, anti-EGFR mAbs (effective for only a small subset of breast cancer patients), olaparib (the first approved PARP inhibitor)

Despite these successes and encouraging achievements, breast cancer still represents a significant health threat, as mortality rates remain high worldwide. According to the World Health Organization (WHO) report, more than 500 000 women died from breast cancer in 2018 (13). In underdeveloped countries, limited access to health care and adequate treatment largely contributes to this problem. However, another major cause is tumor heterogeneity and drug resistance, both of which pose a challenge even in the most advanced clinical settings.

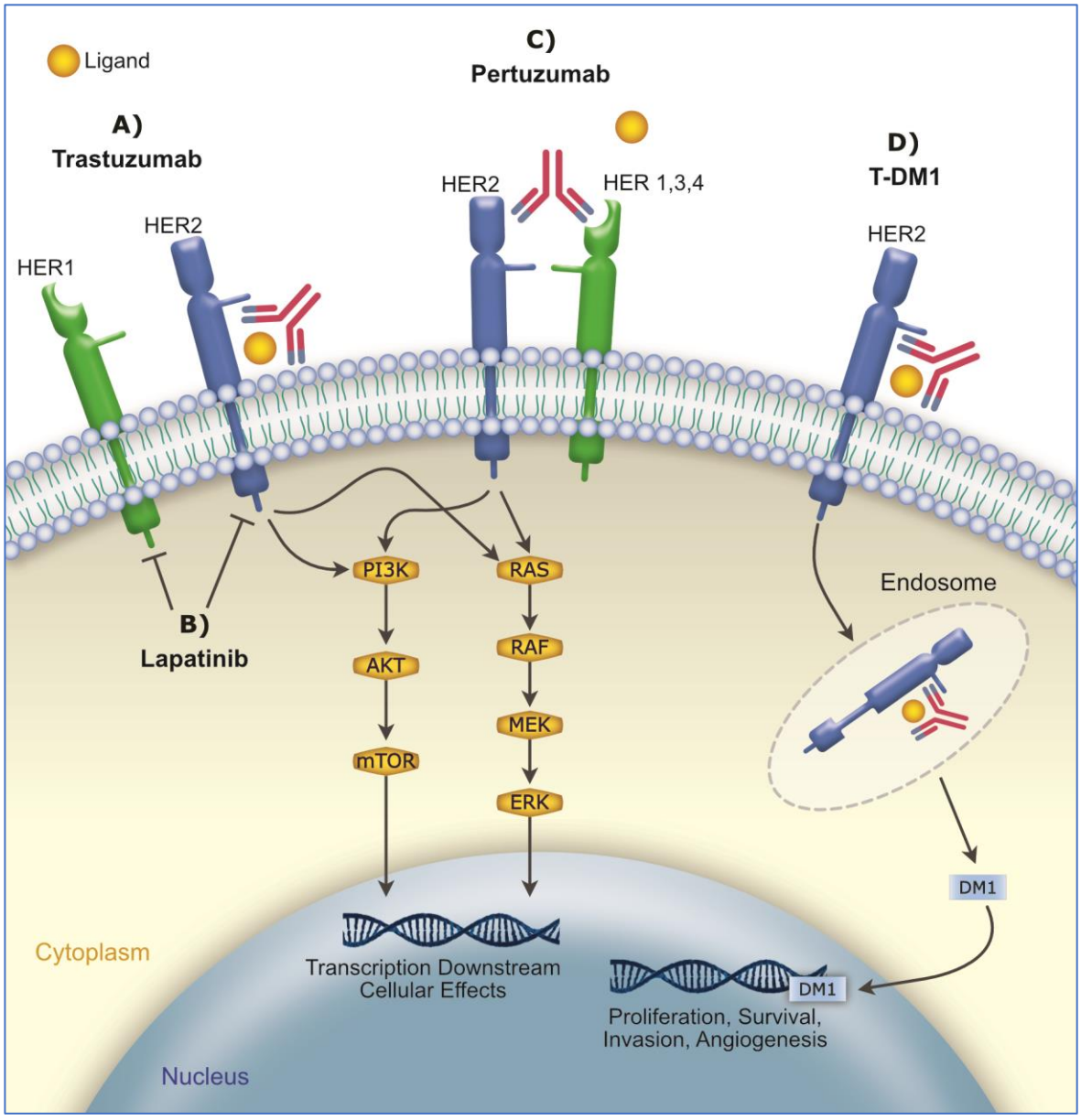

Figure 3: Therapies targeting HER2. A) Trastuzumab inhibits HER2 overexpression effects; B) Lapatinib is tyrosine kinase inhibitor interrupting HER2 and EGFR pathways; C) Pertuzumab inhibits HER2 dimerization with other HER receptors; D) T-DM1 (Trastuzumab emtansine) is a conjugated antibody containing trastuzumab and emtansine which is a cytotoxic agent. Figure modified from (30). Abbreviations are explained in the List of Abbreviation.

Ideally, the therapy of choice should eradicate all of the cancer cells, but this is often not the case. Some tumor cells can evade these treatments and survive in the body below detectable levels (Figure 4). This small population of surviving cells is referred to

as minimal residual disease and is one of the main hindrances in decreasing the breast cancer burden.

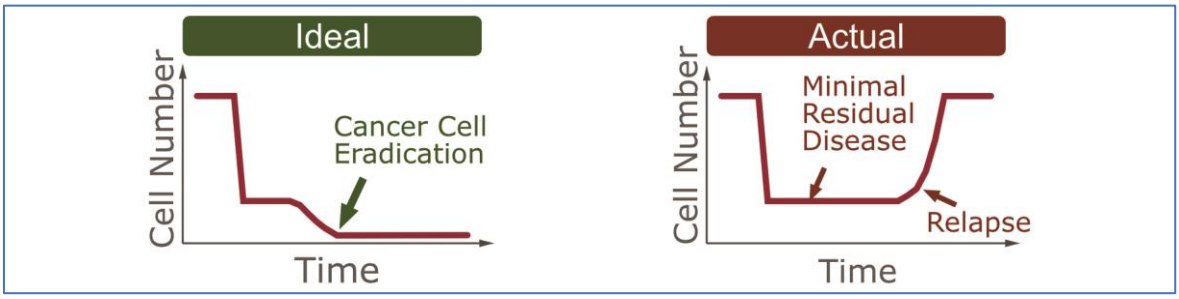

Figure 4: Model of ideal treatment compared to actual clinical outcome that leads to tumor regression. Left: eradication of all cancer cells by therapy; right: remaining residual cell population that can subsequently lead to a relapse. Figure modified from (31).

1.2 MINIMAL RESIDUAL DISEASE

The term "minimal residual disease", first coined in 1980s, refers to the cells that evaded or survived therapeutic intervention but are below detectable levels in the body. This makes the disease asymptomatic, appearing as though the patients have achieved complete remission (32). Conventional diagnostic methods and laboratory tests cannot detect the presence of the remaining cells (32). This surviving residual cell population poses a great threat as it can subsequently drive tumor recurrence which happens in up to 40% of the cases in breast cancer patients (33). A smaller percentage (10-20%) comes back as isolated locoregional relapses, while 60-70% appear as metastases in distant organs or multiple distant sites (33). Even though patients undergo surgery and receive adjuvant therapy in order to minimize the chances of recurrence, many of them still relapse and die (32). Recurrences can occur over a period from 2-3 years, when the risk is highest (33), up to decades (34) following initially successful therapy.

1.2.1 The three faces of minimal residual disease

Minimal residual disease can be present in three different forms (Figure 5):

- locally found residual cells,
- circulating tumor cells (CTCs),
- disseminated tumor cells (DTCs).

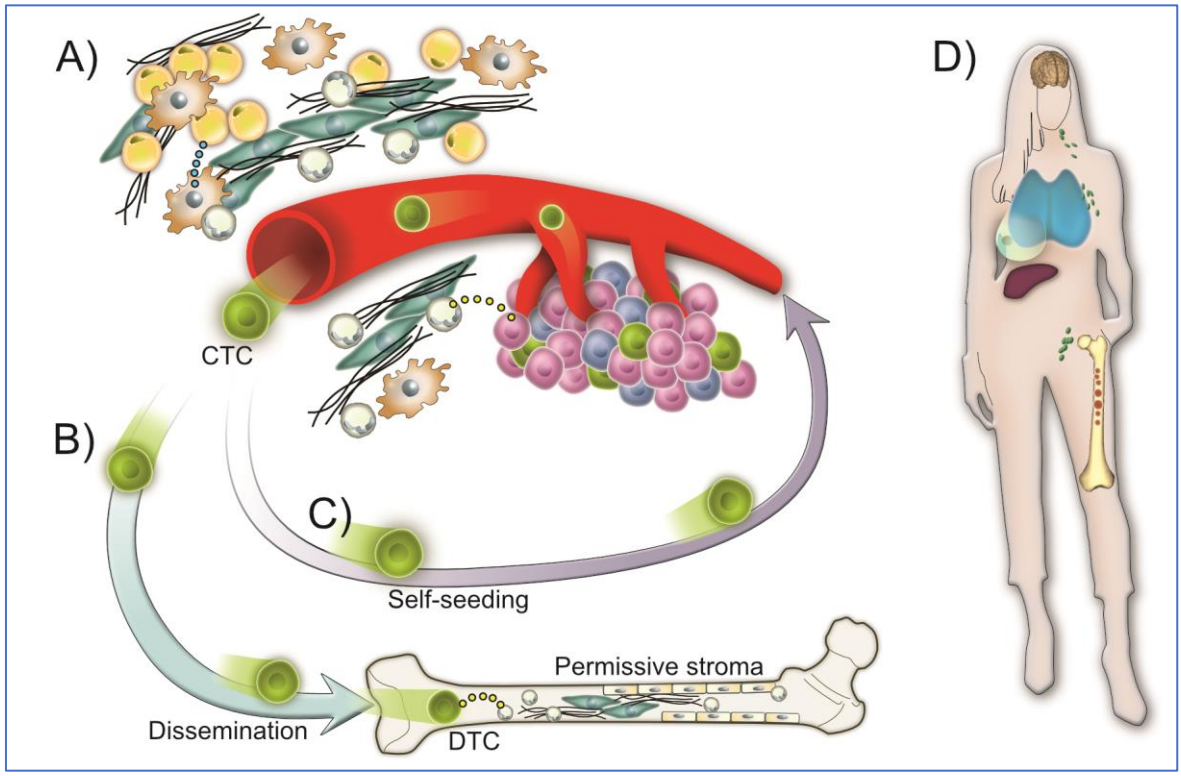

Figure 5: Three forms of minimal residual disease. A) CTCs; B) DTCs; C) Repopulation of the original tumor site; D) MRD can lead to local and regional relapse or metastatic cancer. Abbreviations are explained in the List of Abbreviation.

Locally present residual cells could be cells that were not successfully removed by surgery. This could be the case after breast-conserving surgery when margins were positive or too close; if the tumor was multi-focal or if the cancer cells already invaded deeply into pectoral fascia. However, even a whole mastectomy does not completely remove this risk, as at the time of surgery tumor cells could have been in the chest wall or already reached the regional lymph nodes and metastasized. Above mentioned factors and some other parameters (age, tumor size, grade, molecular subtype) are associated with the local recurrence (35). According to a newly proposed concept, local relapse could be also caused by the disseminated cells that repopulate the original site (36).

Circulating tumor cells are breast cancer cells found in the bloodstream. They shed away initially from the primary tumor or metastasize at later stages and enter the circulatory system (37). Even though a large amount of these cells are present in the patient body, around 10^6 CTCs per gram of primary tumor are detached and entering circulation on a daily basis (38). It is very difficult to detect them in the blood, because they are present in a very low number in the circulation, as the majority are eliminated by the action of

the immune system (39). When these cells manage to invade and infiltrate into other tissues, they are termed as DTCs. However, from the beginning of their travel through the circulation, up to the homing to other distant sites, cells escaping the primary tumor are coping with numerous stress conditions. Research in animal models showed that only a small fraction of these cells (2,5%) are able to survive and adapt to the new tissue microenvironment, while 0,01% could eventually trigger the formation of macro-metastases (37).

Disseminated tumor cells are found at distant sites and invade different organs. The most common site of metastases in breast cancer patients are bone and visceral organs such as the lungs, brain and liver (40). Metastases are incurable and "breast cancer thus remains the most common cause of cancer-related death in women." (33).

1.2.2 Mechanisms behind minimal residual disease

What are the selective advantages of residual cells? What gives them the ability to escape and evade treatment? Mechanisms driving MRD and subsequent tumor relapse remain largely unknown.

Different hypotheses have been established to attempt to explain the survival of these cells (Figure 6).

One of the proposed mechanisms of survival assumes the existence of "cancer stem cells" (CSCs) that are more resilient to therapy through drug-efflux pumps, epithelial-mesenchymal transition (EMT), quiescence and enhanced resistance to pro-apoptotic signals (41). Residual cells often express ATP-binding cassette (ABC) transporters (42) that can pump out many different compounds, which they use to reduce intracellular drug levels. Tumor cells can survive even targeted therapy, as other mutations can occur in the gene encoding the targeted oncoprotein, or alternative pathways emerge to circumvent therapy and enable survival (43). These secondary mutations can lead to substitution of amino acids in the kinase domain of these proteins, changing the structure of the targeted region and thus preventing the binding of the original drug, such as EGFR inhibitors. Another mechanism of resistance involves amplification of the target gene, such as dihydrofolate reductase gene (*DHFR*) which results in the production of its oncoproteins in an amount larger than are able to be targeted by methotrexate (44). Activation of new pathways are nicely illustrated by the example of acquired resistance to phosphatidylinositol 3-kinase (PI3K) inhibitors through the activation of the insulin-like growth factor 1 receptor (IGF1R) pathway (43). The role of epigenetic mechanisms in resistance is also being explored, as many epigenetic changes

take place in tumor cells, allowing them to evolve quickly and evade the therapy (45). For instance, it has been demonstrated that demethylation of ATP-binding cassette subfamily C member 1 gene (*ABCC1*) in cancer cell lines reduces intracellular drug levels (44). Polyploidy of residual cells has also been recently shown to play a role in chemoresistance and initiation of tumors *in vivo.* Namely, these cells avoid chemotherapy aimed at proliferating cells by infrequent cell divisions and they have also evolved budding and bursting as a special mechanism to produce their progeny which is more resilient to stressful conditions (46).

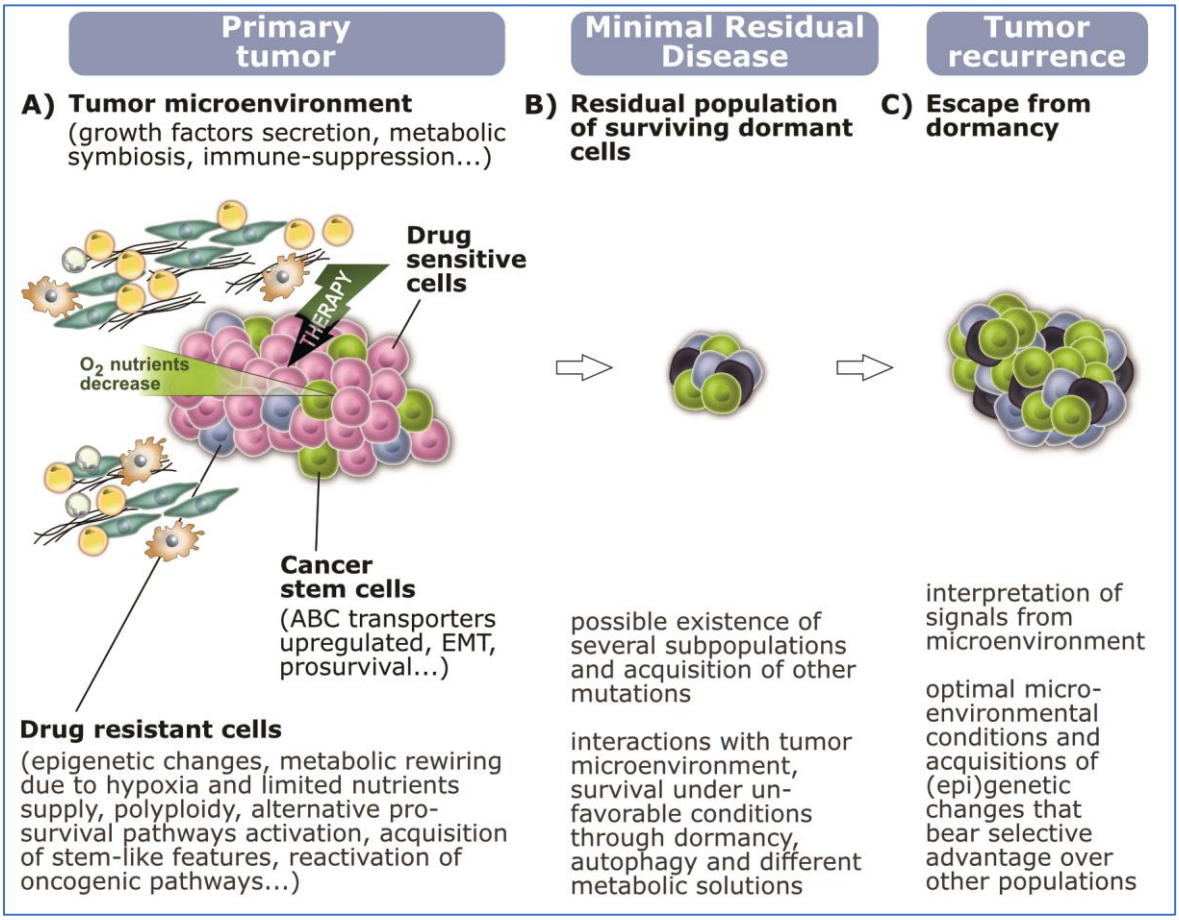

Figure 6: Mechanisms enabling survival of residual cells and tumor recurrence. A) Heterogeneity and adaptations in the primary tumor; B) Residual cells in dormancy; C) Awakening and progression to tumor recurrence. Abbreviations are explained in the List of Abbreviations.

Other studies have shown that residual cells survive through rewiring of their metabolic networks and switching to programs such as autophagy (47). For instance, resistance to a widely used chemotherapeutic drug, Gemcitabine was explored in hormone-

independent prostate cancer (HIPC) cell lines, where it was demonstrated that autophagy decreased sensitivity to the drug (48).

Studies in acute lymphoblastic leukemia (ALL) showed that dissociation of dormant cells from their environment make them exit this state, becoming proliferative and sensitive to therapy (49). Thus, a significant role in influencing the response of tumor cells to the chosen therapy is attributed to the tumor microenvironment (TME), which comprises a multitude of factors including: levels of available nutrients and oxygen, interactions with neighboring cells, presence of immune cells, proximity of blood vessels and different factors secreted from the environment and many others. For instance, certain factors secreted by stromal cells, such as the hepatocyte growth factor (HGF) can help tumor cells survive the drug treatment, as shown in case of BRAF-mutant melanoma, where stromal HGF re-activated MAPK and PI(3)K-AKT signaling pathways, leading to resistance to RAF inhibition (50). Microenvironmental conditions can induce multidrug resistance through different mechanisms, for instance by increasing expression of the epidermal growth factor-like domain 7 (EGFL7) protein and the amount of multi-drug resistance P-glycoprotein under hypoxia in non-small-cell lung carcinoma (NSCLC) cells (51). This can be also achieved by switching metabolic flux towards enhanced glycolysis to obtain energy, as shown in human breast and ovarian cancer cells (52). Through modulating the microenvironment and inducing immunosuppression (53), residual cells can escape from immune-surveillance, survive and eventually drive tumor recurrence.

"Self-seeding" is a paradigm by Comen and Norton (36) which proposes that cancer cells have the ability to move multi-directionally. Hence, they are competent to infiltrate not only distant and regional sites, but also return back to the primary tumor site. This was already experimentally shown in models of some carcinomas, such as breast, colon and melanoma (54). Disseminated cells are attracted by tissue specific factors, inflammatory cytokines, leaky vasculature and microenvironmental conditions that enabled their migration in the first place (36). This view also tries to explain the existence of small, but aggressive cancers, as even such barely detectable tumors can cause metastases as seen in many human malignancies (55). Thus, the cells coming from these cancers could be poor self-seeders, but good distant seeders (36). It would be of utmost importance to reveal specific gene signature of cells able to self-seed, as well as to populate other sites, as this could open a new therapeutic window aimed at the seeding ability of tumor cells.

The "seed and soil" hypothesis, originally proposed by Paget (56), attempts to explain the preference of CTCs/DTCs for certain organs. It assumes that these cells cannot survive microenvironmental conditions in other tissues. What directs these cells to the

preferred sites of metastasis? There have been reports that CTCs express certain proteins and chemokine receptors that drive them to preferred sites, while chemokines produced at the specific site of their dissemination also play a role in their survival in the newly invaded environment (37).

Although the gain of invasive, metastatic properties was before considered as a later event in tumor progression, a study in a mouse model of HER2-driven cancer found that about 80% of metastases were a result of the early dissemination (57). Early dissemination has been also seen in a mouse model of pancreatic cancer. Lineage tracing showed that some cells from pre-invasive pancreatic intraepithelial neoplasia (PanIN) lesions went through EMT acquiring a mesenchymal phenotype and stem cell features and disseminated into the liver (58). This ability of cells to migrate even in the early stages of primary tumor progression can explain tumor recurrences after removal of primary tumors that were small in size and appeared non-invasive, as well as CUP syndrome (cancer of unknown primary) which happens in 3-5% of all cancer cases (59).

The above-mentioned concepts aim to elucidate survival mechanisms of residual cells and have been probed in experimental models. However, they also lack to explain all the observed phenomena concerning MRD and disease progression as they are observed in the clinic (Table 3).

Considering tumor heterogeneity and the influence of tumor microenvironment, as well as the whole organism on a systemic level, it is most likely that all of the proposed mechanisms provide a plausible rationale in a specific context. Namely, while some of the proposed concepts might be true for certain populations of residual cells within a specific tumor subtype and under particular microenvironmental conditions, they might not be applicable for another. In addition, multiple mechanisms could be at work at the same time in a given tumor.

Nevertheless, one apparently universal feature of residual cells seems to be their dormant state. Before progressing to a clinically detectable tumor relapse/metastasis, residual cells can stay in dormancy for decades, such as is the case with dormant breast cancer cells in bone marrow (60).

Table 3: Some proposed concepts aiming to explain the mechanisms behind MRD. Abbreviations are explained in the List of Abbreviations.

Concept	Pros +	Cons -
Drug insensitivity	resistance to initially successful therapy, e.g. erlotinib/gefitinib in lung cancer patients (61)	subset of patients with relapse responding to repeated initial treatment (61)
Different levels of cell damage by the drug caused by its partial delivery	subset of patients with relapse responding to repeated initial treatment (61)	mouse model for BRCA1-deficient breast cancer showed surviving cells despite repeated treatments (61)
Existence of cancer stem cells which are more resistant to therapy	CSCs are found in small numbers in the bulk of tumor, and only a small subset of cells exhibit the ability to cause tumor development when transplanted into immunocompromised mice; CSCs possess escape mechanisms such as up-regulation of drug efflux pumps, increased DNA repair, quiescence, evasion of apoptosis, EMT; CSC model explains some cancers such as mouse acute myeloid leukemias (61)	not tested in many cancers and some of these features are also found in some other cancer resistant cells that are not considered CSC (61) according to the current criteria (established markers, tumor initiation); question whether transplanted CSCs can recapitulate the original tumor heterogeneity from which they come still remains, as (this is typically tested based only on a few surface markers); lack of interaction with immune system in immune-compromised mouse models; the CSC model potentially rationalizes a subset of cancers, not all (62)
Protection from tumor microenvironment	interaction of cells with the TME can be beneficial for their survival, e.g. secretion of HGF from surrounding stromal cells saves cancer cells from therapy (61)	even though the tumor cells are able to modulate TME to favor their growth and survival, they can also be eliminated through the action of the immune system or hostile conditions in the surrounding stroma (62)
Clonal evolution	could explain tumor heterogeneity and existence of several clones by gradual acquisition of genomic and epigenetic alterations (63)	some new models have challenged the idea of gradual changes in cancer cells - later mutations seem to play only a small part in providing selective advantage to tumor cells (64)

1.2.3 Tumor dormancy

Tumor dormancy is a phenomenon observed in clinical practices which refers to a temporary halt in tumor progression, with a prolonged latency (this results in a disease-free survival period after therapy and is manifested by remission of the primary tumor as well as absence of clinically detectable tumor recurrence) (65). Dormancy can extend over a long period of time followed by a relapse. In breast or prostate cancer patients, the recurrence can happen decades after initial treatment (66). Late recurrences have been also frequently observed in melanoma and renal cell carcinomas, 10-15 years after the treatment (67).

Tumor dormancy can refer to both tumor mass dormancy and cellular dormancy. Tumor mass dormancy happens due to the factors that constrain tumor 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 tumor cells (68). Another reason for tumor dormancy is a true mitotic arrest in the tumor cells, a phenomenon known as cellular dormancy (69). MRD is likely to be a result of both processes.

The drivers of cellular dormancy and the molecular entities that push cells in and out of dormancy are not yet fully identified. One of the mechanisms driving cellular dormancy could be senescence, a cellular program that becomes activated as a way to cope against oncogenic stress (66) in normal cells. Even though senescence signatures have been seen in neoplastic cells, it is not clear if this phenomenon could push cells towards relapse, since senescence was considered to be an irreversible process. However, it has recently been demonstrated that the process of senescence could also be reversed, leaving the possibility for its potential role in tumor recurrence (70).

Another more plausible mechanism behind the dormancy could be quiescence. This is characterized by arrest in GO-G1 phase, a state in which most adult cells are and a reversible feature of stem cells, or cells exposed to unfavorable microenvironmental conditions and temporary constraints (66).

Features of senescent and quiescent cells are summarized in Table 4.

Table 4: Differences between quiescence and senescence. Abbreviations are explained in the List of Abbreviations.

Dormancy program	Quiescence	Senescence
Cause	environmental stress and nutrient limitation (71)	replicative or oncogenic stress (71)
Metabolism	in general, low activity, with some exceptions such as quiescent primary human fibroblasts (71)	normal metabolic activity (71)
Reversibility	reversible (71)	mechanism-dependent (71)
Some studied markers	cyclin D1 downregulation (72); p27 and p21 upregulated or stabilized (73,74); CAF-1 strongly downregulated (75)	upregulation of p16 and senescence associated GLB1; p53 activation; senescence-associated heterochromatic foci (71)

1.2.4 Detection of minimal residual disease

In leukemia, certain markers of MRD have been established and some sensitive molecular tools have been developed enabling its detection through blood sampling. However, in solid tumors this is not the case. Lack of standardized methods and markers for reliable detection of MRD currently make it largely undetectable. To overcome the risk of residual disease and relapse, the majority of patients receive adjuvant therapy, but not all of the patients benefit from it (76) and some treatments are also associated with toxicity (77).

There have been some improvements in detection of MRD through isolation of CTCs and DTCs, although their abundance still represents a great hurdle. In breast cancer the ratio is 1 tumor cell in the background of 10^6 to 10^7 mononuclear cells (78). Still, a recent study by Beaver *et al*. (79) offers the exciting hope that MRD could be routinely monitored through plasma tumor DNA (ptDNA) which, albeit present in a very small fraction, can be detected with high specificity and sensitivity by droplet digital PCR (ddPCR) (80). Even though the method was applied in preliminary study on patients bearing PI3KCA mutations, it holds a great promise and hope for monitoring MRD.

1.2.5 Importance of studying minimal residual disease

MRD poses a great threat as remaining cells can lead to tumor recurrence in a form of local and regional relapse, or a distant metastasis. It is estimated, in the case of breast cancer, that around 30% of women diagnosed with the localized tumor (which corresponds to the earlier stages of the disease) and 75% of patients with lymph node involvement will eventually relapse (81). These recurrent tumors can be more aggressive and non-responsive to initial therapies, making these relapses life-threatening (82). This

highlights the importance of studying and understanding features of MRD in order to prevent recurrence by eradicating residual cells or keeping them dormant, following the patients' status, individualizing the treatments and choosing the best therapeutic options available.

In order to survive, tumor cells need to adapt to oncogene activity, new energy demands, constantly changing or new microenvironmental conditions. For this they need to adjust and rewire their metabolism all the time during the course of tumor induction and treatment. In this regard, identifying metabolic dependencies in residual cells and specific metabolic switches they rely on could reveal potential therapeutic targets and bring novel opportunities in the prevention and delay of tumor recurrence.

1.3 TARGETING CANCER METABOLISM – A NEW THERAPEUTIC OPPORTUNITY?

Since the pioneering work of Otto Warburg (83), metabolism has been explored in the context of cancer, as proliferating cells need to rewire their metabolism in order to meet their energy needs. A specialized cancer cell metabolism also ensures the production of sufficient building blocks for biomass increase and needs to respond to changing conditions and environmental challenges.

1.3.1 Altered metabolism as a hallmark of cancer

In their seminal review, Hanahan and Weinberg introduced the concept of cancer hallmarks (84), to better describe the main features that are crucial in the maintenance of a malignant phenotype. This attempt aimed to reduce the immense complexity of cancer when it comes to molecular heterogeneity of different subtypes/tissues and highlighted the most prominent phenotypic, biochemical and cellular interaction traits that distinguish a cancer tissue from normal cells. These properties are shared between many distinct cancer types, as the core molecular machinery of different mammalian cells types regulating these features is, in its essence, the same. Originally proposed cancer hallmarks were: "self-sufficiency in growth signals, insensitivity to growth-inhibitory (antigrowth) signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis and tissue invasion and metastasis" (84). As the body of knowledge on cancer grew, these hallmarks have been revisited and extended to several novel characteristics such as: tumor-promoting inflammation, evasion of immune response, genome instability and dysregulated cellular energetics (85). Changes in cellular energetics proved widespread in cancer. Although originally

thought of as a mere consequence of active proliferation-supporting oncogenic pathways, new discoveries of activating mutations in certain metabolic enzymes show that metabolic changes can initiate oncogenesis (86). Nowadays, altered metabolism has found its place among the accepted and well-studied cancer phenomena, establishing itself as one of the hallmarks of cancer, tightly interconnected with all the other processes that are taking place in a neoplastic cell (Figure 7).

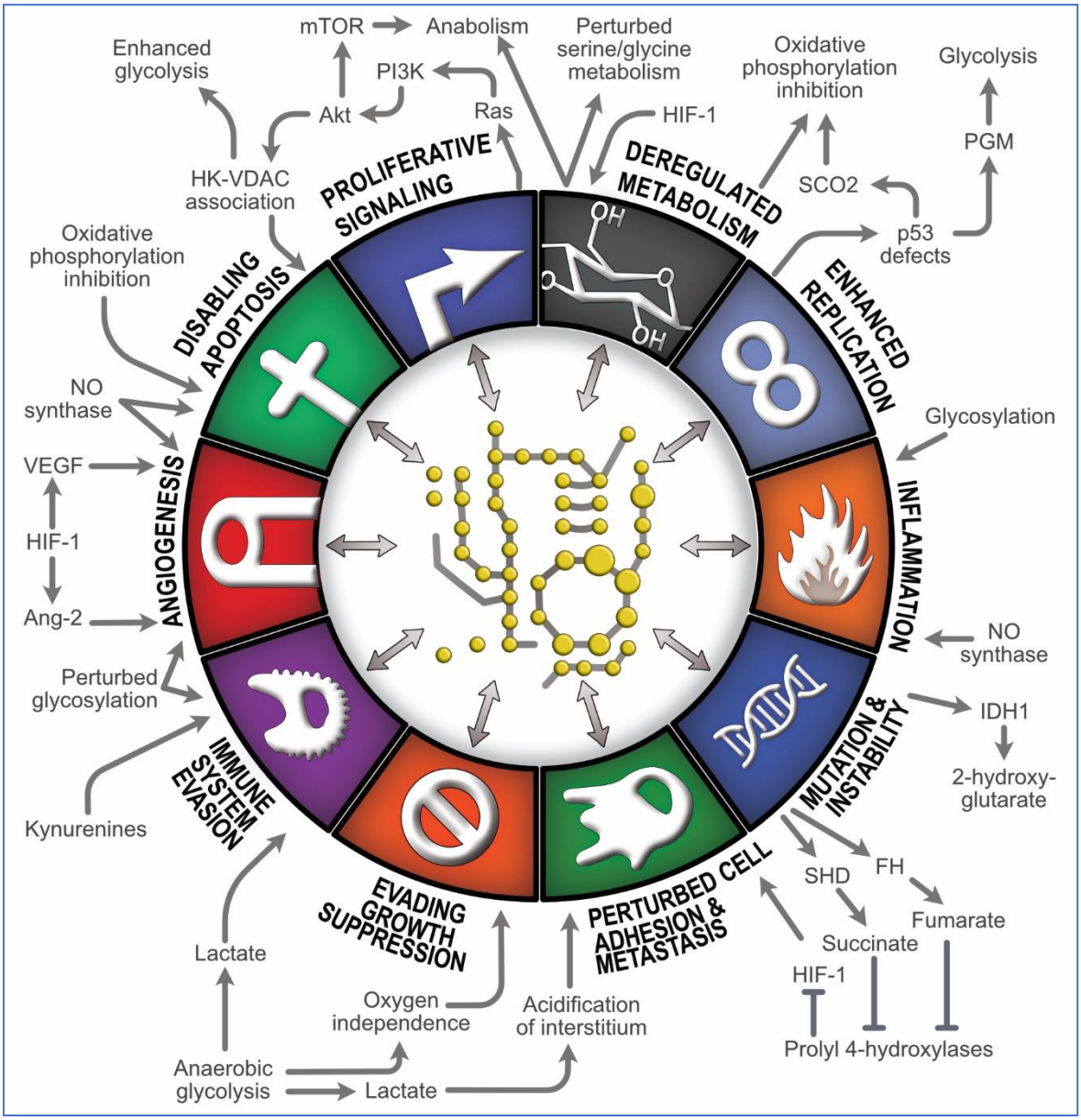

Figure 7: Metabolism as one of the hallmarks of cancer. Abbreviations are explained in the List of Abbreviations. Figure modified from (85,87,88), with permission.

Some of the examples of these relations are also seen from Figure 7. Namely, proliferation requires energy and material for biosynthesis, which is provided through rewiring of metabolic pathways that would meet this demand. Defects in p53 that enable enhanced replication also exhibit effects on energy metabolism, such as inhibition of oxidative phosphorylation (OXPHOS) and upregulation of glycolysis. Modifications of glycan structures in cancer promote cancer-related inflammation that is important in shaping the microenvironment and driving cancer progression. Perturbed glycosylation enables the evasion of immune response and promotes angiogenesis (growth of blood vessels for supplying tumor with oxygen and nutrients). Pathological levels of some metabolites such as fumarate and succinate can affect the activity of certain enzymes such as prolyl 4-hydroxylases, important in sensing cellular O_2. This can induce hypoxia, which affects cell adhesion, facilitating detachment and metastasis. Switch to glycolysis render the cancer cells independent of O_2, helping them evade tumor growth suppression. Excreted lactate enables the evasion of immune response.

Pavlova and Thompson further define six hallmarks of altered metabolism in cancer cells. These are:

- "… deregulated uptake of glucose and amino acids,
- use of opportunistic modes of nutrient acquisition,
- use of glycolysis/tricarboxylic acid (TCA) cycle intermediates for biosynthesis and NADPH production,
- increased demand for nitrogen,
- alterations in metabolite-driven gene regulation, and
- metabolic interactions with the microenvironment". (86)

Each of the above-mentioned six hallmarks enable metabolic adaptations of cancer cells and their survival, which results in the tumor growth, progression and invasion.

1.3.1.1 Deregulation in the uptake of nutrients

While division of normal cells and their uptake of metabolites is regulated by growth factors, proliferation of cancer cells is enabled through deregulation of signaling pathways that renders them independent (89). Nutrients might be scarce, but tumor cells are still able to find the means to use them and satisfy their demands by engaging alternative pathways. To meet the demands of proliferation, cancer cells (as well as

normal mammalian proliferating cells) display anabolic metabolism (89). As glucose and glutamine are the main components supporting biosynthesis and maintaining cellular redox homeostasis, their uptake is particularly increased; however, many other metabolites such as methionine, cysteine, choline, essential fatty acids and micronutrients, play various roles in tumor maintenance and progression and their function is yet to be understood (86). While glucose is the main molecule for catabolism and anabolism of carbohydrates, glutamine is used for *de novo* synthesis of proteins and nucleotides or non-essential amino acids via glutamate (90). Glutamate can enter the TCA cycle as α-ketoglutarate and thus support OXPHOS and lipid biosynthesis, generate NADPH and replenish intermediates (90). Through conversion into glutathione, glutamate also plays role in dealing with the oxidative stress (90). Since glutamine plays a role in so many processes supporting cancer cells, its uptake is particularly increased in many cancers, which is often termed "glutamine addiction".

1.3.1.2 Acquisition of nutrients by engaging alternative opportunistic pathways

As tumor cells proliferate and the growing tumor mass gets further away from the blood vessels, supply with oxygen and nutrients might become critical. Cancer cells also experience hypoxia and nutrient scarcity in the new microenvironment where they metastasize to. Certain therapies, modulation in the diet, calorie restriction and changes on the systemic level in the organism can lead to a lack of certain compounds important in tumor metabolism. In order to survive starvation periods and to overcome these unfavorable conditions, cancer cells engage alternative mechanisms that enable them to seize the necessary material and energy from different, normally unexpected sources. These opportunistic modes are illustrated by micropinocytosis of extracellular macromolecules and their further degradation, entosis of entire cells, phagocytosis of apoptotic bodies and self-catabolism through autophagy (86). Cancer cells also engage both arms of pentose phosphate pathway (PPP) – oxidative (used by normal cells) and non-oxidative, to ensure components for nucleotides when nutrients are scarce. Further, cancer cells manage oxidative stress and induce angiogenesis (91), building a route for obtaining their supplies, which leads to further tumor growth. In order to ensure lipid supply, they also often enhance *de novo* fatty acid synthesis (91). Another pathway frequently increased to provide material for biosynthesis and energy is fatty acid oxidation (FAO) that yields acetyl CoA, NADH, FADH and ATP (91). This whole spectrum of adaptations and diverse metabolic solution is made possible through the action of certain oncogenes and their impact on different cellular processes such as actin cytoskeleton remodeling or capability of cells to invade the cytoplasm of other neighboring cells (86).

1.3.1.3 Generation of important intermediates through engagement of glycolysis and TCA cycle

While non-proliferating, differentiated cells switch to enhanced lactate production only in anaerobic conditions, cancer cell increase lactate levels irrespective of the O_2 availability which is a phenomenon known as the Warburg effect or aerobic glycolysis (89). Even though less effective in terms of energy production, proliferative cancer cells favor this metabolic pathway probably because ATP itself is not a limiting factor most of the time and the cancer cells' need for reducing equivalents and building blocks originating from glycolytic intermediates overcomes their need for a greater ATP yield (89). In order to compensate for the lower efficacy of ATP production, cancer cells overexpress glucose transporters to increase glucose uptake (91). Overexpression is also seen in many glycolytic enzymes. In addition, cancer cells often engage usage of alternative enzyme isoforms, such as pyruvate kinase muscle isozyme M2 (PKM2), embryonic isoform of pyruvate kinase which is frequently expressed and responsible for shunting glycolytic metabolites for biosynthesis of glycine and serine (83). Metabolic intermediates coming from glycolysis are used for anabolic processes such as biosynthesis of nucleotides, amino-acids and biological macromolecules. Shunting metabolites to the PPP also results in production of NADPH which is the major reducing agent used in maintenance of cellular redox homeostasis (91).

Besides ensuring a faster metabolic rate and proliferation, glycolysis can also provide a fuel for mitochondrial OXPHOS through excess lactate (89). Lactate secreted in the microenvironment could be consumed by the cancer cells that are near blood vessels and have oxidative metabolism (83). Contrary to the misconception that cancer cells do not use mitochondrial OXPHOS, emerging evidence shows that a significant portion of ATP is being produced through this pathway (92) – in some cancer cell lines 50 to 70% (83). In other cell lines, only 10% of ATP is produced this way (83) and mitochondrial enzymes are dominantly utilized for generating precursors that would be used in biosynthetic processes (92). This is done via the TCA cycle which provides intermediates that are important for biosynthesis, such as: oxaloacetate that generates aspartate further involved in pyrimidine synthesis, citrate for lipid biosynthesis and acetylation and α-ketoglutarate for biosynthesis of amino acids, nucleotides, hexosamines and glutathione (93). In addition, enhanced mitochondrial activity ensures control over $Ca2^+$ concentration and overall cellular redox homeostasis (91). Increased glucose uptake and glycolysis, as well as glucose carbon flux into anabolic pathways require functional mitochondria. This increase is often triggered through activation of the PI3K/Akt pathway that is a very frequent event in human cancers (92).

1.3.1.4 Increased requirement for nitrogen

Besides carbon, cancer cells exhibit an increased requirement for nitrogen for their biosynthetic processes. Nitrogen is released when amino acids are broken down, as a result of protein catabolism. The main source of reduced nitrogen is glutamine, which is also rate-limiting factor, since it is required for progression of the cell cycle and utilized in the biosynthesis of nucleotides and non-essential amino acids (86). Nitrogen in the cell is often present in the form of ammonia, which was recently discovered to be more than just a waste product. Namely, breast cancer cells are able to recycle ammonia into central amino acid metabolism, supporting the growth of tumor biomass (94). The interest in the role of nitrogen is growing and, while alterations in nitrogen metabolism in cancer are still not largely explored, recent studies highlight its importance for cancer progression, mainly through production of polyamines (86) or dysregulation in urea cycle enzymes (95). The latter has been found to correlate with tumor invasiveness and reduced survival of the patients (95).

1.3.1.5 Changes in gene expression through metabolic networks

Metabolic reprogramming is a result of oncogenic signaling, but at the same time it can drive oncogenic changes, mainly through an action of regulatory enzymes that can further shape the cell's epigenetic profile and thus affect gene expression (86). Some of the known metabolites involved in epigenetic alterations are (86):

- cytosolic acetyl-CoA – has a role in histone acetylation, as a substrate of histone acetylases that renders the genome more accessible for transcriptional enzymes;
- crotonyl-CoA – a substrate of p300 histone acetyltransferase that activates gene expression;
- S-adenosylmethionine (SAM) – has a role in methylation, as a donor of methyl groups;
- NAD+ – has a role in de-acetylation of histones and non-histone proteins, as a cofactor of sirtuins (class of deacetylases);
- α-ketoglutarate – has a role in post-translational modifications, through α-ketoglutarate-dependent dioxygenases;
- succinate and fumarate – have an inhibitory effect on α-ketoglutarate-dependent dioxygenases.

1.3.1.6 Metabolic reprogramming through dynamic interaction with the tumor microenvironment

Tumor cells have the ability to change their microenvironment and use it to meet their needs, fueling the tumor growth. For instance, they can ensure survival through this type of "metabolic symbiosis" by consuming the products of metabolically altered stroma. Tumor cells exhibit effects on normal neighboring cells, surrounding stromal cells and vasculature, as well as the components of the immune system by reprogramming them through secretion of growth factors and dynamic changes of the extracellular matrix (ECM) (86). A well-known consequence of an increased glucose and glutamine uptake and lactate secretion by cancer cells is the acidification of the microenvironment that results in immunosuppression, generation of the new blood vessels (angiogenesis) and tumor invasion (86). Another example of microenvironmental effects on the tumor concerns stromal cells that had lost caveolin-1 (*CAV1*); these cells enhance glycolysis and secrete the metabolites that cancer cells consume as a fuel for their mitochondrial metabolism, which is a phenomenon known as a "reverse Warburg effect" (96). Similarly, molecules secreted from cancer associated fibroblasts (CAFs) as well as physical barriers of a rich stroma in pancreatic ductal adenocarcinoma (PDAC) contribute to chemoresistance (97).

The relationship between cancer cells and their environment is bidirectional. Metabolic stress can induce dormancy in cancer cells in different ways: through chronic hypoxia (98), cell cannibalism (99) and increased activity of certain pathways that enhance compensatory metabolic routes (100). This enables the survival in a hostile environment. Hypoxic conditions, through activating hypoxia-inducible factor 1-alpha (HIF1-α), can also lead to EMT. EMT results in the loss of polarity and cell-cell adhesion, giving the tumor invasive properties (101). Conditions in their microenvironment thus also affect tumor cells, challenging them to find the best metabolic solution, shaping their metabolic network (86).

Previously described metabolic changes – from increased nutrient uptake, over engagement of opportunistic pathways, generation of important intermediates through upregulation of key pathways, and metabolic crosstalk with the environment reflect a metabolic plasticity of tumor cells. Extensive metabolic reprogramming that takes part in cancer cells gives them a selective advantage and enables their survival in harsh environmental conditions, eventually driving tumor progression through invasion of surrounding tissue and metastasis.

1.3.2 Metabolic reprogramming in cancer – more than just an epiphenomenon of oncogenic signaling and uncontrolled proliferation?

Traditionally, cancer has been thought of as a disorder of aberrant, uncontrolled proliferation. In addition, as the role of altered metabolic pathways behind many oncogenic processes is becoming apparent, this view is expanding and metabolic aspects of this disease start to receive attention (83).

Cancer metabolism is characterized by increased glucose uptake and glycolysis, lactate production and export, enhanced PPP, upregulated fatty acid synthesis and lipid metabolism, higher glutamine uptake and glutaminolysis, mitochondrial biogenesis, FAO increase, alterations in the TCA cycle, reactions in electron-transport chain (ETC) and an increase in overall anabolic processes for generating energy and supporting biomass production. A global metabolic landscape of cancer cells is illustrated in Figure 8.

Overexpressed oncogenes (gain of function) or missing tumor suppressors (loss of function) affect not only cell cycle/proliferation, but also induce changes in metabolism through aberrant signaling (Figure 9). For instance, MYC, HIF-1α and AKT upregulate the expression of key glycolytic enzymes enhancing glycolysis, while p53 exhibits the opposite effect by inhibiting the expression of glucose transporters.

V-myc avian myelocytomatosis viral oncogene homolog (*c-MYC* or *MYC*) is a particularly important proto-oncogene, a master regulator of metabolism. This gene encodes a transcription factor (TF) which, through regulation of microRNAs (miRNAs) and thousands of genes, controls many cellular processes, integrating them with cellular metabolism (102). Besides driving cell proliferation, oncogenic *MYC* shapes the cancer metabolic network by increasing glucose uptake, glycolysis and mitochondrial biogenesis, glutaminolysis and other anabolic pathways (Figure 10). The MYC protein is also involved in sensing the cells nutrient status and response to hypoxic conditions, as hypoxia reduces its levels, turning the switch from mitochondrial metabolism towards glycolysis (102).

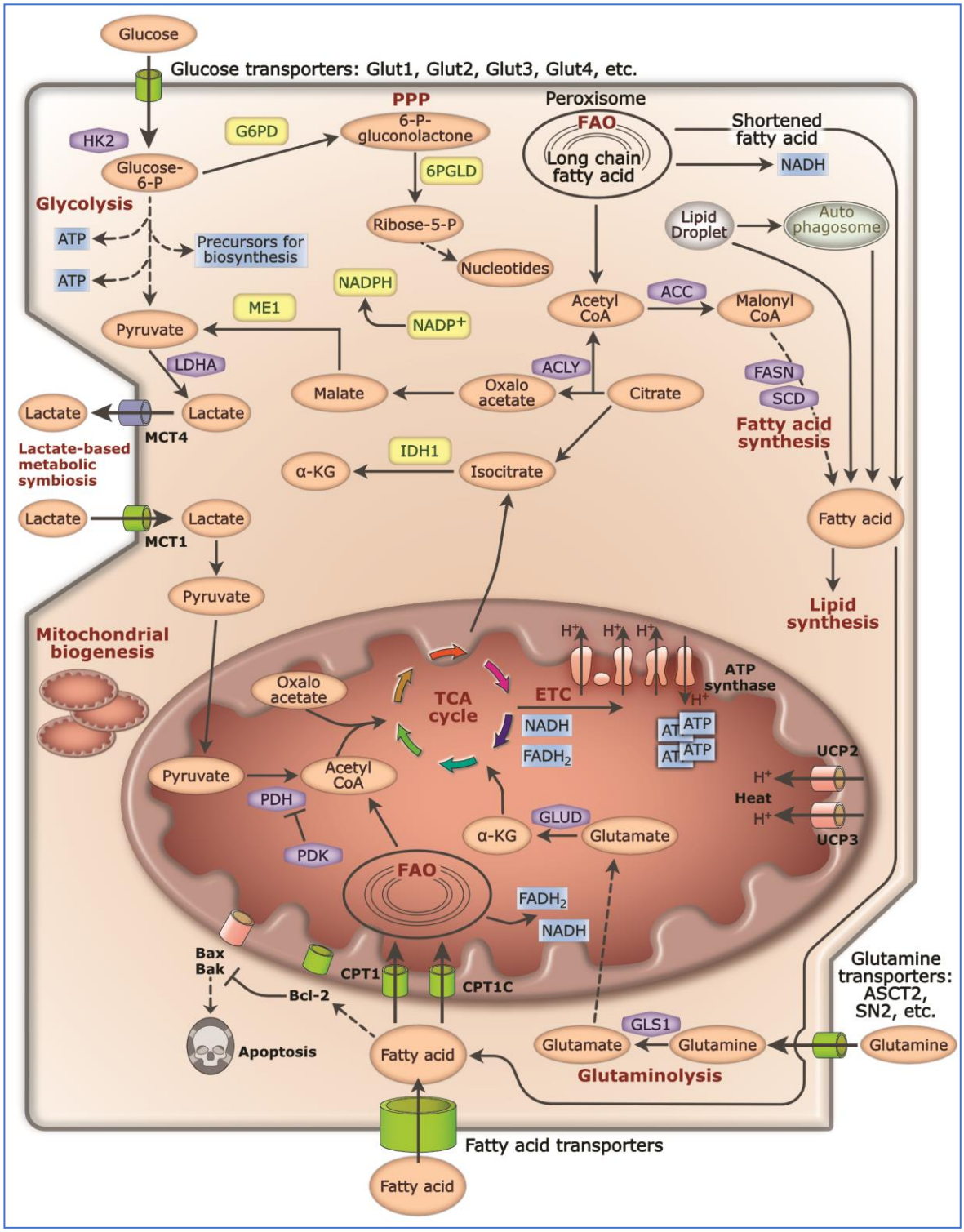

Figure 8: Global metabolic landscape of cancer cell with key metabolic changes summarized. Figure modified from (91). Abbreviations are explained in the List of Abbreviations.

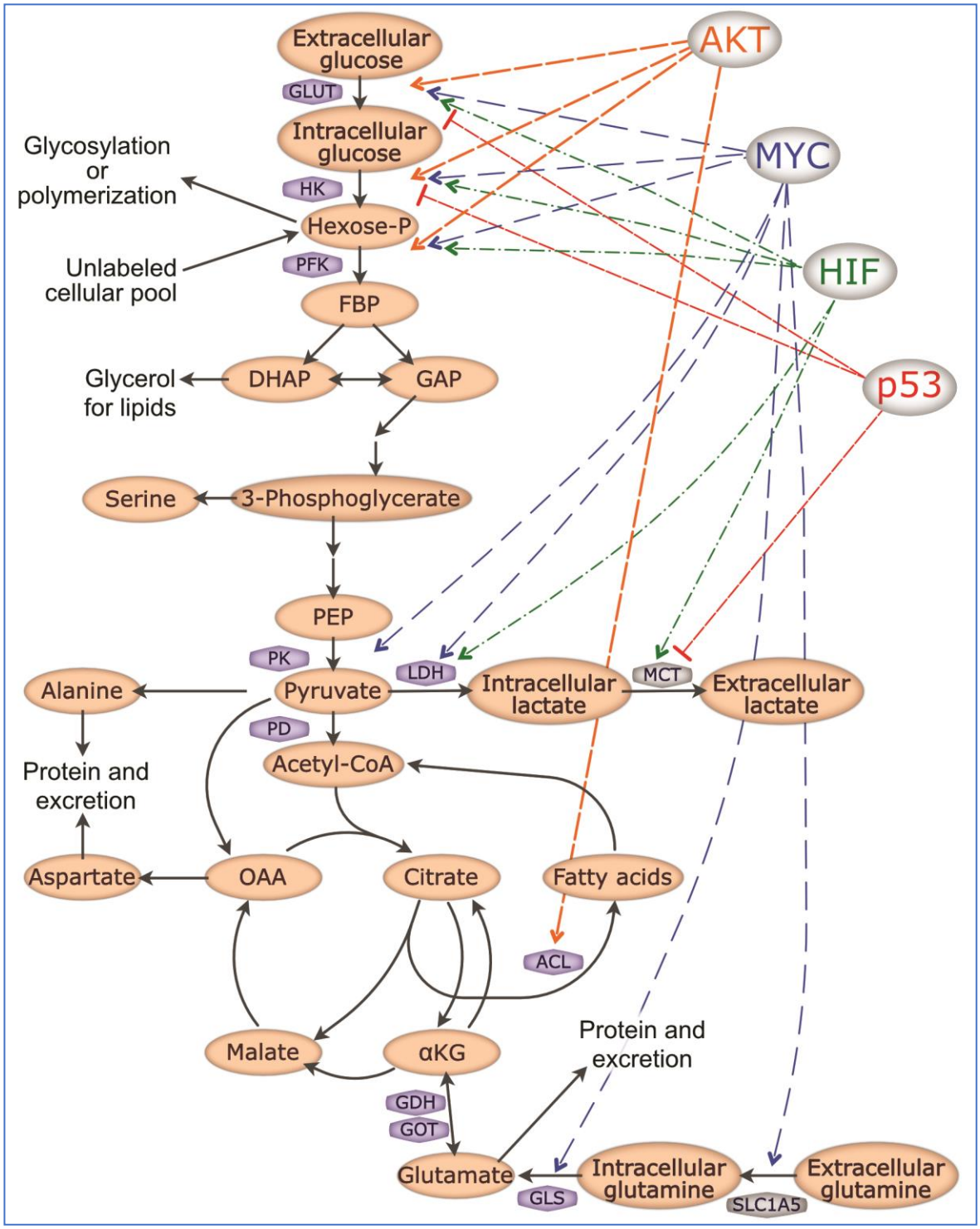

Figure 9: Changes in central carbon metabolism resulting from dysregulated oncogenes (AKT, MYC, HIF) and tumor suppressors (p53). Activation of certain enzymes in indicated in arrows, inhibition in lines. Figure modified from (83). Abbreviations explained in the List of Abbreviations.

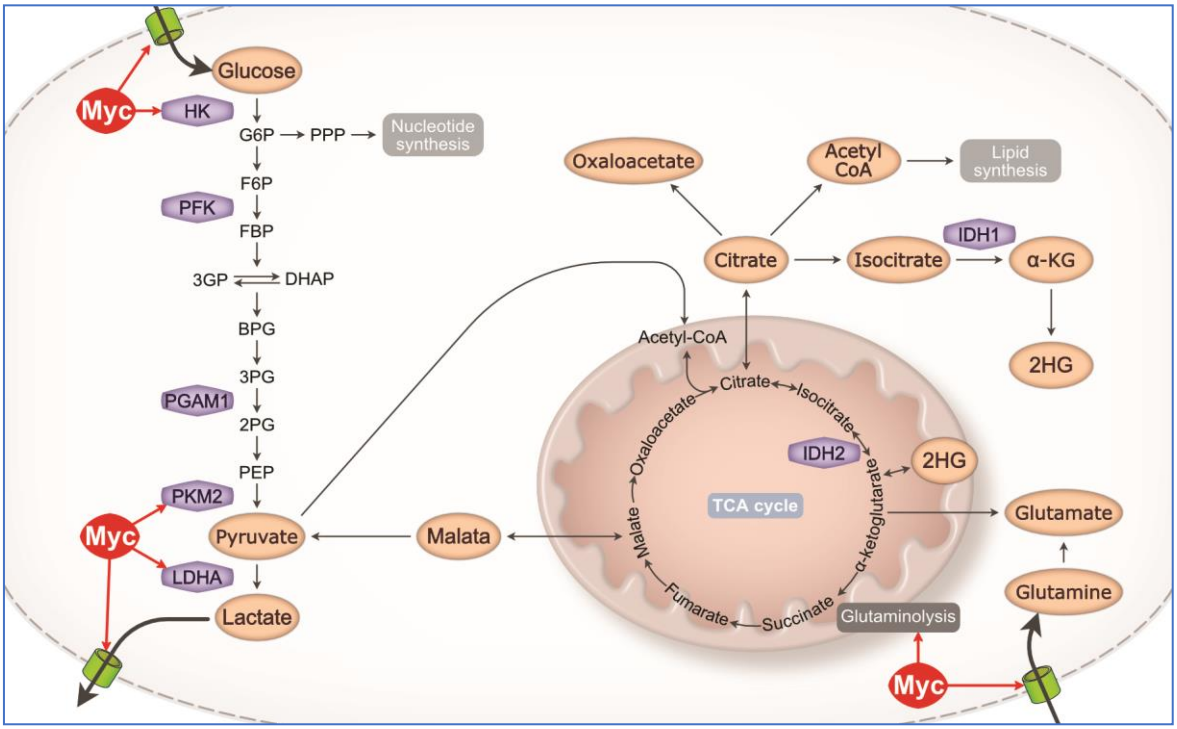

Figure 10: Metabolic pathways under regulation of MYC in cancer metabolism. Figure modified from (103). Abbreviations are explained in the List of Abbreviations.

Nevertheless, cancer metabolism is not just an epiphenomenon, a result of altered gene expression and aberrant oncogenic pathways, but an active player in oncogenesis (104). Metabolites are substrates or cofactors of enzymes catalyzing chemical modifications of chromatin. Changes in the abundance of cellular metabolites also affect nutrient-sensitive posttranslational modifications (104). Oncometabolites, such as 2-hydroxyglutarate (2HG), change gene expression through their effect on chromatin organization (104). The role of altered metabolic pathways and effect of changes in abundance of metabolites in modulating and shaping the cells' epigenetic profile is being acknowledged and intensively explored. Discovery of oncometabolites resulting from mutations in genes encoding enzymes in some cancer types such as glioma, demonstrated the active role of altered metabolic pathways in tumorigenesis and cancer progression. One such example are mutations in isocitrate dehydrogenases (*IDH1* and *IDH2*). In normal cells, normal variants of IDH1 and IDH2 catalyze reactions producing NADPH and α-ketoglutarate, which have a prominent role in many metabolic processes. Instead of α-ketoglutarate (α-KG), mutant forms of IDH1 and IDH2 catalyze the reaction that converts it to its structural analog 2-hydroxyglutarate (2HG) which is only present in low concentrations in normal cells (below detection), while it accumulates in tumor cells (104). 2HG is considered an oncometabolite because it binds to the enzymes that

normally require α-KG as a cofactor and thus inhibits their action. Since many of these enzymes are involved in shaping an epigenetic landscape, 2HG drives epigenetic changes through inhibition of Jumonji-C domain histone demethylases (JHDMs) and TET (Ten-eleven translocation) enzymes involved in DNA demethylation (104). A similar phenomenon is seen in other cancer types where mutations in succinate dehydrogenase (SDH) and fumarate hydratase (FH) lead to the pathological accumulation of these two metabolites. Fumarate and succinate thus become oncometabolites that also inhibit α-KG-dependent enzymes, further driving a range of other effects such as promoting HIF1 action and causing epigenetic alterations (104).

Products of metabolism previously considered to be a cellular toxic waste, keep emerging as players in tumor progression. One of them is ammonia, recently shown to support tumor growth in breast cancer (94). Lactate is another metabolite that plays an important part in tumor progression, mainly through inducing changes in the tumor microenvironment through acidification or serving as a fuel for a subset of cancer cells that appear to be addicted to mitochondrial metabolism (104).

Besides glycolysis and mitochondrial metabolism, which have been pathways of great interest in cancer research since the discovery of the Warburg effect, other metabolic pathways and their implications in tumor maintenance are being comprehensively explored. One such example is one-carbon metabolism which is connected with the main anabolic pathways and, along with affecting metabolic processes and redox homeostasis, has been linked to epigenetic changes in tumor cells (104).

Epigenetic modifications such as methylation and acetylation are influenced by metabolic changes (Figure 11), as the abundance of methyl donors, glucose or acetyl CoA can have an effect on their regulation and thus impact gene expression (104). Similarly, altered expression levels of deacetylases which are dependent on NAD^+ (sirtuins) have been demonstrated in many diseases, including cancer where loss of certain sirtuins promoted neoplastic transformation through various mechanisms. Low levels of certain sirtuins have been associated with increased DNA damage or ineffective DNA repair. In particular, sirtuin 1 (*SIRT1*) downregulation has been found in breast cancer, especially associated with breast cancer 1, early onset gene (*BRCA1)* mutations (104).

While oncogenic signaling affects and regulates many metabolic pathways, the opposite is also true. Dynamic exchange between the cancer cell and their microenvironment, their exposure to external cues and factors of the surrounding stroma and physiological conditions affecting the cellular metabolic state result in the changes of metabolite abundance. This subsequently alters the cells epigenetic landscape and gene expression

patterns. Together, oncogenic signaling and perturbed metabolism further shape the cancer cells' phenotype. Thus, altered metabolism is not just a mere byproduct, but a prominent player in tumorigenesis worth exploring as a potent source of many therapeutic targets.

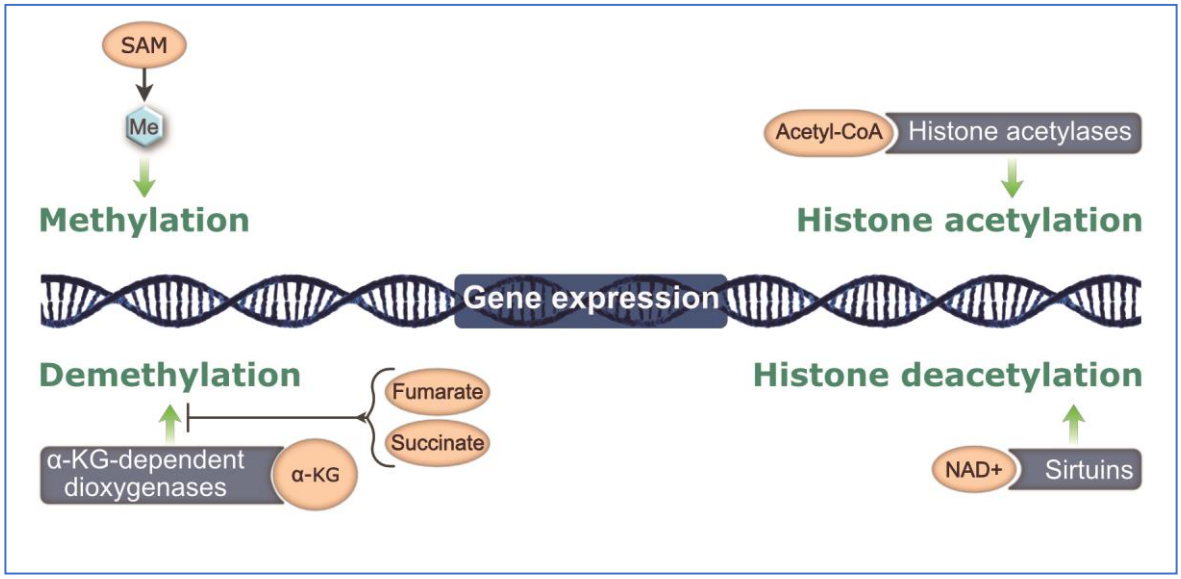

Figure 11: Epigenetic control as a link between metabolism and gene expression profiles in cancer. Abbreviations are explained in the List of Abbreviations.

1.3.3 Metabolic changes – clinical implications and therapeutic targets

Metabolic profiling of cancer cells holds a promise for new approaches in treating cancer. Its exploration is aimed at identifying novel biomarkers for various cancer subtypes and the development of new metabolic drugs. With the improvement of existing metabolomics techniques for the precise and fast detection of a broad range of metabolites, the body of knowledge in the field of cancer metabolism is growing. This knowledge has already found some practical application in a clinical setting. One of the fine examples is 2-[^{18}F] fluoro-2-deoxy-D-glucose positron-emission tomography (FDG-PET) scan – an imaging technique that detects tumors which exhibit the Warburg effect employing radiolabeled glucose analogues used to visualize the process of an increased glucose uptake, revealing the tumor sites (105). Non-invasive techniques such as magnetic resonance imaging (MRI) and magnetic resonance spectroscopy imaging (MRSI) have been used in diagnosis of cancers as well as monitoring and assessment of the treatment (105). Application of MRI and MRSI led to significant improvements in breast cancer management – while mammography and MRI displayed better sensitivity

and accuracy compared to mammography alone, detection of choline by MRSI displayed a 100% sensitivity distinguishing between benign and malignant breast tumors (105).

Besides metabolic biomarkers used in imaging, there are many metabolites that could be measured in the samples which could be obtained in a relatively non-invasive and easy way. Metabolites could be measured in breath, saliva, serum, plasma and urine. These markers could be used in monitoring cancer progression and assessing the cancer stage, as they have been found to correlate with cancer aggressiveness (105). For example, three phosphatidylinositols were found to be reduced in the plasma of pancreatic cancer patients; lysophosphatidylcholines were decreased in the blood of lung and liver cancer patients; lysophosphatidic acid was elevated in ovarian cancer (105). Employing metabolic biomarkers in clinical practice has a lot of potential, but requires optimization, reproducibility and standardization of specific and sensitive laboratory tests.

Observed changes in cancer metabolism are also finding their way to a clinical setting, in terms of extensive searching for the novel targets that could be hit with a specific therapy while sparing the normal cells. Even though tumor heterogeneity exists also on metabolic level, it is still a more efficient way compared to targeting diverse genetic features – there are certain common metabolic dependencies seen in many tumors. One of the common metabolic pathways upregulated in cancer cells is glycolysis, so glycolytic enzymes and transporters emerged as an interesting option for targeting. This is especially true for glucose transporters and hexokinases (enzymes that catalyze the first step of glycolysis) as they are frequently over-expressed and studies on several mouse models on different cancer types show that their inhibition slows the tumor growth (90). To target 6-phosphofructo-1-kinase (PFK1) which is often upregulated in cancer and is an enzyme in the upper part of the glycolytic pathway, small inhibitors of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase isoform 3 (PFKFB3) are being developed, as PFKFB3 inhibition prevents the formation of a potent allosteric activator of PFK1 and slows down cancer proliferation (90). Interfering with the end part of glycolysis through pyruvate dehydrogenase kinase (PDHK) inhibition increases pyruvate dehydrogenase (PDH) activity, disrupting aerobic glycolysis and leading to cell death *in vitro* and a better outlook in animal models (90). PDHK is frequently overexpressed in cancer cells and its inhibitor, dichloroacetate (DCA), is being intensively explored in the clinics (90).

Increased uptake of glutamine is another property of cancer cells, particularly of the ones driven by oncogenic MYC. Glutamine transporters: solute carrier family 1 member 5 (SLC1A5) and large neutral amino acids transporter small subunit 1 (LAT1) are

frequently overexpressed and their inhibition slowed tumor growth and induced autophagy; decreased proliferation and cell death was also shown when glutaminase (GLS) was inhibited (90). Another interesting target is glutamine dehydrogenase (GDH) whose inhibition blocks the entrance of glutamine in the TCA cycle and thus prevents OXPHOS, replenishment of important intermediates and lipid synthesis – however, there is a need for more specific inhibitors (90).

Certain success has also been seen in targeting some parts of a deregulated TCA cycle – a small molecule inhibitor aimed at mutant IDH reversed the cancer phenotype towards a more normal one by reducing 2HG generation (90).

Altered lipid metabolism in cancer has also received attention – targeting *de novo* fatty acid synthesis through interference with fatty acid synthase (FASN) using available obesity drug Orlistat or by utilizing inhibitors of ATP citrate lyase (ACLY) represents another therapeutic opportunity (90).

Despite the challenges imposed by metabolic flexibility and plasticity of cancer cells, potential side-effects, suboptimal specificity and toxicity, drugs targeting metabolism are a subject of extensive research and hold a promise for more efficient treatments as new and improved generations of therapeutics are being developed. Identifying metabolic specificities of residual cells that could be targeted, while leaving metabolism of normal cells functional, would be a crucial step towards finding a therapeutic window in treating MRD and preventing tumor recurrence.

1.3.4 Exploring metabolic characteristics of MRD

Many specific metabolic traits and adaptations of tumor cells are being revealed and cancer metabolism emerges as an important player driving oncogenesis. But what is known about metabolism of dormant residual cells? How similar it is to the primary tumor? How different it is from the normal cell?

Studying this population still poses a challenge due to our inability to follow their fate and evolution in patients. Cells coming from relapses/metastatic disease already exited dormancy and could have acquired additional genetic insults, as most metastases appear to be driven by mutations different than the ones maintaining the primary tumor (106,107).

One of the ways to study residual cell populations is through samples obtained after neoadjuvant treatments whose purpose is to make some locally advanced tumors operable. The treatment leads to tumor regression which is often incomplete – a resistant

cell population survives, representing MRD (108). Cells from the primary tumor that entered the circulation (CTCs), sampled after neoadjuvant therapy in breast cancer patients, showed cancer-stem like features associated with resistance (109). Cancer stem cells (CSCs) or CSC-like – tumor initiating cells (TICs) are the cells that have been implicated in aggressiveness of cancer and its poor prognosis (110,111) and thus assumed to be responsible for drug-resistance and subsequent tumor recurrence. Being stem cells, they are able of self-renewal dividing asymmetrically, yet they are slow-cycling (112). Because of their restricted proliferation and differentiation, they are usually refractory to standard chemotherapy aimed at proliferating cells. Alterations in metabolic programs emerge as important in the survival and resistance to therapy (113). As many parallels can be drawn between tumor dormancy and CSC models (114), studying the metabolic characteristics of CSCs/CSC-like dormant cell population provides important cues about metabolic properties of MRD.

Scientists are employing additional ways and novel platforms for studying the residual, dormant population. Three-dimensional cell cultures with matrix of different physical properties that regulate dormancy (115), *in vitro* model of tumor dormancy and recurrence after short-term chemotherapy that enriches for therapy resistant, slow-cycling cells (116) and platforms that induce dormancy through encapsulation of the cells in the special nondegradable gels (117) are some of the tools that will expand our knowledge about MRD.

Another interesting opportunity to closely follow tumor regression and the surviving residual population comes from the inducible mouse models and employment of the corresponding 3D cultures which enable detailed insights in these cells. They have been shown to drive relapse and thus represent a correlate of MRD (118,119). So far, our previous study on the residual cells in breast cancer model revealed enhanced fatty acid metabolism as a striking feature associated with oxidative stress that renders them susceptible to DNA mutation and subsequent relapse, that was also verified in MRD samples obtained from different cancer subtypes (119). Further exploration of residual cells on a metabolic level could reveal additional interesting metabolic targets or biomarkers of clinical importance in prevention of tumor relapses and metastatic spread.

Chapter 2

OBJECTIVES

OBJECTIVES

The current knowledge on MRD in breast cancer is limited, as most samples available from clinics are either primary tumors or relapses, where many intricate changes have already taken place. Studying MRD in breast cancer is limited by the availability of patient samples and by the difficulty to obtain residual cells, which leaves them largely unexplored.

It is clear that knowledge of this elusive cell population would be valuable, as its features could potentially be exploited as targets for better interference against tumor recurrences. **The overall goal of this thesis is to characterize this cell population, employing a tractable mouse model of MRD.** Understanding **the metabolic traits of MRD** would be particularly important, as this could provide valuable information on the potential metabolic dependencies and vulnerabilities of the residual cells.

2.1 HYPOTHESIS

Since altered metabolism is one of the hallmarks of cancer, it is likely to be of great importance in the context of MRD. One oncogenic insult can lead to the dysregulation of multiple pathways. Metabolic reprogramming in cancer is an extensive and profound process that happens as a consequence of oncogenic signaling and defects in the function of tumor suppressors, but also as a result of the exposure of cancer cells to different cues and conditions and their dynamic interaction with the environment. Altered metabolism is not just a byproduct of these processes, but emerges as an active player that could drive further changes in tumor progression, leading the cells to evolve new features. Changes in gene expression and metabolic pathways during tumor progression could eventually leave a fundamental imprint on the cellular metabolic network, which would remain altered despite the inactivation of the driver oncogenes. These alterations could have a vast, intense and lasting effect on the rewiring of the cancer cell.

Metabolic reprogramming of cancer cells could have a more permanent effect on the biochemical phenotype of the cell, persisting as a metabolic memory in the dormant residual cells.

Identifying metabolic features specific for the residual cell population could bring us closer to the information on the phenotype of these cells as it would reflect their biochemical characteristics. Narrowing the immense tumor heterogeneity to a finite number of metabolic solutions could pave a way towards new targets and therapeutic opportunities in the prevention of tumor recurrences.

OBJECTIVES

In this thesis, the following questions will be answered:

- What are the overall characteristics of surviving residual cells following oncogene inactivation?
- Is there the lasting metabolic imprint following extensive metabolic rewiring during tumorigenesis on the residual cells after oncogene inactivation? Are there oncogenic remaining features preserved in the absence of oncogenic signaling?
- Can we identify specific metabolic features – potential metabolic targets of the residual cell population, that set them apart from healthy cells?
- Can we identify a time-point during tumor progression where irreversible changes take place; find the "point of no return"?
- What could be a mechanism behind potential metabolic memory?

2.2 AIMS

The knowledge on residual cells would be obtained through the following aims:

- in depth characterization of a 3D culture system that yields a correlate of MRD;
- a multi-omics approach and set up of metabolic methods to obtain the metabolic landscape of the residual cell population;
- verification of the most prominent findings in the mouse models *in vivo* and in correlation with datasets from breast cancer patients;
- following the kinetics of oncogenesis and regression as a mean to identify a potential time-period when the "metabolic memory" is established.

Chapter 3

RESULTS

RESULTS

CONTRIBUTIONS

This thesis contains experiments performed in collaboration with the groups of Dr. Kiran Patil, Dr. Theodore Alexandrov (both Structural and Computational Biology Unit, EMBL Heidelberg) and the group of Prof. Britta Brügger (BZH Heidelberg, University of Heidelberg).

I designed the experiments, optimized and characterized the 3D culture system. I genotyped the animals and performed *in vitro* experiments: phenotype characterization through bright-field and confocal microscopy, immunofluorescence, Real-Time qPCR, Western blot, RNA collections and extractions over timepoints during oncogenesis and regression, DNA and protein collections and extractions, intracellular and extracellular metabolite collections, extractions and runs on GCMS and Q-Exactive MS; collections of the extracellular and intracellular samples for the shotgun lipidomics. I did the conceptual integration of the transcriptomics, lipidomics and metabolomics data, Over Representation Analysis (ORA), identification and selection of the most prominent targets. I validated selected targets and most prominent findings through *in vivo* experiments: maintained the mouse colony, collected mammary gland tissues, performed IHC and IF stainings on FFPE tissue sections, performed and ran metabolic experiments *in vivo* (GCMS metabolic measurements, measurements of ^{13}C glucose labelling, NOS and BCA assays). I analyzed NOS assay data and the *in vivo* fluorescence staining data at the Tissue-gnostic microscope with the help of Dr. Yuanyuan Chen (group of Dr. Sotillo, DKFZ Heidelberg). I set up the methods for IMS on mammary gland tissues and performed the cryo-sectioning, prepared and ran the samples for IMS.

Eleni Kafkia (Patil group, Structural and Computational Biology Unit, EMBL Heidelberg) set up metabolic protocols, optimized GCMS for data acquisition and analyzed GCMS data, contributing also to its interpretation. Eleni contributed also to harvesting and collecting of the *in vivo* metabolite samples, and NOS and BCA assays experiments.

Katharina Zirngibl (Patil group, Structural and Computational Biology Unit, EMBL Heidelberg) did bioinformatic analysis and integration of RNA-sequencing data (GO enrichment, KEGG reactions and metabolite predictions), as well as correlation of mouse and patients' (publicly available) datasets.

Ashna Alladin (Jechlinger group, Cell Biology and Biophysics Unit, EMBL Heidelberg) contributed to the optimization of TetO-Myc/TetO-Neu/MMTV-rtTA system and its

characterization: doxycycline titration experiments, IF stainings, collection and extraction of RNA during oncogenesis and regression. Ashna Alladin, Savannah Jackson, Matthew Boucher and Marta Garcia Montero further maintaining the mouse colony and contributed to collecting the tissues to be used for *in vivo* analysis.

Sylwia Gawrzak analyzed the data for the Kaplan-Meier curve.

Marta Garcia Montero provided technical assistance with histology.

Federico Villa contributed to the *in vitro* setup of metabolic methods out of the 3D culture matrix.

Dr. Daniel Sevin (GSK Cellzome, Heidelberg) ran and analyzed Q-Exactive MS data from untargeted metabolomics experiments performed at Cellzome, GSK Heidelberg.

Christian Lüchtenborg (Brügger group, BZH Heidelberg, University of Heidleberg) performed lipid extractions from collected samples and shotgun lipidomics together with data analysis on these experiments.

RNA sequencing and bisulfite-free DNA sequencing were performed at the Genomic Core Facility at EMBL. Andrew Palmer (Alexandrov group, Structural and Computational Biology Unit, EMBL Heidelberg) assisted to setting up the methods for IMS on mammary gland tissues.

Dr. Martin Jechlinger supervised the project and designed experiments.

Manuscripts in preparation:

I am co-first author of the first manuscript in preparation, along with the co-first authors Eleni Kafkia and Katharina Zirngibl; with Ashna Alladin, Federico Villa, Marta Garcia Montero, Savannah Jade Jackson, Daniel Sevin, Christian Lüchtenborg, Britta Brügger, Kiran Raosaheb Patil and Martin Jechlinger: In depth multi-omics analysis reveals an oncogenic metabolic memory in a surviving cell population in a breast cancer model of minimal residual disease (in preparation).

The second manuscript in preparation will deal with the kinetics of oncogenesis and tumor regression – highlighting the point of no return from oncogene addiction. I will be co-first author of this manuscript, along with the co-first authors Ashna Alladin, Eleni Kafkia and yet to be determined contributing scientists that will include the lead authors Kiran Raosaheb Patil and Martin Jechlinger.

3.1 INDUCIBLE MOUSE MODELS AND THREE-DIMENSIONAL PRIMARY CELL CULTURES FOR STUDYING MINIMAL RESIDUAL DISEASE IN BREAST CANCER

Mouse models have been of a great importance in studying many human disease related processes, helping us to elucidate various aspects of human biology and establishing themselves as a valuable tool in research, biotechnology and the pharmaceutic industry. Conditional mouse models, commonly used in studies exploring cancer, represent a leap forward in better modeling of human malignancies. In particular, tetracycline-inducible mouse models bring us closer to the patients' situation, as they allow for temporal control of oncogene expression. Oncogenesis is induced in the adulthood in case of breast cancer as modeled disease.

Experimental design and approaches to study MRD in this project are outlined in Figure 12. We explored MRD in breast cancer utilizing inducible mouse models and primary mouse mammary cells grown in 3D cultures. To fully characterize MRD, we optimized and set up methods for a 3D culture system that enabled us to follow it in depth: from observation of phenotypes to the application of a multi-omics approach. To test the hypothesis of metabolic memory and to get a detailed, full picture on the features of MRD, a systems biology approach was employed, integrating obtained information from different layers of biological system (transcriptomic, metabolomic, lipidomic and methylome profiles). Key features of MRD observed *in vitro* were verified in mouse models *in vivo* and then correlated with available human datasets of MRD.

As a model of MRD, we employed *TetO-cMYC/TetO-Neu/MMTV-rtTA* inducible mice that have two transgenes – oncogenes that play an important role in breast tumorigenesis: *c-MYC (MYC)* which contains human exon 2 and 3 (120) and activated rat *Neu* (121). *MYC* is a gene that encodes MYC protein which is a transcription factor and a master regulator of metabolism (122). It is also a potent proto-oncogene implicated in many tumor types. While it represents a primary oncogene in some human cancers (for instance, Burkitt's lymphoma), 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 (122). *Neu* (*Her2*) encodes HER2 – human epidermal growth factor receptor that plays a role in regulation of many cellular processes, mainly cell proliferation, differentiation and survival (123). *Her2* gene overexpression is seen in around 15-30% of primary breast cancer cases (121). When these two oncogenes are overexpressed together, which is a common case in breast cancer, prognosis is particularly poor and tumors are aggressive (124).

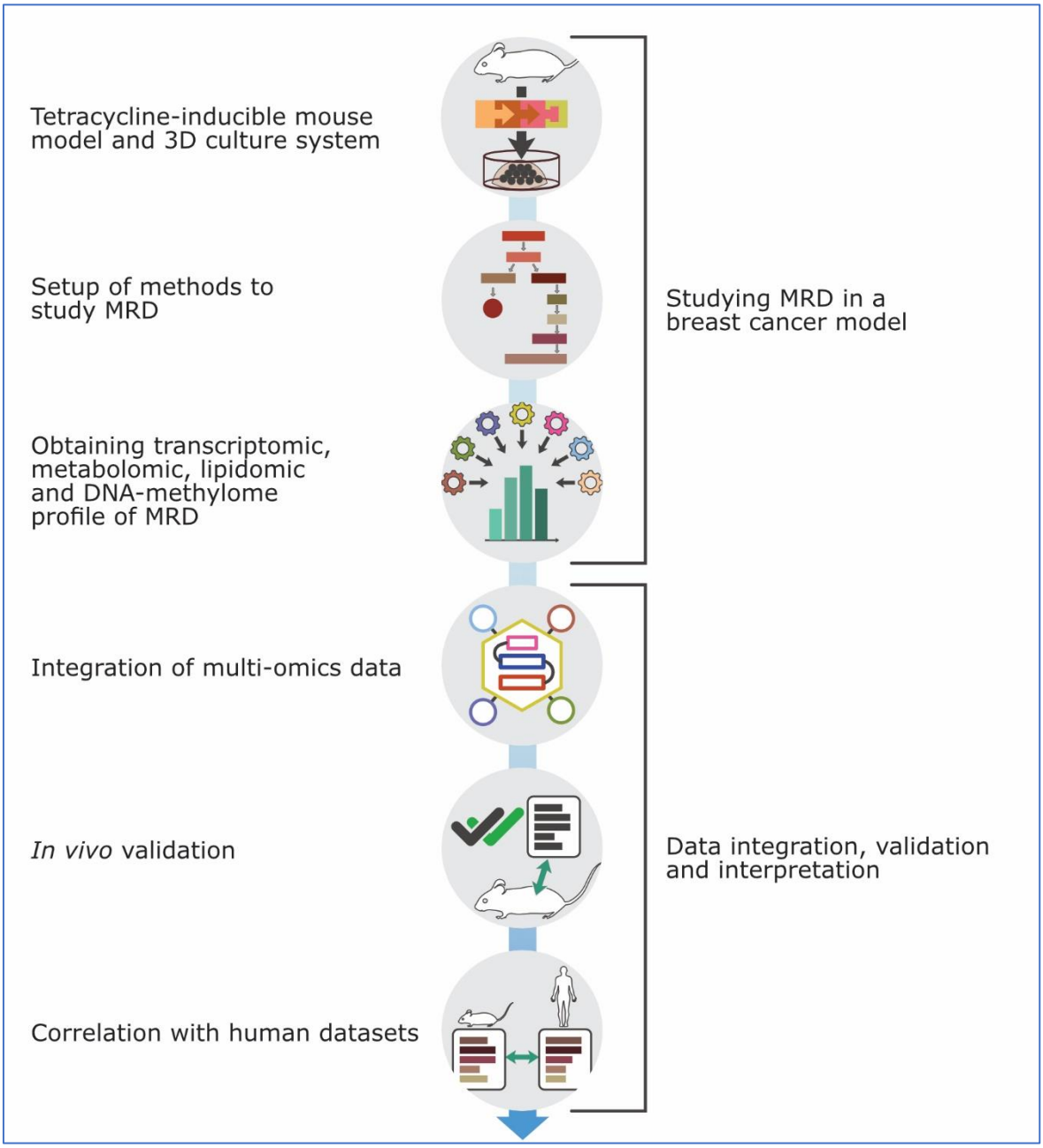

Figure 12: Experimental design to study MRD

Mouse mammary tumor virus long terminal repeat (MMTV-LTR) sequence enables spatial control of these transgenes; their specific expression in the mammary gland (120). Reverse tetracycline-dependent transcriptional activator (rtTA) regulates temporal expression of the oncogenes which are under control of the tetracycline-dependent minimal promoter (120). The system is regulated by doxycycline (antibiotic from

tetracycline class). Upon doxycycline (Dox) administration, rtTA binds to tetracycline operator (TetO) sequences, driving oncogene transcription and inducing tumorigenesis (120) (Figure 13A). Doxycycline removal from the system results in very fast inactivation of oncogenes. They are "switched off"; their expression is not detectable anymore (120,121). The system mimics the phenomenon of oncogene addiction, where tumor maintenance is dependent on the action of one or a few oncogenes (17). Switching off the oncogenes by doxycycline withdrawal resembles perfect targeted therapy aimed at the driver oncogenes.

In order to follow the complex dynamics of tumorigenesis and regression, as well as to model MRD, we employed a 3D culture system, obtained from the mouse mammary epithelial cells, as previously established in our group (118). Mammary glands from adult virgin female mice were dissected and dissociated to single cells placed in 2D collagen-coated plates for selection of epithelial cells (Figure 13B). Epithelial cells were then seeded in 3D spheres of Matrigel-collagen mixture which represented the elements of the microenvironment and enabled the growth of the cells in physiological, 3D conditions, so they formed basic mammary gland ductal units – acini. Each acinus consisted of luminal epithelial cells, surrounding the lumen. Figure 13C shows a schematic representation of phenotypes followed in 3D culture. Normal hollow acini (termed "NI" from "never induced" as they were not exposed to doxycycline and thus oncogenic transgenes were never expressed) represent a healthy duct. When doxycycline (Dox) was introduced through the cell media, *MYC* and *Neu* oncogenes were activated and cells began to proliferate uncontrollably. After 5 days, the lumens were filled and these structures (termed "tumor" or designated "DOX") acted as an experimental tumor correlate. From the previous research in our group, it is known that, upon switching off the oncogenes, the majority of these cells die yet the surviving rim remains (119). These residual structures (designated "OFF") represent a model of MRD, as some of the cells bear tumorigenic potential and were able to drive relapse (118,119). Studying MRD in this model first required its characterization and optimization, as well as the set-up of new protocols for obtaining different types of material from 3D cell cultures (Figure 13D).

This clear and synchronous culture represents a reductionist, clean system without noise and complexity one would encounter *in vivo*, enabling also more controlled experiments and an overall shorter time-frame for oncogenesis and regression. An additional advantage of our 3D culture system lies in the fact that, while recapitulating 3D cellular organization *in vivo* and interactions with extracellular matrix components, we can still closely follow the structures, monitor their phenotypes throughout different stages of oncogenesis and regression, and collect the material for various types of analyses.

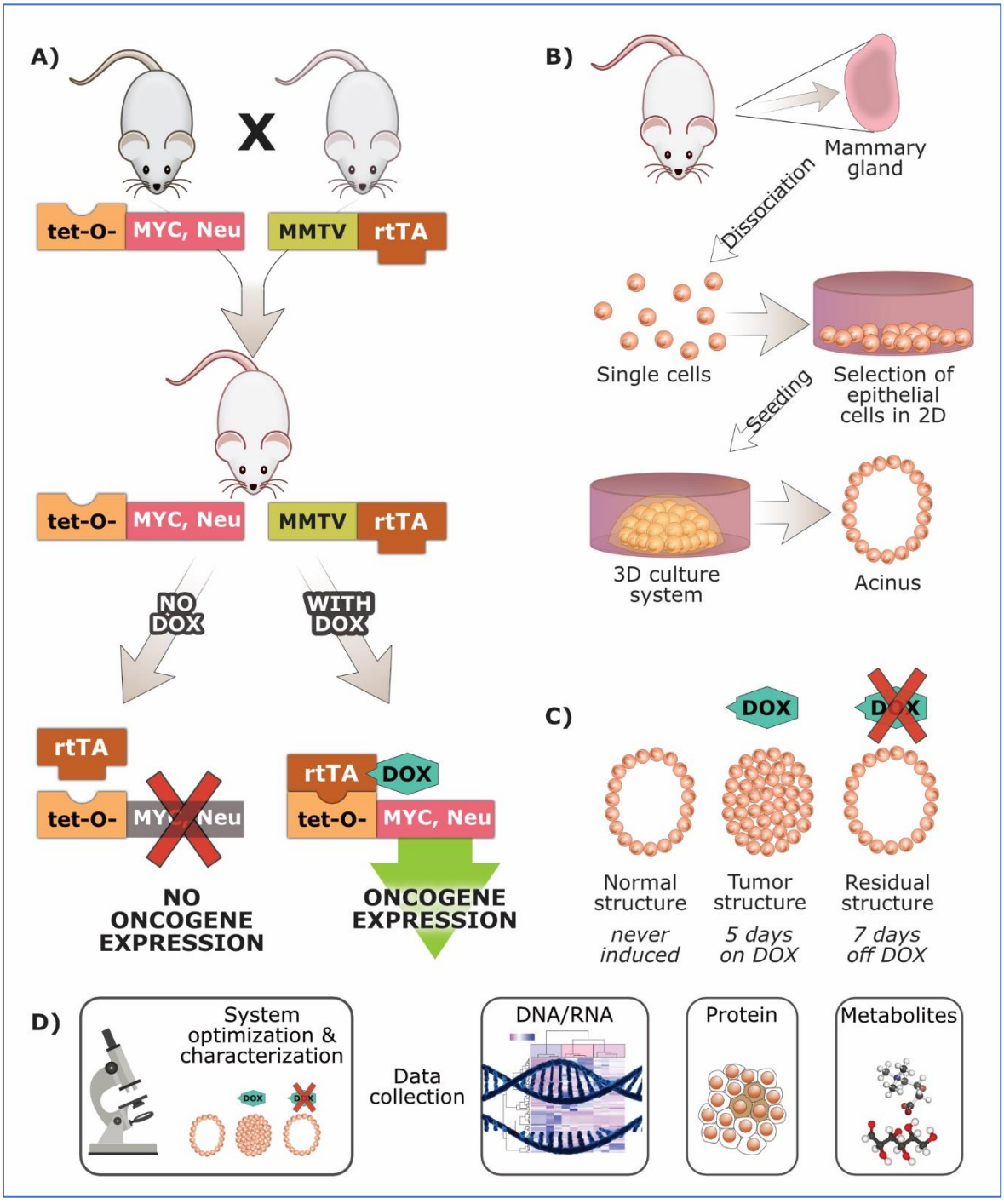

Figure 13: Studying MRD in a breast cancer model. A) Tetracycline inducible mouse models; B) 3D culture system enabled primary mammary mouse epithelial cells to grow and form basic ductal functional unit – acinus; C) Schematic representation of the phenotypes at the most critical timepoints: healthy, never induced (NI) structure, tumor (DOX) and residual (OFF). D) Setting up methods and obtaining information from 3D culture system.

Through collaboration with the Patil group, bioinformatic models were built and developed to help us gain a further understanding of processes that were taking place in the residual cell population *in vitro* (Figure 14A). For *in vivo* validation of metabolic memory and the most prominent features found in 3D model of MRD, regressed mammary gland tissue at the timepoint of 9 weeks after doxycycline withdrawal was used, along with the corresponding age-matched controls (Figure 14B). Namely, mice were fed with doxycycline food until they reached the full-blown tumors (4-6 weeks on Dox). When the tumor burden was large, they were put off doxycycline food and taken when tumors were regressed and non-palpable (9 weeks off Dox). Age-matched controls that were wild type, non-inducible mice, were treated in the same manner, being put on and off Dox food at the same time as animals with tumors. In this way, comparison between these different states was possible.

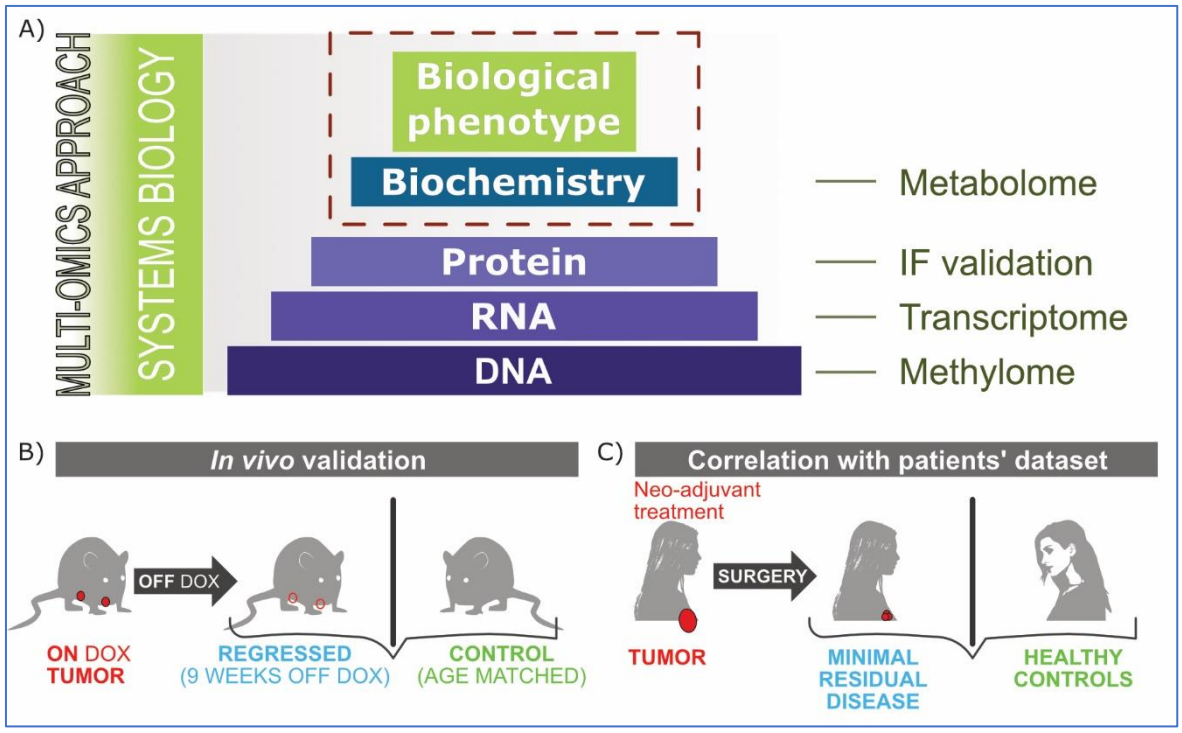

Figure 14: Data integration, validation and interpretation. A) Multi-omics approach for global overview of the system; B) In vivo validation of the most prominent findings from 3D culture; C) Correlation of main findings with human datasets obtained samples after neoadjuvant treatment and healthy controls.

While mouse models still represent an invaluable tool in exploring human diseases, as regulation of genome seems to be very similar between the two species, there are certain differences, particularly in metabolic processes. To check if the identified metabolic

features could have some relevance in human breast cancer, findings from the mouse model were overlaid and compared to the human datasets. Namely, relevance of the observations verified *in vivo* was further assessed through correlation with available databases from patients that were undergone neo-adjuvant treatment (Figure 14C). Following neoadjuvant treatment, often given to make large tumors operable, residual cells can remain as a result of resistance to therapy. Publicly available gene expression profiles of residual cells were compared with the ones from healthy control samples obtained from women who had reduction mammoplasty.

3.2 Phenotypic characterization and optimization of *TetO-cMYC/TetO-Neu/MMTV-rtTA* 3D culture system

Our 3D culture system faithfully recapitulates the dynamics of tumorigenesis and regression, as shown in comparison to mammary gland biopsies taken from *TetO-cMYC/TetO-Neu/MMTV-rtTA* transgenic mice (119) (Figure 15).

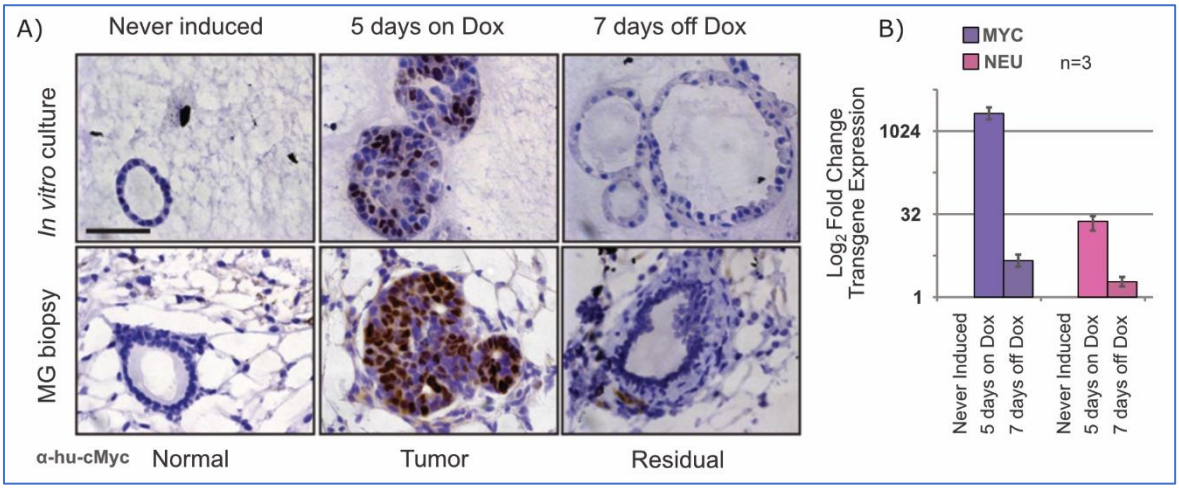

Figure 15: Model of tumorigenesis and regression, immunohistochemistry (IHC) on MYC protein. A) Upper panel: in vitro images - 3D culture, from left to right: never induced structures; structures 5 days on doxycycline (Dox) (tumor correlate); structures 7 days off doxycycline (residual structures). Lower panel: in vivo images - mammary gland biopsies, from left to right: tissue section showing duct from a healthy gland, tumor, regressed gland. Scale bar: 50 µm. B) Expression of transgenes: human MYC and rat Neu at 5 days on Dox and 7 days off Dox (β-actin used as a reference). Error bars represent the standard error of the mean (±SEM). Figure taken and adapted from (119).

In our *TetO-cMYC/TetO-Neu/MMTV-rtTA* 3D cultures, several days during growth in 3D culture, cells form acinar-looking, polarized structures with the lumen (Figure 15A, upper

panel, left), resembling and mimicking basic ductal units of mammary glands, representing normal structures without oncogene expression (Figure 15A, lower panel, left). Upon induction with the doxycycline, cells start proliferating in an uncontrolled manner, filling up the lumen of the duct (Figure 15A, middle) and expressing the oncogenes (MYC stain shown in Figure 15A, middle; *MYC* and *Neu* mRNA expression shown in B). Doxycycline withdrawal results in a rapid regression and clearance of the lumen (Figure 15A, right) and loss of oncogene expression (absence of MYC in Figure 15A, right; low mRNA expression of *MYC* and *Neu* in B), but not in the complete tumor eradication, as the surviving rim remains.

One of the first things that needed to be assessed and optimized in our 3D culture system was the appropriate concentration of doxycycline for induction of the oncogenes: human *MYC* and rat *Neu*. Doxycycline belongs to tetracyclines, antibiotics that bind to the small subunit of ribosomes of prokaryotes, thus interfering with their protein synthesis (125). Since mitochondria resemble prokaryotic cells in their ribosome subunits, doxycycline could also affect mitochondrial protein synthesis (125). This is why several doxycycline concentrations were assessed, whereby a potential effect on the phenotype was examined *in vitro* and *in vivo* (Figure 16). In the 3D cultures, same structures were followed over time, using bright field microscopy, starting from their normal, never induced phenotype (time 0) (Figure 16A). Some 24 hours upon induction, structures became thicker. Proliferation continued and structures were getting larger. Proliferating cells were filling up the lumen and at 5 days on doxycycline, structures represented a correlate of a full blown-tumor, while uninduced structures remained normal, representing a healthy phenotype. In both commonly used doxycycline concentration of 1000 ng/ml (Figure 16A, middle panel) and the lowest tested concentration of 150 ng/ml (Figure 16A, lower panel) there were no differences in the induction kinetics. Lower concentrations of doxycycline did not affect oncogenesis and the tumor phenotype of the induced structures. In addition, structures at lower doxycycline concentration appeared healthier, as not much cell death was observed around them.

Efficiency of different dosage of doxycycline on oncogene activation was also assessed through counting the number of induced structures in 3D culture, expressed in percentage (Figure 16B). While the percentage of induced structures was the highest at 1000 ng/ml (90-95%), it remained high even at the lowest doxycycline concentration tested (77-87%). The percentage of induced structures was higher at a slightly higher doxycycline concentration of 200 ng/ml (81-94%), almost reaching the percentage seen at 1000 ng/ml.

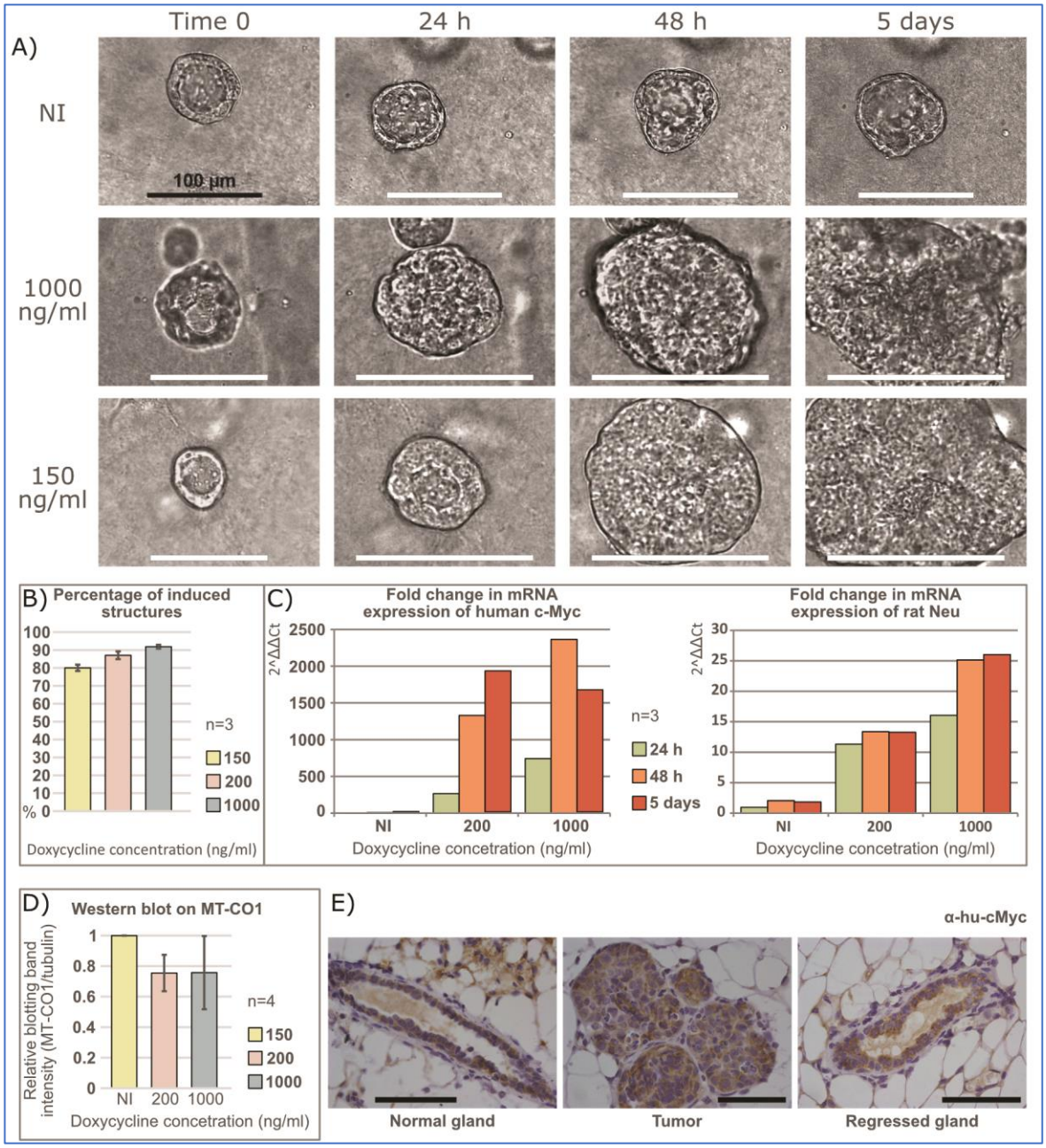

Figure 16: Examining the effect of doxycycline concentration on phenotype, oncogene expression and mitochondrial function. A) Bright-field (BF) images of 3D cultures at indicated timepoints and Dox concentrations (ng/mL); B) Percentage of induced structures in vitro (5 days on Dox) at the indicated Dox concentrations; C) Expression of the transgenes in vitro at the indicated timepoints and Dox concentrations: left – human MYC, right – rat Neu; D) Western blot on MT-CO1 protein in vitro (5 days on Dox) at the indicated Dox concentrations; E) IHC on MT-CO1 protein in vivo (from left to right: control, tumor at 3 weeks on Dox, regressed gland at 2 weeks off Dox).
Error bars represent the standard error of the mean (±SEM). Scale bar: 100 µm.

Next, we employed qPCR to measure the mRNA expression of human *MYC* and rat *Neu* oncogenes in the cells harvested from 3D culture. The analysis showed approximately 1,7 times lower *MYC* expression over 24 and 48 hours upon induction at the lower doxycycline concentration compared to the 1000 ng/ml. Nevertheless, on the endpoint of fifth day on doxycycline, the expression was actually approximately 1,7 times higher than the one at 1000 ng/ml doxycycline, which could be explained by potential apoptotic effect of high *MYC* expression (Figure 16C, left). As for the *Neu* oncogene, its expression at the lower concentration of doxycycline at 24 hours was around 1,4 times lower than at 1000 ng/ml. At 48 hours and 5 days the expression was about 2 times lower than at 1000 ng/ml (Figure 16C, right).

Mitochondrially-encoded cytochrome C oxidase 1 (MT-CO1) protein, which is a part of the respiratory complex IV, was used for assessing the potential effect of doxycycline on mitochondrial function in 3D cultures (Figure 16D) and *in vivo* (Figure 16E). Western blots were conducted on protein lysates from 3D cultures. Relative blotting band intensity of MT-CO1 compared to tubulin showed some effect of doxycycline on MT-CO1, as a decrease of intensity (from 1 down to 0,8) was observed in the samples treated with doxycycline in comparison to never induced controls (Figure 16D). However, this was not significant (One-Way ANOVA, F=1,49, p=0,25 at CL=95%). In addition, IHC staining for MT-CO1 was performed on the FFPE tissue sections obtained from mice with tumors (3 weeks on Dox food), mice with regressed mammary glands (2 weeks off Dox food) and age-matched controls (Figure 16E). The staining showed that there was no dramatic effect of doxycycline on MT-CO1 expression at the commonly used concentrations of 625 mg/kg, supplemented in the food of the experimental mice for inducing fully palpable tumors.

Overall, lowering doxycycline concentration did not affect the induction of oncogenes or tumor phenotype. Based on these experiments, we decided that a doxycycline concentration of 200 ng/ml would be used for oncogene induction, as it was high enough to induce oncogenesis, yet could be potentially safer for metabolomics experiments, regarding the effect of doxycycline on the mitochondria.

To further investigate gene expression profiles and obtain full transcriptomic picture of the residual cells, protocols for RNA sequencing were tested. The established RNA isolation protocol using the mirVana kit proved highly effective for obtaining high quality RNA for deep sequencing employed for obtaining transcriptomics datasets. Immunofluorescence protocols were successfully modified and optimized for *in vitro* and *in vivo* material, which enabled full phenotypic characterization of the respective structures in a normal, physiological state, and during tumorigenesis and regression.

3.3 Residual population exhibits transcriptional profile distinctive from tumor and normal physiological state

We employed immunofluorescence (IF) staining and confocal microscopy on 3D cultures to observe and compare the phenotypes of normal (never induced), tumor (5 days on doxycycline) and residual (7 days off doxycycline) structures in greater detail. IF staining for polarity markers and the *MYC* oncogene (Figure 17A) showed that tumor structures lost their lumen and polarity (upper panel, middle), as opposed to normal which exhibited a clear lumen and a single, polarized rim (upper panel, left). In normal condition, alpha-6 integrin (ITGA6, shown in red) had a basal localization; Zonula Occludens (ZO-1, shown in green) was apical, staining occludins at the tight junctions; Golgi matrix protein (GM-130, shown in magenta) was restricted to apical part of the cell, marking Golgi. This was not the case in the tumor, where the expression of these markers was less consistent. However, a great phenotypic similarity was observed between normal and residual structures, as they regressed back to polarized single rim with the lumen (upper panel, right). Immunofluorescence staining for transgenic MYC showed its expression in tumor cells (lower panel, middle; in green) and its absence in the regressed cells (lower panel, right) similarly to the controls (lower panel, left).

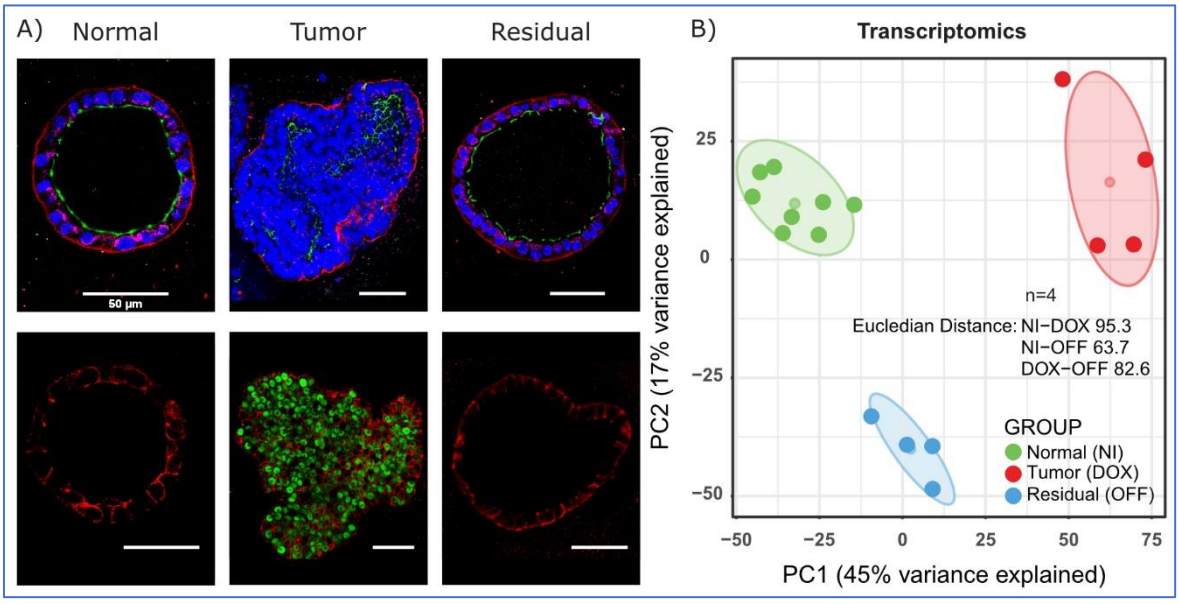

Figure 17: Comparison of phenotypes and transcriptional profiles between normal, tumor and residual structures. A) Upper panel: ITGA6 (red), ZO-1 (green), GM-130 (magenta). Lower panel: human MYC oncogene (green); E-cadherin (CDH1, in red) as a counterstain. From left to right: normal, tumor and residual structures. Scale bar: 50 µm. B) PCA plot based on transcriptional profiles from RNA sequencing data of normal (NI), tumor (DOX) and residual (OFF) structures.

Despite this similarity, it was demonstrated that the residual structures exhibited a specific gene expression profile, observed from microarray data in our previous study (119). To further obtain the in-depth picture of the whole transcriptome, we performed RNA sequencing on material collected from 3D cultures. This confirmed that the residual cell population (OFF, shown in blue) indeed had a distinctive transcriptional profile that distinguished it from both tumor (DOX, shown in red) and normal (NI, shown in green), as seen in the Principal Component Analysis (PCA) (Figure 17B). The first and the second principal components accounted for 45% and 17% of the variance, respectively. Centroids were calculated based on two principal components and showed the center of each sample group (control, tumor, residual) represented on the PCA plot. Distance was calculated considering all principal components, as Euclidean distance between the centroids of the samples. In this regard, residual structures (OFF) were closer to the normal (NI) ones (distance between them was 63,7). The distance between residual structures (OFF) and tumors (DOX) was larger (82,6). The largest distance was seen between tumor (DOX) and normal (NI) cells (95,3). The confidence interval of a multivariate normal distribution is represented in the ellipses (CI = 0,95).

To gain an understanding of the affected pathways, 1925 differentially expressed genes in residual cells (compared to never induced control, q value<0,05) were taken for GO (Gene Ontology) Enrichment Analysis. GO terms obtained from genes with lower expression showed downregulation of processes mostly connected to cell division, which reflected the dormant state of the residual population (Figure 18A). GO terms coming from upregulated genes reveal activation of the cell surface receptor signaling pathway, cell adhesion and locomotion, probably reflecting cellular movements towards integration into the single rim, establishment of contact with the basement membrane and restoration of polarity. Interestingly, metabolic processes such as ATP generation, nucleotide phosphorylation and glycolysis were seen upregulated in residual cells (Figure 18B) despite their non-proliferative state, hinting at the potential significance of metabolism in the MRD. To check for altered pathways in tumor cells, GO enrichment of the 5274 differentially expressed genes in comparison to the control (p_{adj}<0,1, Bonferroni correction for multiple testing) was also performed (Figure 19). In contrast to the residual cell population, cell adhesion, motility and cell communication were downregulated in tumor cells compared to the control (Figure 19A), while cell division was upregulated (Figure 19B). This was expected for the proliferating cells that did not have contact with the extracellular matrix anymore and lost contact inhibition, filling up the lumen. In addition, many metabolic and biosynthetic processes were enhanced in the tumor state, which is unsurprising, considering that these cells have high demands for energy and support of proliferation.

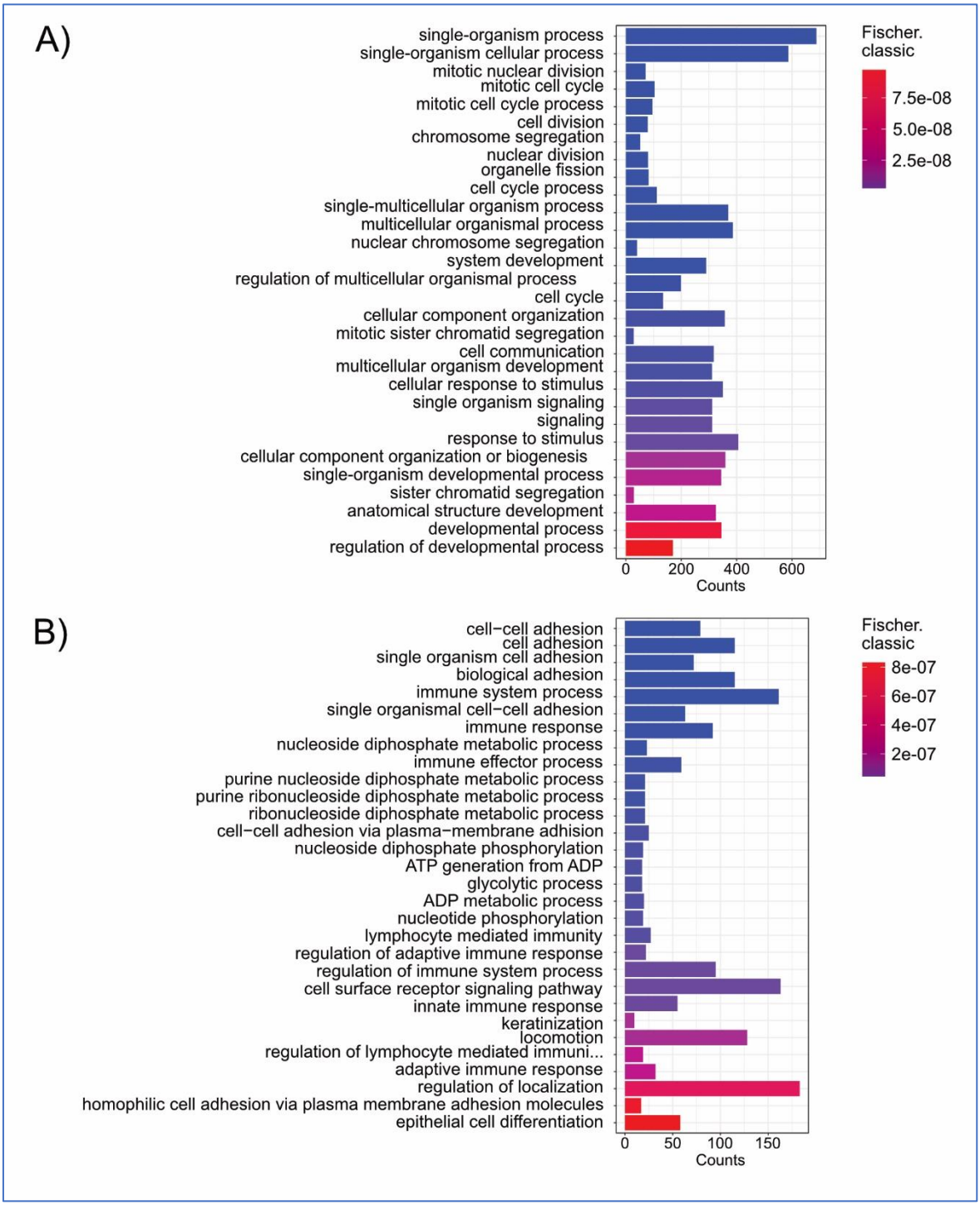

Figure 18: Altered biological processes in residual population compared to normal (Fisher classic test; p-value shown in the legend). A) GO terms: top 30 (p<0,01) downregulated processes in residual (OFF) compared to control (NI) population. B) GO terms: top 30 (p<0,01) upregulated processes in residual (OFF) compared to control (NI) population.

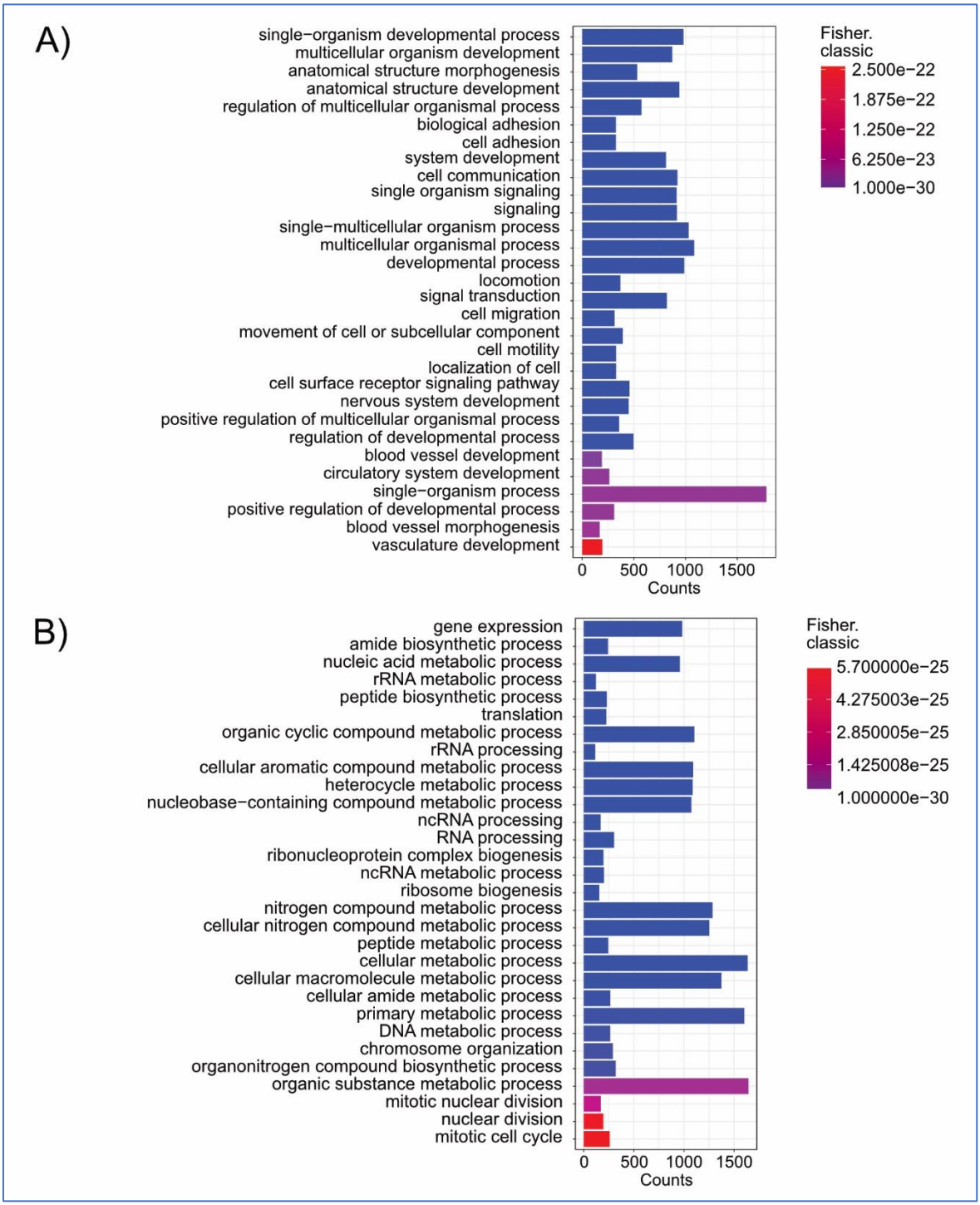

Figure 19: Altered biological processes in tumor population compared to normal (Fisher classic test; p-value shown in the legend). A) GO terms: top 30 (p<0,01) downregulated processes in tumor (DOX) compared to control (NI) population. B) GO terms: top 30 (p<0,01) upregulated processes in tumor (DOX) compared to control (NI) population.

3.4 METABOLIC FEATURES OF RESIDUAL CELL POPULATION

3.4.1 Establishment of metabolomics methods for studying *TetO-cMYC/TetO-Neu/MMTV-rtTA* 3D culture system

Seeing the changes on the transcriptional level (described in Chapter 3.3) that pointed out to the alterations in metabolic processes happening in the residual cells (shown in Figure 18B), we wanted to further examine these findings on the metabolic level. In addition, acknowledging altered metabolism as one of the hallmarks of cancer, we wanted to explore it in the context of MRD. The goal was to obtain a metabolic landscape that acts as a "signature" of MRD, and also to explore the impact that metabolic rewiring could have on the residual cells.

Measuring metabolites also brings us closer to the cellular phenotype, as it reflects real-time processes taking place in the cell and is a direct read-out of its biochemistry. For this purpose, we employed untargeted, high-throughput metabolomics using high-resolving mass spectrometer (flow injection Q-Exactive MS) that enables detection of up to 10 000 ions and profiling of thousands of metabolites. This approach was untargeted, allowing detection of changes across wide range of molecules, which yielded global profiles of our normal, tumor and residual population. However, it was semi-quantitative with putative metabolite annotations based on accurate mass and isotope distribution, without the possibility to resolve isomers. Thus, as a more precise method we employed gas chromatography-mass spectrometry (GCMS). GCMS provided retention time and mass that, along with the specific fragment pattern, allowed for precise metabolite identification and quantification. We took advantage of both untargeted metabolomics and GCMS features to characterize our 3D cultures on a global level and at the same time gain insights into more specific metabolic changes.

Employing metabolic methods *in vitro* first required establishment of metabolic methods for our 3D culture system. The great challenge was imposed by the fact that the structures were growing in the Matrigel and initially the Matrigel signal was too strong in the GCMS profile, making it impossible to see other metabolites. Since there were no established protocols for 3D cultures in Matrigel, we firstly needed to set up the methods for efficient quenching and collecting of the metabolites. Several approaches were tried and eventually 1,5-hour long digestion of Matrigel using collagenase and liberase was shown to be the most successful in freeing the structures from the Matrigel, without interference from the Matrigel signal. We also had to try different strategies regarding cell number that would give a good signal, but also to tackle growth conditions and regularity of media changes. Eventually, we chose to change media every day at the

same time. This strategy was very important as it did not dilute and affect good signal from metabolites (which could have been potential problem especially in case of extracellular metabolites) and did not additionally stress the cells, as they had enough nutrients. Experiments were designed in such a way that the metabolites from different states (normal, tumor, residual) were collected and processed at the same time. Before harvesting intracellular metabolites, extracellular metabolites were also collected. After fast washing the structures in PBS, metabolites were quenched with absolute cold methanol and stored at -80°C. Upon assessing and deciding on solvents concentrations to be used during extraction, samples were further processed - extraction of metabolites was successful using a methanol-water 1:1 extraction protocol, with a 15-minute-long denaturation of cellular proteins at 72°C. GCMS data acquisition was optimized by fine-tuning different concentrations of derivatization agents, split ratio, detector voltage and ion transitions. The established method was suitable for harvesting the metabolites for different metabolic analyses.

3.4.2 Residual population is metabolically similar to tumor state, despite the absence of oncogene expression

To get the wide metabolic picture across different states (control, tumor, residual), both intracellular and extracellular metabolites samples from 3D cultures were run using high-throughput discovery metabolomics. Total of 7941 intracellular and 6791 extracellular raw ions were detected.

Upon performing quality control and checking the consistency of raw ion detection across injections, ions that had been detected reproducibly were taken for further analyses (Appendix, Figure 52). Their number was 2832 for intracellular and 3208 for extracellular samples. Clustergrams were obtained by hierarchical bi-clustering of samples based on measured intracellular (Figure 20) or extracellular (Figure 21) ions and metabolite ion clusters.

Strikingly, the clustergram of samples with measured intracellular metabolites (Figure 20) showed that, on a metabolic level, residual population (shown in blue) clustered with tumor samples (shown in red), suggesting that it retained some metabolic features acquired in the tumor state. Analysis also included metabolites taken from cells of wild type (WT), non-inducible animals (shown in dark green). These cells were treated with media containing doxycycline, as an important control of its potential effect on metabolic profiles. Clustering of these samples with normal (NI) controls (shown in green) showed that this clustergram pattern did not come from the treatment with doxycycline.

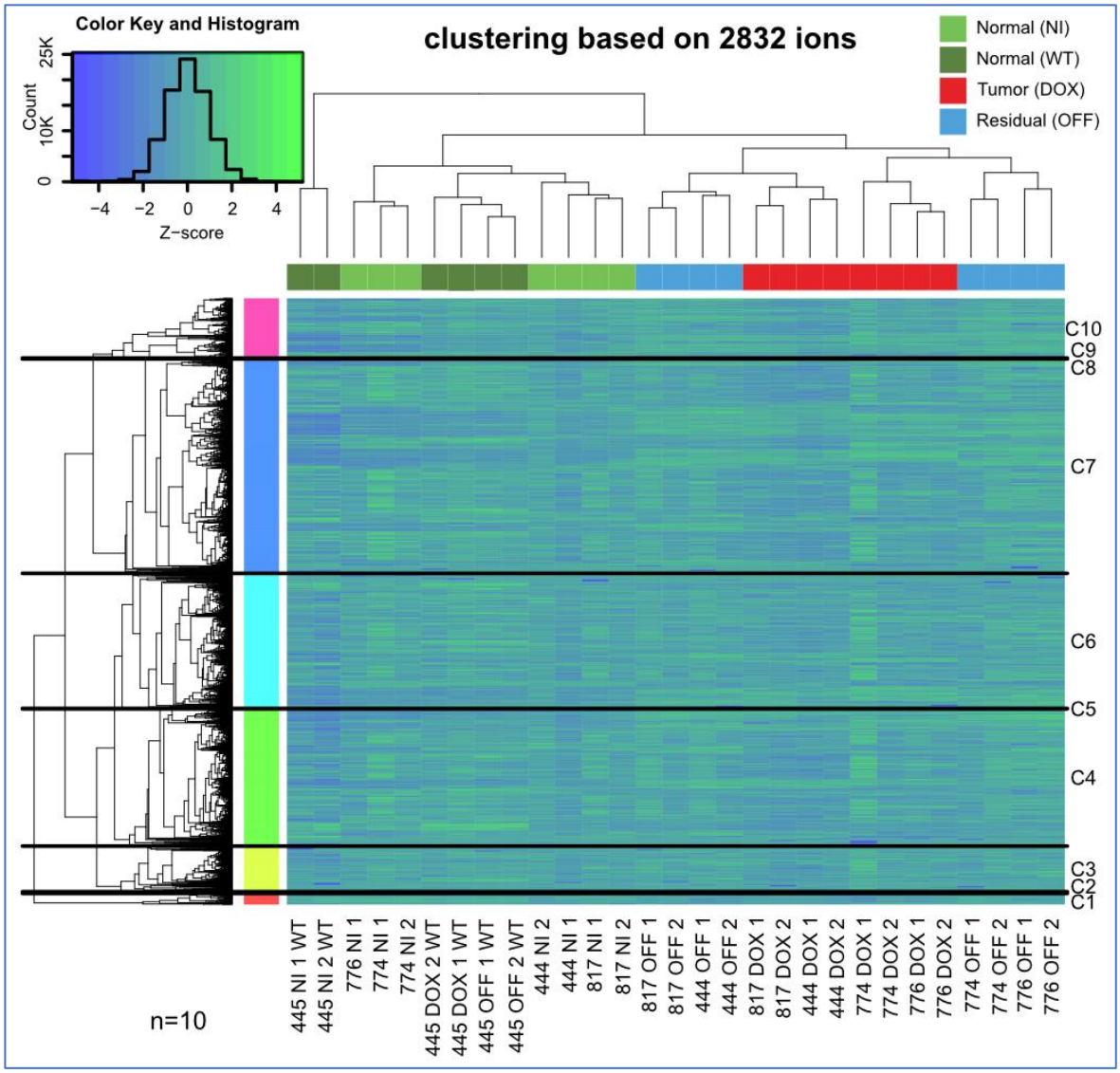

Figure 20: Clustergram of normal (NI, WT), tumor (DOX) and residual (OFF) cell culture samples, based on measured (Q-Exactive MS) intracellular ions and metabolite ion clusters (C1-C10), obtained by unsupervised hierarchical bi-clustering and Manhattan distances.

Bi-clustering of extracellular metabolites showed that the separation of the samples based on their condition (normal, tumor, residual) was not really clear (Figure 21). This was probably due to the detection of many more non-significant metabolites in the media.

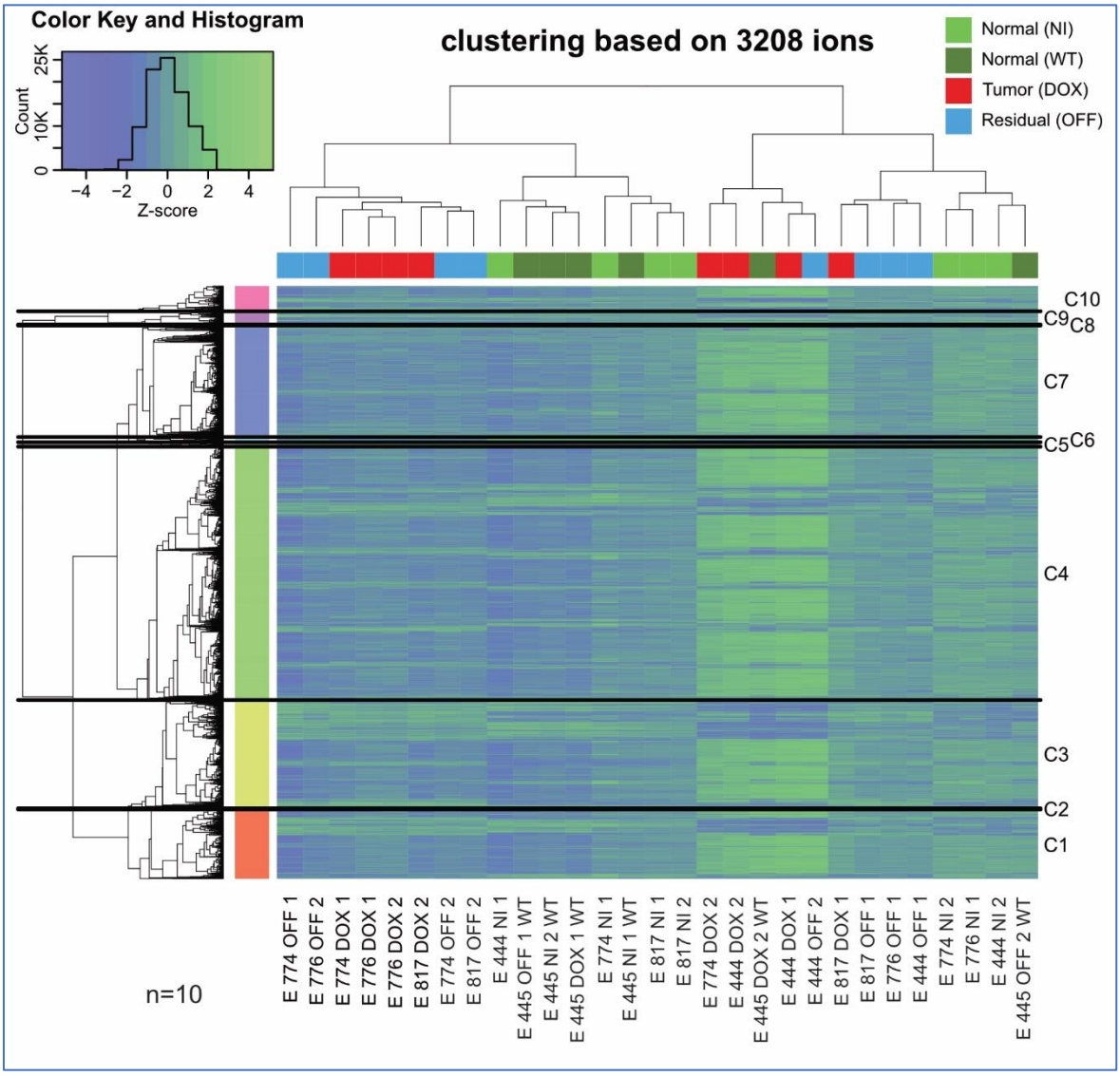

Figure 21: Clustergram of normal (NI, WT), tumor (DOX) and residual (OFF) cell culture samples, based on measured (Q-Exactive MS) extracellular ions and metabolite ion clusters (C1-C10), obtained by unsupervised hierarchical bi-clustering and Manhattan distances.

For intracellular samples, 813 ions were annotated as metabolites. Number of unique metabolites (without double/several annotations) was 2235. Metabolites in clusters were used in metabolic pathway enrichment analysis (left-tail hypergeometric test), where p-values marked the probability that having up to a certain number of metabolites in a certain pathway was more than expected by chance. Enrichment analysis was conducted using KEGG metabolic pathway definitions. Metabolites that were not detected have been removed before the analysis and ambiguous annotations were counted individually. In

five clusters (2, 3, 5, 8 and 9) there were no significant changes with respect to a higher number of metabolites appearing in 85 metabolic pathways. Significant enrichment was seen in: tyrosine and tryptophan metabolism in cluster 1; fructose and mannose metabolism, beta-alanine metabolism and amino sugar and nucleotide sugar metabolism in cluster 4; phenylalanine metabolism in cluster 6; pentose and glucuronate interconversions, purine metabolism; valine, leucine, isoleucine biosynthesis, D-glutamate and D-glutamine metabolism and glyoxylate and dicarboxylate metabolism in cluster 7; histidine metabolism, tryptophan metabolism and vitamin B6 metabolism and folate biosynthesis in cluster 10.

For extracellular samples, 1039 ions were annotated as metabolites and 2501 were unique metabolites. Consistent with unclear separation of the extracellular samples on clustergram, metabolic pathway enrichment analysis of metabolites in clusters did not reveal any significantly upregulated pathway.

Analysis of intracellular ions (unpaired t-test) was further performed to assess which metabolites could be differentiating residual from the normal population. Some of the highest log2 fold changes (FC) of residual versus normal values were observed in ions that could not be annotated, with log2 FC reaching up to 5,3; $p<0,05$, as well as certain dipeptides with log2 FC reaching up to 4,3; $p<0,05$. Besides dipeptides, annotated ions contributing significantly ($p<0,05$) to the differences between the two populations could be corresponding to metabolites such as the ones presented in Figure 22 (residual population shown as boxplots in blue; normal in green; tumor in red). These examples included leucine (m/z 130.087; log2 FC=0,82), proline (m/z 114.056; log2 FC=1,06), succinic acid (m/z 117.019; log2 FC=1,48), methionine (m/z 148.044; log2FC=1,50) and adenosine monophosphate (AMP) (m/z=346.056; log2 FC=1,95) that were present in higher levels in the residual cells (and tumor); 5,10-methenyltetrahydrofolic acid (m/z=227.072; log2 FC=-1,11) was found in lower amounts. Elevated metabolites were also observed to be significantly higher in the tumor cells (boxplots in red) compared to the normal: leucine (log2 FC=0,75), proline (log2 FC=1,73), succinic acid (log2 FC=0,94), methionine (log2 FC=1,52) and AMP (log2 FC=2,35). Ions corresponding to 5,10-methenyltetrahydrofolic acid were also recorded significantly lower in the tumor samples (log2 FC=-1,0). In addition, significantly higher levels of glutathione (m/z 306.077) and oxidized glutathione (m/z 305.069) were found in the residual (log2 FC=0,95; log2 FC=0,97, respectively) and tumor (log2 FC=2,54; log2 FC=1,54, respectively) cells. Changes were also seen in beta-alanine (m/z 88.041; residual log2 FC=0,88; tumor log2FC=1,35), lysine (m/z 145.098; residual log2 FC=0,64; tumor log2

FC=0,70), tyrosine (m/z 180.067; residual log 2FC=0,66; tumor log2 FC=1,0) and various other metabolites.

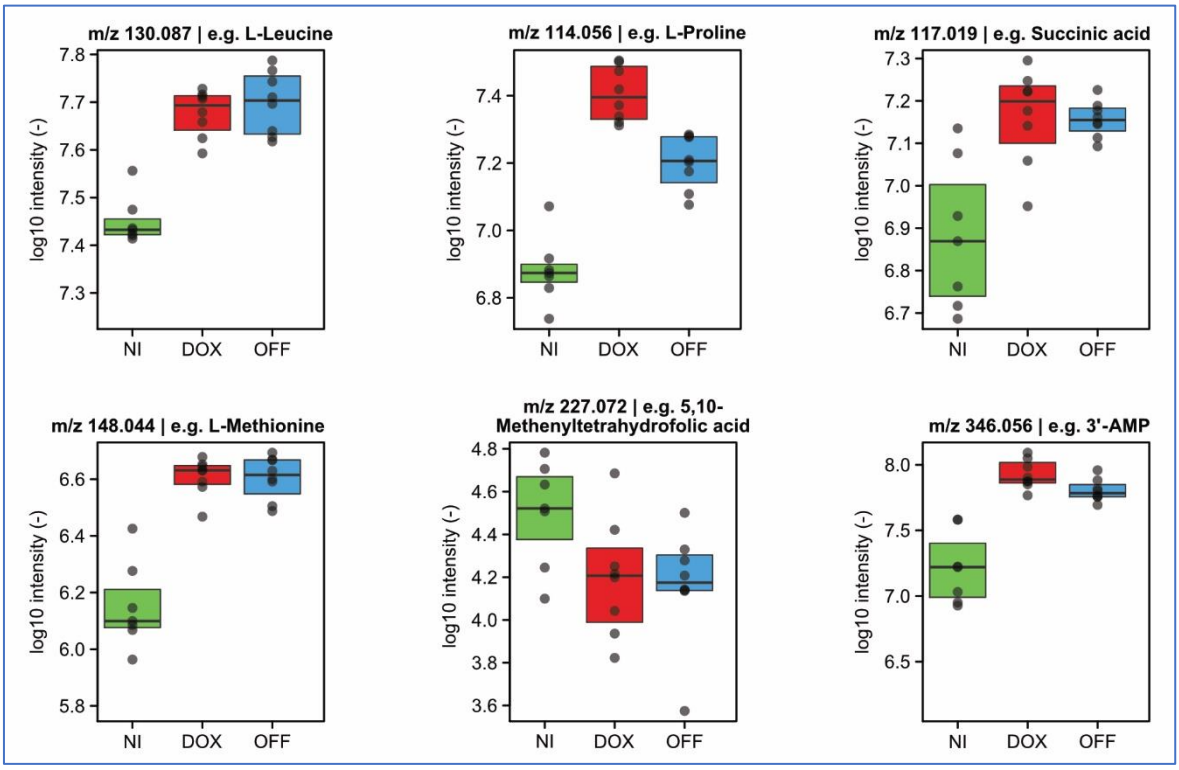

Figure 22: Selected intracellular ions significantly different (unpaired t-test; p<0,05) between residual (OFF) and normal (NI) cells, measured by Q-Exactive MS.

The same analysis was performed to compare residual with tumor samples. Of the annotated ions, some with the significantly changed levels correspond to metabolites shown in Figure 23.

Glutathione (log2 FC=-1,6), 5'-methylthioadenosine (m/z 296.082; log2 FC=-0,78) and diguanosine triphosphate (m/z 393.03; log2 FC=-1,6) were found in lower amount in the residual compared to the tumor cells, but still higher than in the controls. Pantothenic acid (m/z 218.103) was found in significantly lower amount in residual cells compared to the tumor (log2 FC=-0,86), but also to the control (log2 FC=-0,6). Glutamine (m/z 145.062) was significantly higher in residual cells compared to tumor (log2 FC=0,66), as well as anserine (m/z 241.12), found higher than in the tumor (log2FC=1,75) and control (log2 FC=1,30). In addition, around 30 ions without annotations and several dipeptides with log2 FC ranging from 2 to 4,6 were significantly different between residual and tumor cells.

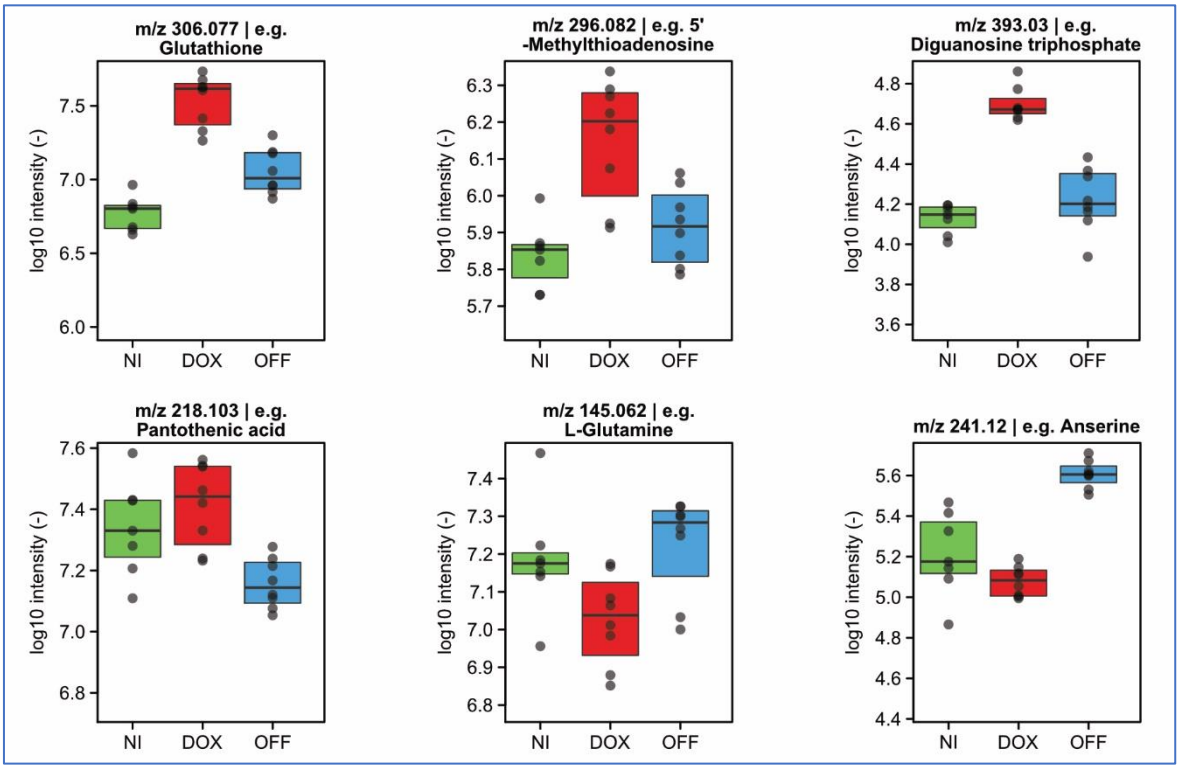

Figure 23: Selected intracellular ions significantly different (unpaired t-test; p<0,05) between residual (OFF) and tumor (DOX) cells, measured by Q-Exactive MS.

Analysis of extracellular ions showed that some of the metabolites contributing to differences between the residual population and control were (Figure 24): lactic acid, ornithine, arginine, glutamine, methionine and ribose 1,5-bisphosphate.

Lactic acid (m/z 89.025, log2 FC=1,09) and ornithine (log2 FC=1,57) were found in significantly higher amounts in residual cells, while arginine (m/z 173.104; log2 FC=-2,07) was found in lower amounts. In the tumor cells, compared to the control, extracellularly located lactic acid (log2 FC=0,90) and ornithine (log2 FC=1,36) were also significantly higher. Arginine was also significantly lower extracellularly in tumor cells compared to the control (log2 FC=-2,99). Additionally, levels of extracellular glutamine were significantly lower in the tumor population compared to both residual (log2 FC=-1,19) and normal (log2 FC=-1,11), which could be due to the proliferative state of tumor cells as opposed to non-proliferative residual and normal cells. The same was observed in the case of extracellular methionine (m/z 148.044) which was significantly lower in the tumor compared to residual (log2 FC=-1,68) and normal (log2 FC=-0,55) cells. Ribose 1,5-bisphosphate (m/z 308.979) was found to be significantly lower in residual cells compared to normal samples (log2 FC=-3,55).

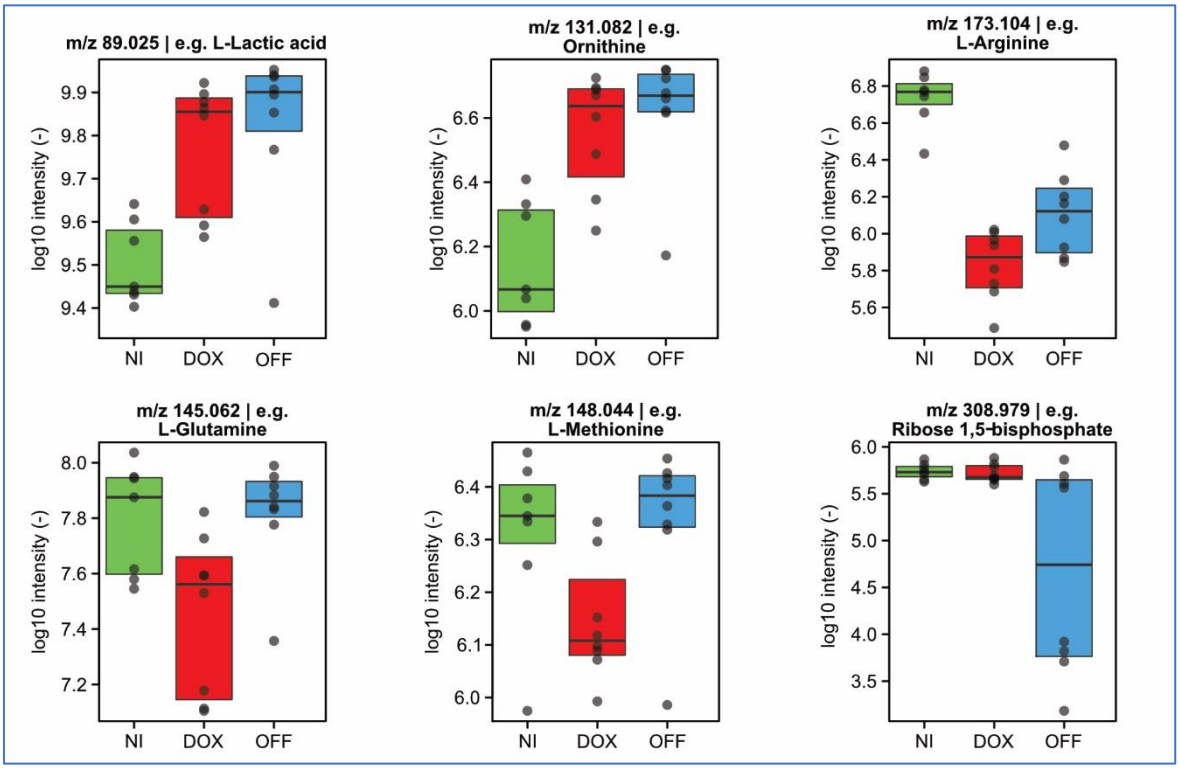

Figure 24: Selected extracellular ions significantly different (unpaired t-test; p<0,05) between residual (OFF), tumor (DOX) and normal (NI) cells, measured by Q-Exactive MS.

All the described significant changes in intracellular and extracellular ions shown in figures 22-24, were not found to be significantly different in the wild-type samples between the never induced (NI) state compared with the doxycycline treated (DOX, OFF) cases.

Taken together, untargeted metabolomics revealed a metabolic resemblance of residual and tumor population. We further confirmed this observed similarity by performing GCMS analysis on the cell culture samples of the three respective states (normal, tumor, residual). PCA plot based on GCMS measurement of intracellular (Figure 25A) and extracellular (Figure 25B) metabolites demonstrated again the clustering of residual population (shown in blue) closer to tumor (shown in red) with a Euclidean distance of 5,6 (for the intracellular metabolites plot) and 4,4 (for the extracellular metabolites plot). Both populations were clearly distinctive from normal (NI, WT shown in green and dark green, respectively), with a Euclidean distance of 8,8 (for both NI-OFF and NI-DOX on intracellular metabolites plot) and 8,7 (for both NI-OFF and NI-DOX on extracellular metabolites plot). Centroids for both PCA plots were calculated based on two principal components, indicating the center of each sample group. The first principal component

in the plot of intracellular metabolites explained 37% of variance and the second18 %, while the first principal component in plot of extracellular metabolites explained 68% of variance and the second 11%. Ellipses indicate the confidence interval of a multivariate normal distribution (CI = 0,95).

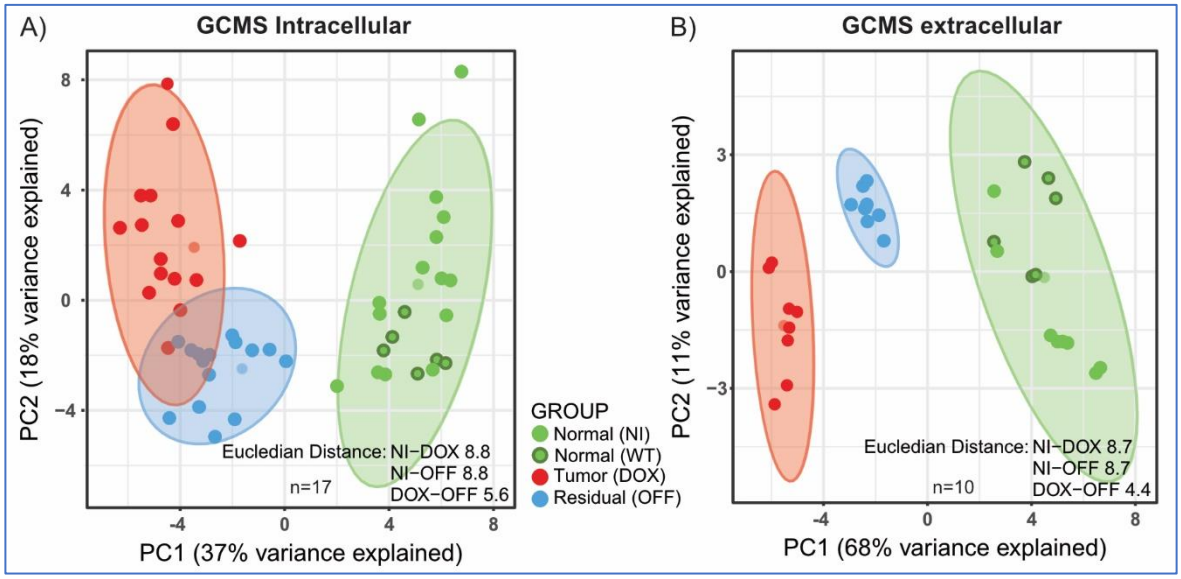

Figure 25: PCA plots from GCMS measurement of the respective states: residual (OFF), tumor (DOX) and controls (NI, WT). A) intracellular and B) extracellular metabolites.

Seeing the metabolic changes and comparing PCA plots of metabolic and transcriptomic profiles, we speculated whether the distinct transcriptomic profile of residual cells could be attributed to changes in non-metabolic genes. To see if their transcriptomic profile would come closer to tumor if we exclude non-metabolic genes, we performed a PCA only on metabolic genes. However, this did not affect the transcriptomic profile pattern (Appendix, Figure 53), indicating that transcriptional differences in genes encoding metabolic enzymes do exist between the tumor, residual and normal populations. To have an insight into what gene expression changes could potentially underlie clustering residual cells closer to tumor, we performed the PCA on the subset of genes differentially expressed in both tumor and residual cells compared to normal, which resulted in the transcriptional clustering more similar to what we observe on metabolic level (communication with Katharina Zirngibl).

3.4.3 One third of detected intracellular metabolites is specifically changed in residual and tumor population

From the heatmap corresponding to 54 intracellular metabolites detected by GCMS, it was clear that the residual (OFF) and tumor (DOX) samples showed metabolic similarity, separating distinctly as one huge cluster away from the normal samples (NI, WT) (Figure 26).

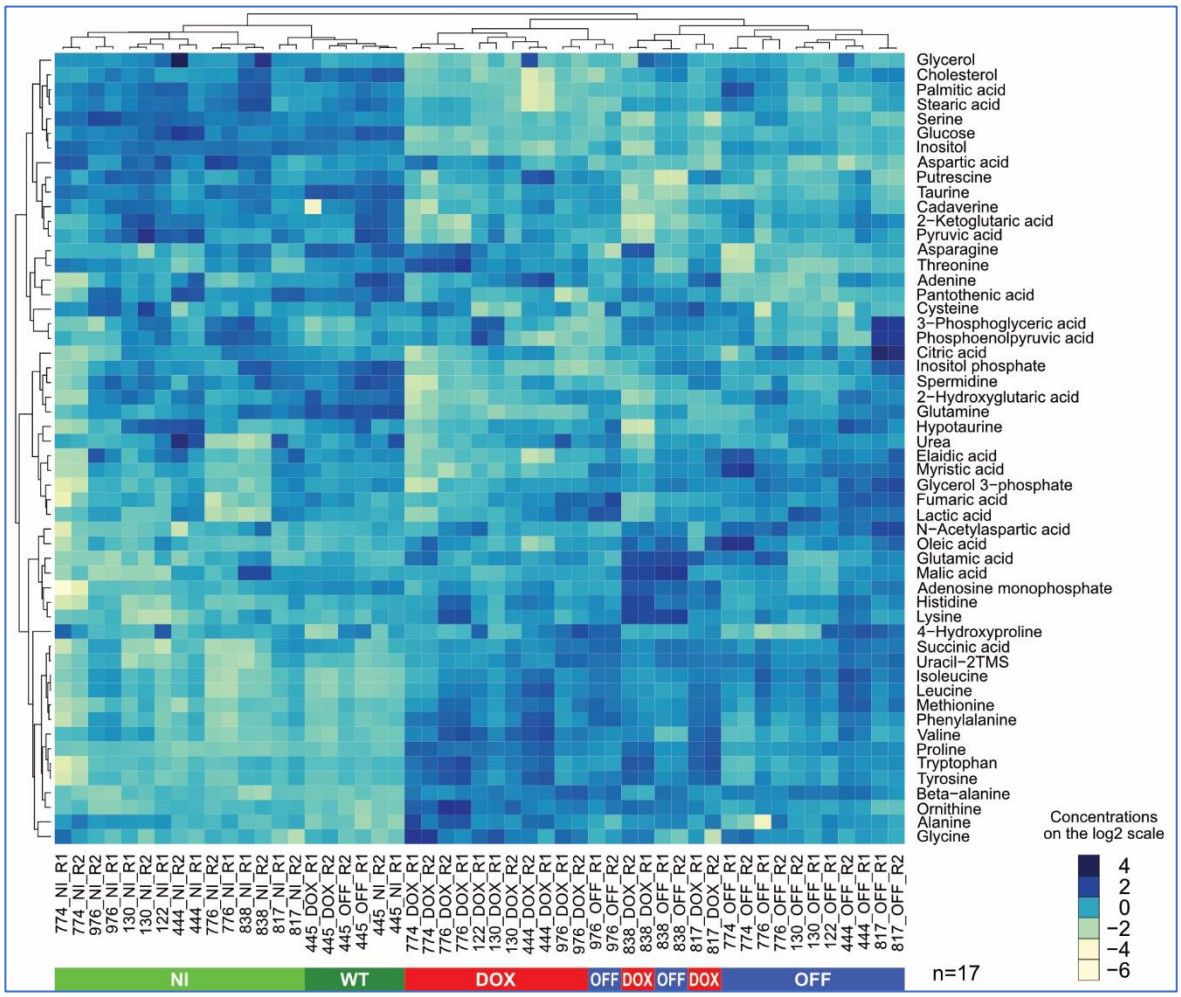

Figure 26: Heatmap of all intracellular metabolites measured by GCMS across the three populations – normal (NI), tumor (DOX) and residual (OFF).

From this number, about one third of intracellular metabolites were at least more than 1-log2 fold change (FC) significantly different (limma linear regression model, Benjamini-Hochberg p_{adj}<0,01) between the three populations (Figure 27). Both residual and tumor cells exhibited higher levels of malic, succinic, N-acetylaspartic and lactic acid,

tryptophan, histidine, uracil, proline and beta-alanine. Metabolites found in lower levels compared to the normal were: aspartic, palmitic, stearic, 2-ketoglutaric and pyruvic acid, inositol, glucose, taurine, glutamine, glycerol and glycerol 3-phosphate. Observed metabolic differences between the residual and tumor population were slight – the most prominent were higher levels of myristic acid (log2 FC=1,14), inositol (log2 FC=0,79) and glucose (log2 FC=1,01) in the residual samples.

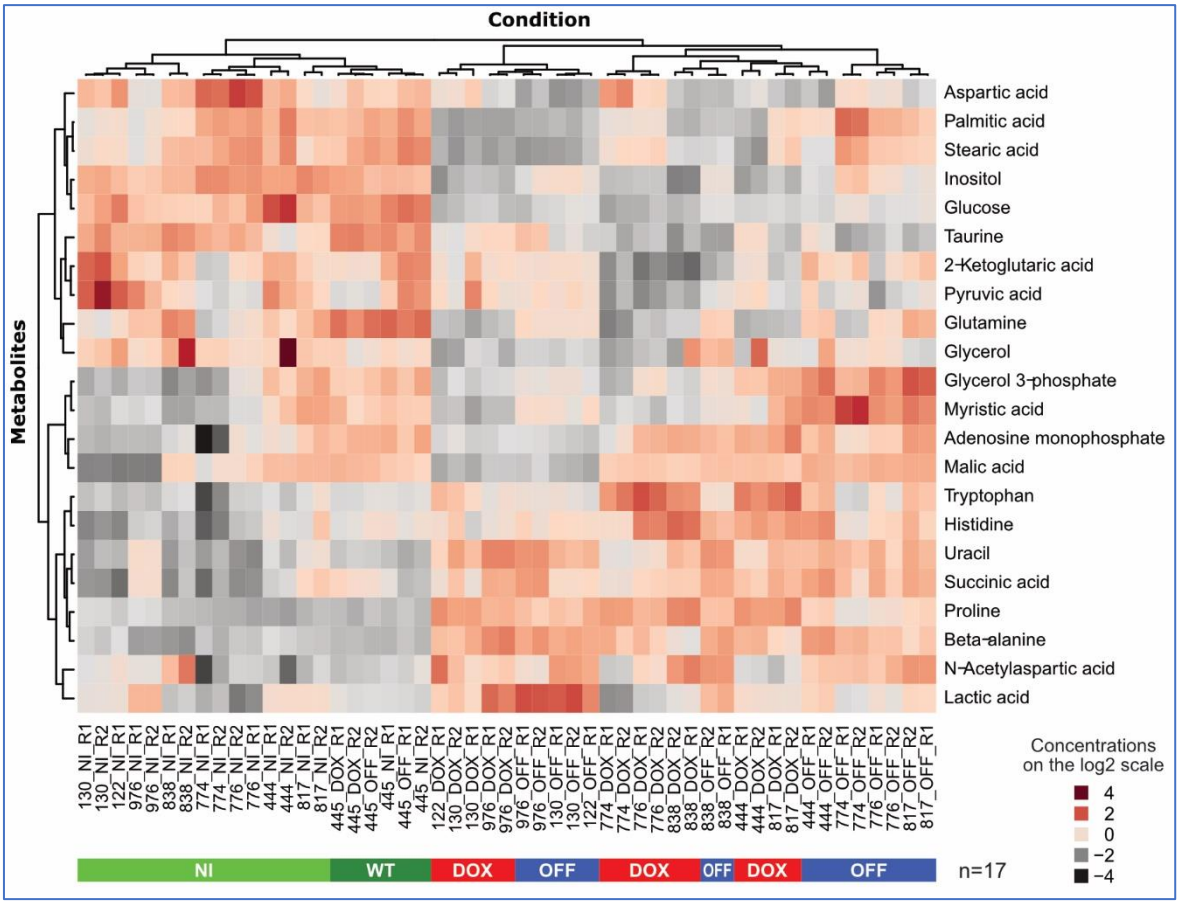

Figure 27: Heatmap of intracellular metabolites that were significantly changed (log2 FC>1, padj<0,01) in residual (OFF) and tumor (DOX) cells compared to controls (NI).

Next, we applied Over Representation Analysis (ORA) using MetaboAnalyst 4.0, a tool for statistical, functional and integrative analysis of metabolomics data (126). The analysis, based on all the significantly changed (log2 FC>0,5; p<0,01) intracellular metabolites measured in the residual samples compared to normal, was performed to facilitate the insight into altered metabolic pathways in residual population (Figure 28). ORA hypergeometric test indicated that several pathways were significantly enriched: amino acid metabolism, ammonia recycling and urea cycle, sugar metabolism,

mitochondrial ETC and polyamine biosynthesis (Figure 28). Raw p-values are indicated in the legend.

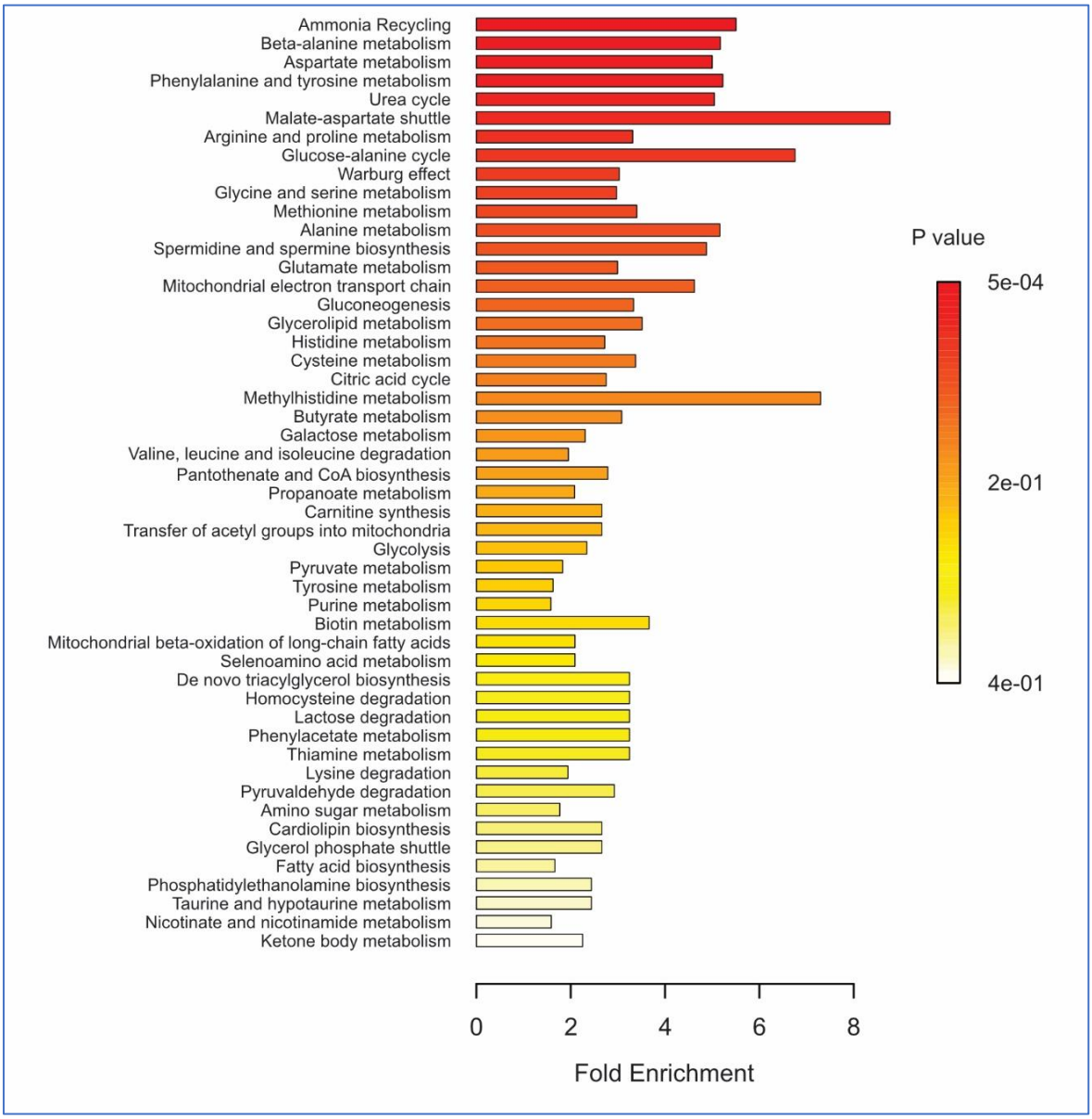

Figure 28: Over Representation Analysis (ORA) based on all significantly changed (log2 FC>0.5, padj<0,01) metabolites in residual population compared to the normal

Extracellular metabolites analysis showed that around one half of metabolites were more than 1-log2 FC significantly different (limma linear regression model, Benjamini-Hochberg p_{adj}<0,01) in residual and tumor cells compared to the normal (Figure 29). The residual population differed more from the tumor than in case of intracellular

metabolites, hinting at alternative pathways and decreased metabolites uptake by non-proliferating residual cells. This is particularly visible in the case of glutamine/glutamate pool which was the lowest in the tumor cells, as expected, since they were highly proliferative and driven by the *MYC* oncogene that controls glutamine metabolism. Higher amounts of extracellular methionine, phenylalanine, isoleucine, serine and glucose was observed in residual cells when compared to the tumor, but it was lower in comparison to the normal cells. Extracellular succinic, fumaric and lactic acid, urea, ornithine, putrescine, proline and alanine were found to be elevated in residual population compared to the normal, albeit not to the extent seen in the tumor.

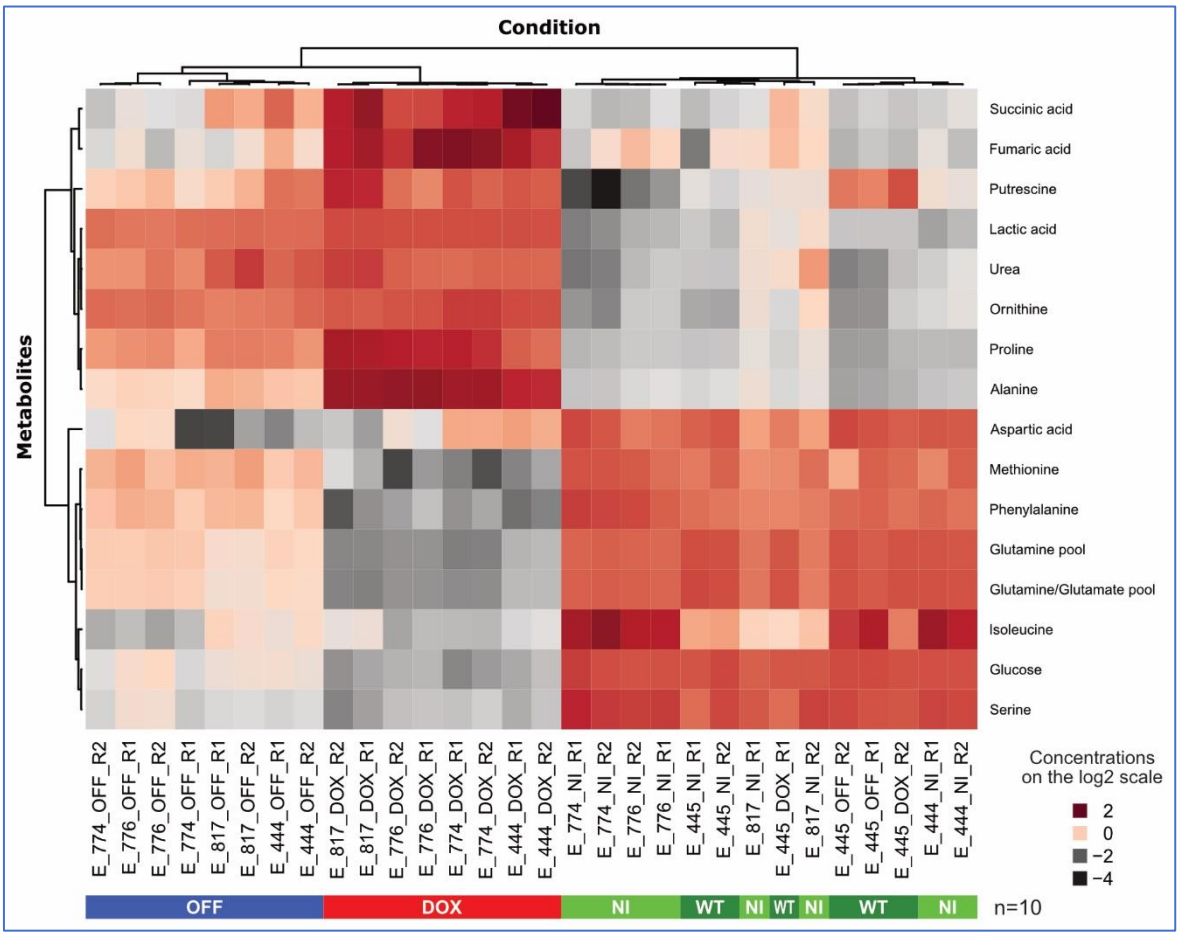

Figure 29: Heatmap of extracellular metabolites that were significantly changed (log2 FC>1, padj<0,01) in residual (OFF) and tumor (DOX) cells compared to controls (NI)

Decreased amount of extracellular glucose and high amount of extracellular lactate, as well as lower levels of intracellular glucose and higher levels of extracellular lactate were observed in both tumor (DOX; boxplots shown in red) and residual (OFF; boxplots shown

in blue) population compared to compared to normal (NI, WT; boxplots shown in green and dark green, respectively), reflecting their glycolytic phenotype (Figure 30). Urea, putrescine and ornithine were not significantly changed intracellularly, but have been found in significantly higher amounts extracellularly (Figure 30) in residual and tumor cells indicating their higher export from the cells and hinting at the alterations in the urea cycle.

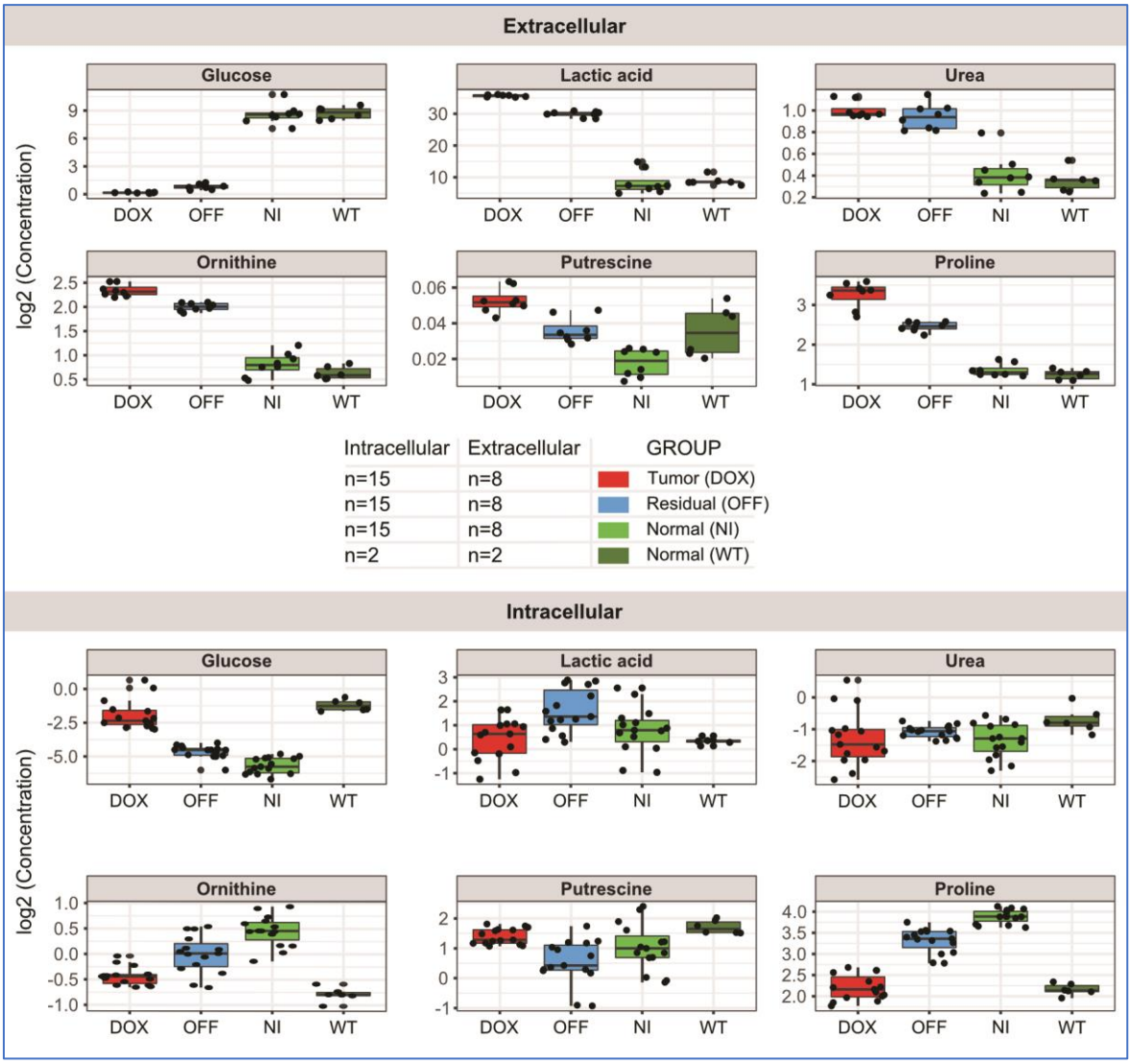

Figure 30: Boxplots showing the most prominent significant (padj<0,01) changes in levels of extracellular and intracellular metabolites between the three populations: tumor (DOX), residual (OFF), normal (NI). The concentration is measured as log2 (area under the curve = AUC).

3.4.4 Multi-omics data integration shows urea cycle and glycolysis as the prominent features of residual cell population

In order to get an overall, systemic view of the changes happening in residual cells, we combined transcriptomic and metabolic data using bioinformatic tools. This included integration of datasets obtained by RNA-sequencing (described in Chapter 3.3) and metabolic measurements acquired by GCMS (described in Chapter 3.4.2 and 3.4.3). GCMS datasets were chosen because the metabolite could be precisely identified. Even though metabolites obtained by untargeted metabolic analysis (Q-Exactive MS, described in Chapter 3.4.2), were not directly integrated because of the limitations imposed by putative annotations, they could still be considered in the conceptual data integration and further hypotheses generation.

In addition, to gain more detailed insight and information about metabolic alterations in tumor and residual cells, reporter metabolite analysis was performed. For this purpose, the Human Metabolic Reaction 2.0 (HMR2), a genome-scale metabolic model containing thousands of genes associated with thousands of reactions and unique metabolites (127), was revised. The Platform for Integrative Analysis of Omics Data (piano, an R package for enriched gene sets analysis) was applied (128) and the analysis was based on the differentially expressed genes where the q-values ($p<0,05$) and log 2 fold changes were taken to predict significantly changed metabolites (reporter metabolites).

Analysis using the piano vignette distinct-directional class was used to obtain directionality of the predicted metabolite changes, by cancelling out gene sets containing both significantly upregulated and downregulated genes. Over Representation Analysis (ORA) in MetaboAnalyst 4.0 was performed on the top metabolites predicted (original p_{adj} <0,1, but maximally 5% of the total list of tested metabolites) to be found in higher levels in residual compared to the normal state. This analysis indicated changes in the metabolic pathways comprising sugar, nucleotide, amino sugar, amino acid, nitrogen and folate metabolism; mitochondrial ETC and TCA cycle; sphingolipid and glycerolipid metabolism; cardiolipin, phosphatidylcholine and phosphatidylethanolamine biosynthesis (Figure 31). Hypergeometric tests were performed to calculate the probability that a certain number of metabolites point to the pathway in the tested metabolites list. The legend indicates the raw p-values.

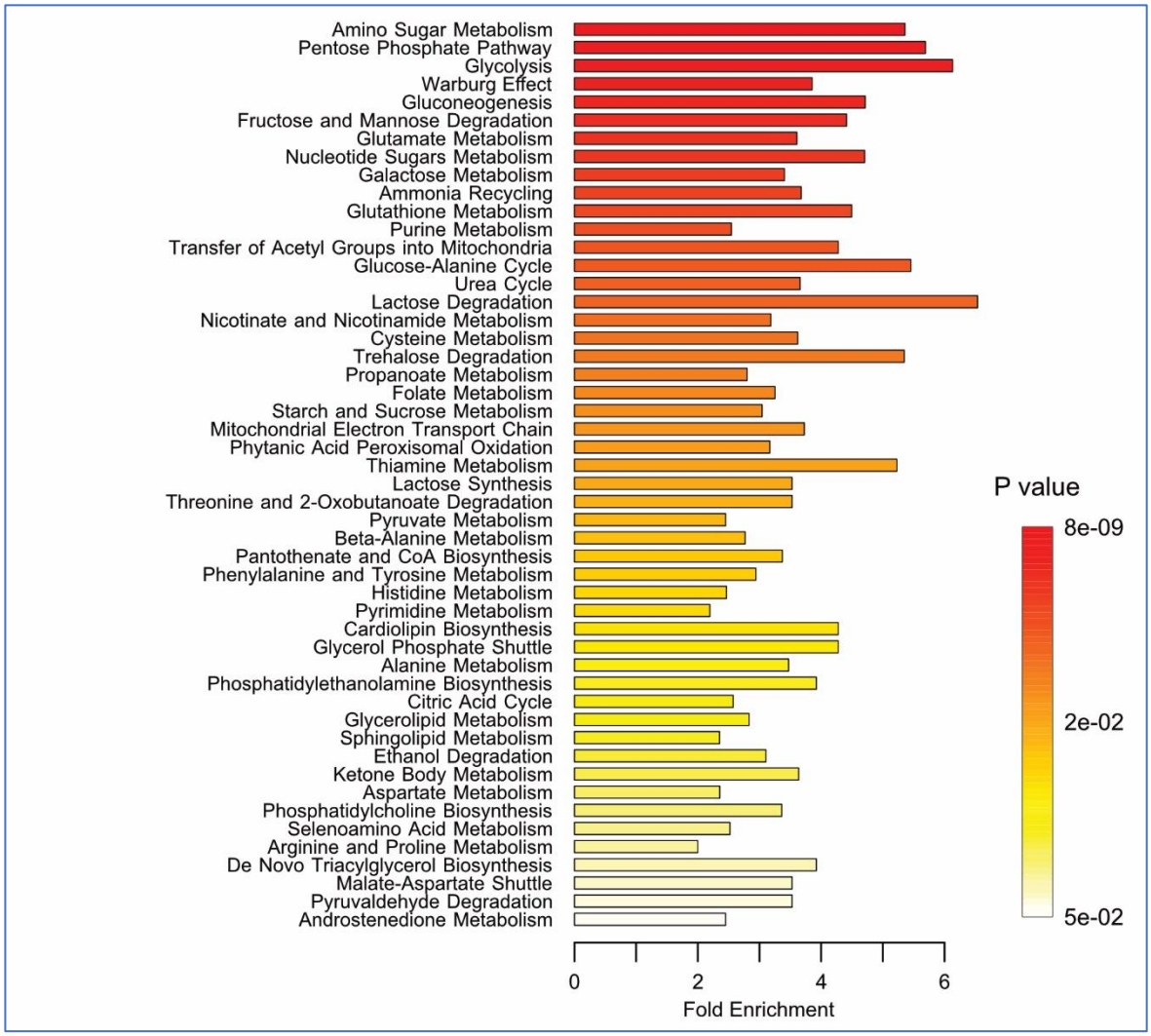

Figure 31: Over Representation Analysis (ORA) based on the top 5 % metabolites predicted to be upregulated in the residual cells

Next, to visualize the global changes in the tumor and residual cells in comparison to the normal, an overlay of reactions inferred from transcriptomics data (see Chapter 3.3) with actual metabolites measured by GCMS (see Chapter 3.4.3) was made. For adding more information to this overall picture, reporter metabolite analysis (described above) was also included: this time taking 5% of changed metabolites according to the p-value after multiple testing adjustment (Benjamini-Hochberg). Global changes in the tumor cell compared to the normal are shown in the Figure 32. The inferred reactions were based on the genes, whose transcript levels were significantly changed. These genes were mapped to biochemical reactions according to the KEGG annotation of their encoding enzymes. When there were multiple enzymes annotated to the same reaction,

with a conflicting sign of change, the sign of the gene with the strongest log2 fold change was depicted. Measured metabolites that were significantly changed (p_{adj}<0,01; ones found to be elevated shown in red, lower in blue) with log2 FC>0,5 and metabolites that were predicted to be changed (reporter metabolites, marked with asterisk, p_{adj} <0,1) in the central metabolic pathways were integrated in the picture. This overlay showed that the tumor cells had high uptake of glucose and high glycolytic flux and redirection to lactate, but at the same time also the increased reactions in tricarboxylic acid (TCA) cycle. Alterations in intracellular abundance of some fatty acids were observed, such as decrease in palmitate and stearate. Intracellular cholesterol was also found in lower amount compared to normal. Increase in abundance of intracellular amino acids was seen, probably in relation to their enhanced anabolism for proliferative needs, reflecting changes typical for cancer metabolism. One novel interesting feature that showed up was an upregulated urea cycle. Extracellular urea levels, but also putrescine, were found to be higher in tumor cells. Urea cycle is independent, but connected to TCA cycle. The branch where arginosuccinate is converted to fumarate that feeds into TCA cycle was enhanced, with excreted fumarate found in higher amount. Another interesting observation were predicted changes in glutathione metabolism and indicated higher conversion of S-adenosyl methionine (SAM) to S-adenosyl-homocysteine (SAH) which could hint towards epigenetic changes.

But what happens when we compare non-proliferative, residual cells to the non-proliferative, normal cells? Quite surprisingly, despite their similarity to never induced structures, they still kept some prominent features of tumor metabolism (Figure 33). This was above all visible in higher glucose uptake, prominently increased glycolysis and lactate export and enhanced urea cycle, visible from upregulated genes but also metabolic measurements showing increased excretion of extracellular urea, ornithine and putrescine. In addition, pathway associated with nitric oxide (NO) production was found upregulated. Overlay also showed deregulated TCA cycle and some changes in metabolism of fatty and amino acids, albeit with differences compared to the tumor. One part of TCA cycle (from isocitrate to succinate) showed downregulation. In another part, succinate, fumarate and malate were accumulated. While palmitate, stearate and cholesterol were still low in the residual cell, we observed the increase in oleate and myristate. Levels of some amino acids were back to the normal, while others (leucine, isoleucine, tyrosine, phenylalanine, histidine and tryptophan) remained elevated. Changes in glutathione metabolism were also seen in residual cells, but conversion of serine into glycine and cystathionine was lower compared to the normal cells (while in tumor this was higher); however, conversion of homocysteine to cystathionine remained high.

This integration revealed the two most interesting pathways that emerged in both tumor and residual cells: glycolysis and urea cycle.

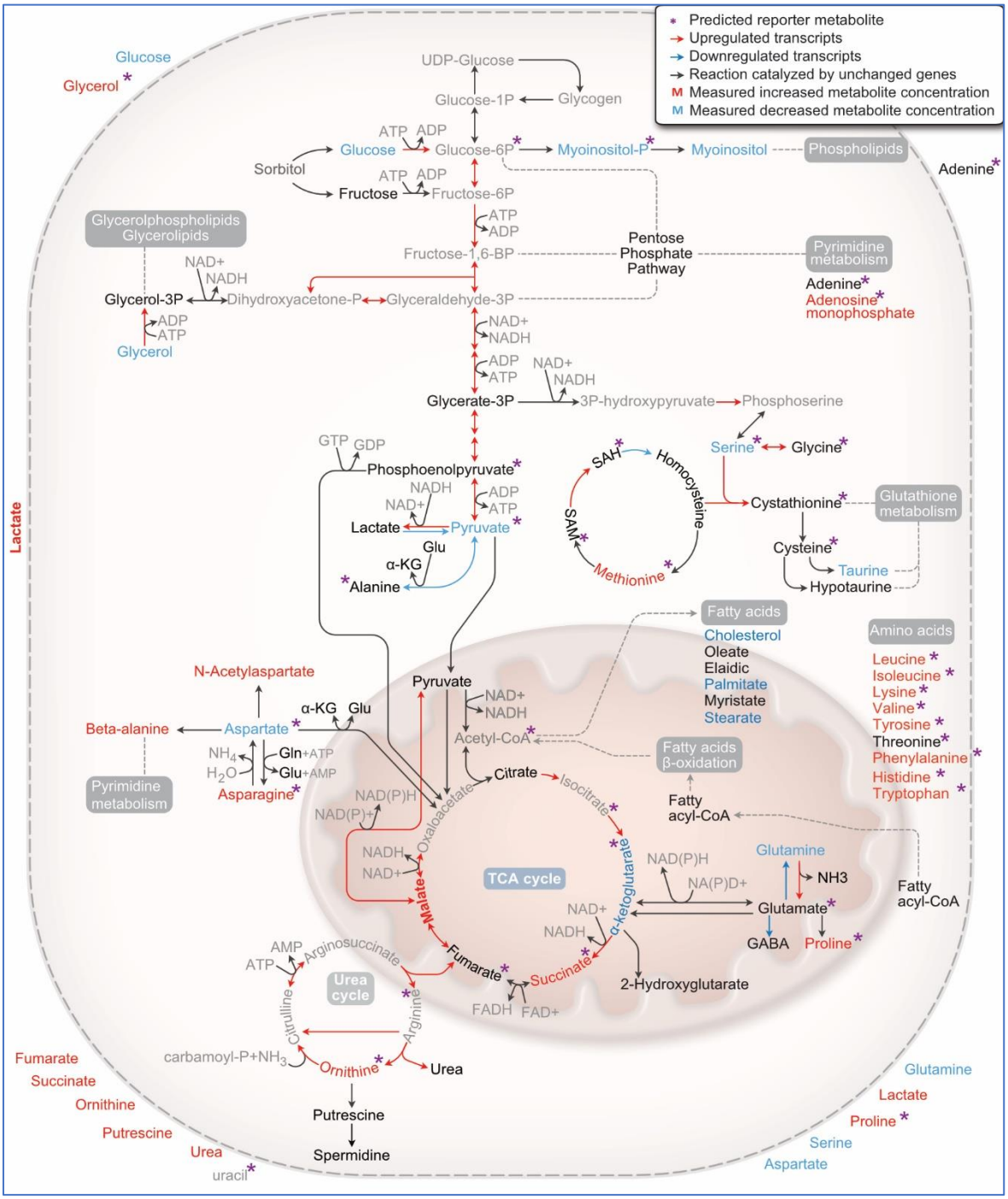

Figure 32: Global overview of altered pathways in tumor versus normal cells. Graphical design was inspired by figure from (129). *Abbreviations are explained in the List of Abbreviations.*

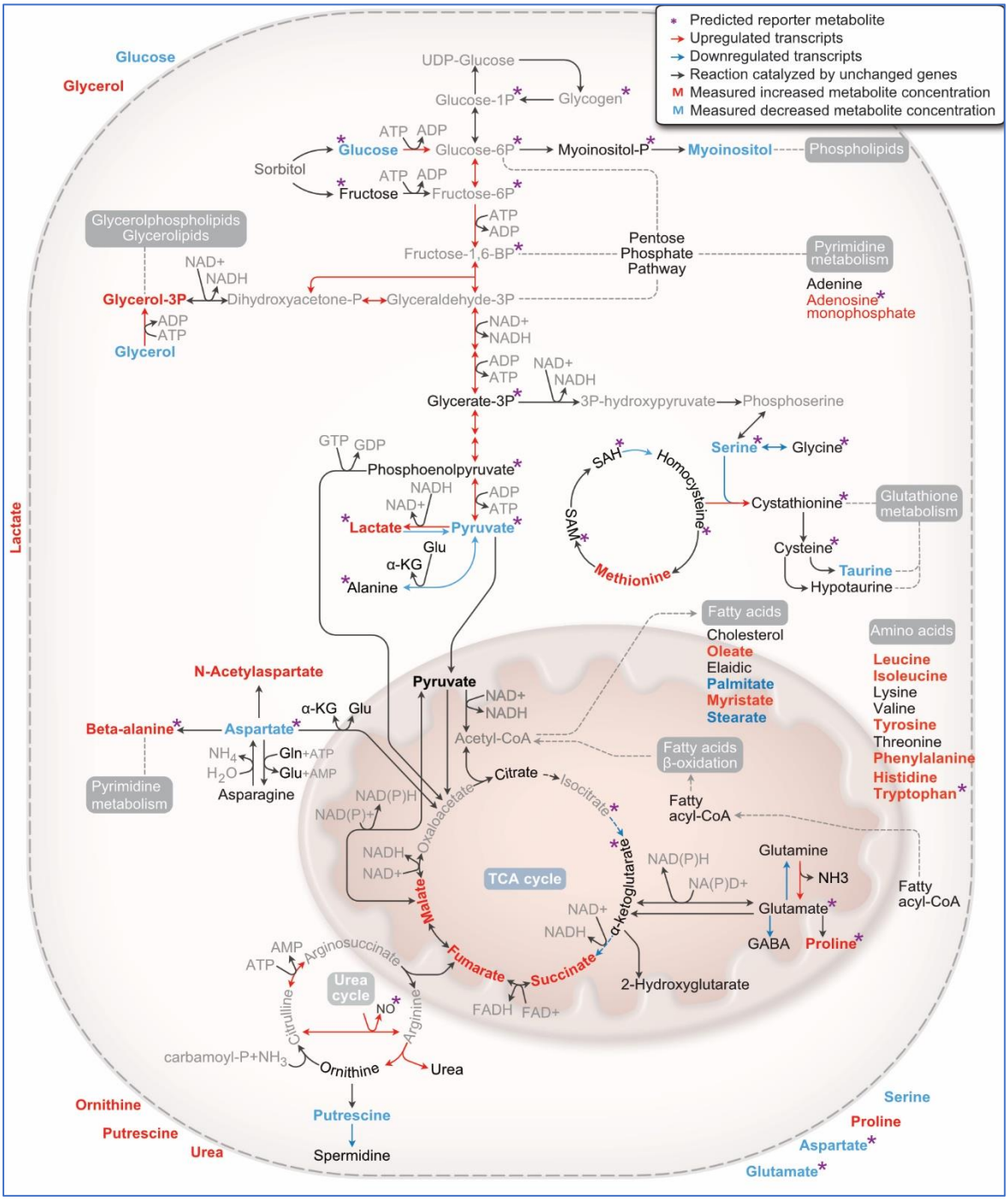

Figure 33: Global overview of altered metabolic pathways in residual versus normal cells. Graphical design was inspired by figure from (129)*. Abbreviations are explained in the List of Abbreviations.*

This helped us narrow down potentially interesting metabolic targets in residual cells, so we looked back to the RNA-Seq data, focusing on genes encoding key enzymes in these pathways (Figure 34).

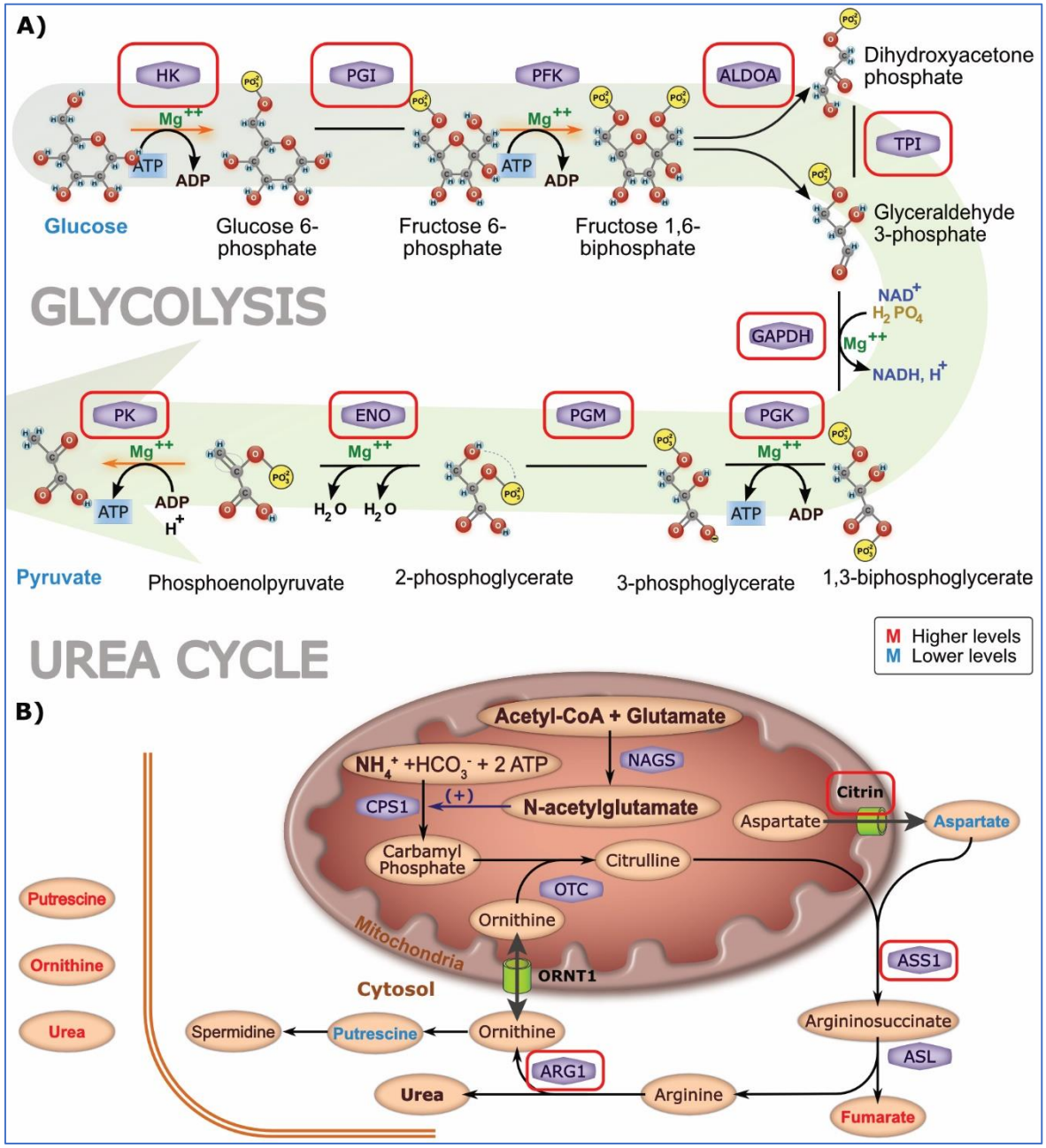

Figure 34: Genes encoding key enzymes found to be overexpressed in residual cells compared to controls (in red). A) Glycolysis; B) Urea cycle. Measured metabolites indicated by color. Schemes of the cycles modified from (130,131). *Abbreviations are explained in the List of Abbreviations.*

Apart from phosphofructokinase (Pfk), all genes encoding enzymes catalyzing steps in glycolysis were upregulated in the residual cells (Figure 34A). Genes encoding key enzymes in urea cycle: arginosuccinate synthase 1 (Ass1) and arginase 1 (Arg1) were also overexpressed compared to the levels of these genes in the normal cells (Figure 34B). In addition, Citrin (Slc25A13) encoding aspartate transporter was upregulated. Log2 FC of the key significantly changed (p_{adj}<0,05; with the exception of Lin28a with original p=0,003 and p_{adj}=1) genes involved in these two pathways (glycolysis marked in yellow, urea cycle in green) are shown in Table 5.

Table 5: Upregulated metabolic genes in the residual cells, encoding important enzymes in glycolysis and urea cycle. Abbreviations are explained in the List of Abbreviations.

Gene	Enzyme	Log2 FC (residual)	Log2 FC (tumor)
Hk2	Hexokinase 2	1,60	2,00
Gck	Glucokinase	2,64	-
Pgi	Phosphoglucose isomerase	0,96	1,52
Aldoa	Aldolase A	1,23	1,28
Tpi	Triosephosphate isomerase	1,38	1,40
Gapdh	Glyceraldehyde 3-phosphate dehydrogenase	1,13	1,38
Pgk1	Phosphoglycerate kinase 1	1,52	1,89
Pgm1	Phosphoglucomutase 1	1,21	1,28
Eno1	Enolase 1	0,66	1,24
Pkm2	Pyruvate kinase	0,77	0,98
Pdk1	Pyruvate dehydrogenase kinase 1	1,73	1,73
Pdk4	Pyruvate dehydrogenase lipoamide kinase isozyme 4	2,31	-
Pfkfb3	Fructose-2,6-biphosphatase 3	3,14	-
Ldha	Lactate dehydrogenase	0,58	0,91
Mct4	Monocarboxylate transporter 4	1,55	-
Lin28a	Lin28 homolog A	3,08	9,59
Ak4	Adenylat kinase 4	2,51	2,89
Ass1	Arginosuccinate synthase 1	0,84	1,43
Arg1	Arginase 1	1,26	1,65
Slc25A13	Citrin	1,26	1,52
iNos	Inducible nitric oxide synthase	6,67	3,36

Overall data integration pinpointed at metabolic similarities between residual and tumor cells. It also hinted at potential metabolic vulnerabilities of residual population, revealing possible points of interference. Some of these genes encoding metabolic enzymes were

found upregulated in both tumor and residual cells and some were exclusively seen in the residual population. Interfering with the latter ones might represent a new opportunity to target significantly upregulated pathways in MRD.

3.5 *In vivo* verification

While our 3D culture system is clean and synchronous, enabling us to follow in detail dynamics of tumor initiation, progression and regression, while providing a reductionist approach without noise and complexity encountered in living organisms, it was important to validate these findings *in vivo*. This was done using tissues of the regressed mammary glands (animals that had full blown tumors, but these tumors regressed after Dox withdrawal) and their corresponding age-matched controls, whereby following experiments were performed:

- obtaining extracellular metabolic profiles using GCMS measurements of metabolite levels excreted after culturing them for 8 hours,
- validating the most interesting enzymes involved in glycolysis and urea cycle (shown in Chapter 3.4.4., Table 5) at the protein level using IF stains,
- measuring enzyme activity of NOS in the tissues,
- measuring flux to lactate after culturing for 8 hours in glucose- and pyruvate-free media supplemented with ^{13}C glucose.

All the age-matched controls were healthy, wild-type animals, that were also fed with doxycycline supplemented food, to exclude the potentially confounding effect.

GCMS measurements of the extracellular metabolites from the regressed and healthy tissue confirmed our observation from 3D cultures (see Chapter 3.4.2, Figure 25B) – even after tumor regression, 9 weeks of doxycycline, mammary glands (Figure 35A, shown in blue) still clustered away in their extracellular metabolic profiles from the healthy glands of age-matched controls (Figure 35A, shown in green). PC1 explained 63,6% of variance, while PC2 explained 16,9%.

We measured 37 metabolites by GCMS. The most prominent changes ($p_{adj}<0,05$) were seen in the levels of: urea, ornithine, putrescine, proline and glycine that were higher in the extracellular media of the regressed glands and pyruvate that was lower (Figure 35B). Higher levels of urea, ornithine, putrescine and proline were in concordance with

the extracellular profiles obtained from the 3D cultures (see Chapter 3.4.3, Figure 30), hinting at changes connected with urea cycle.

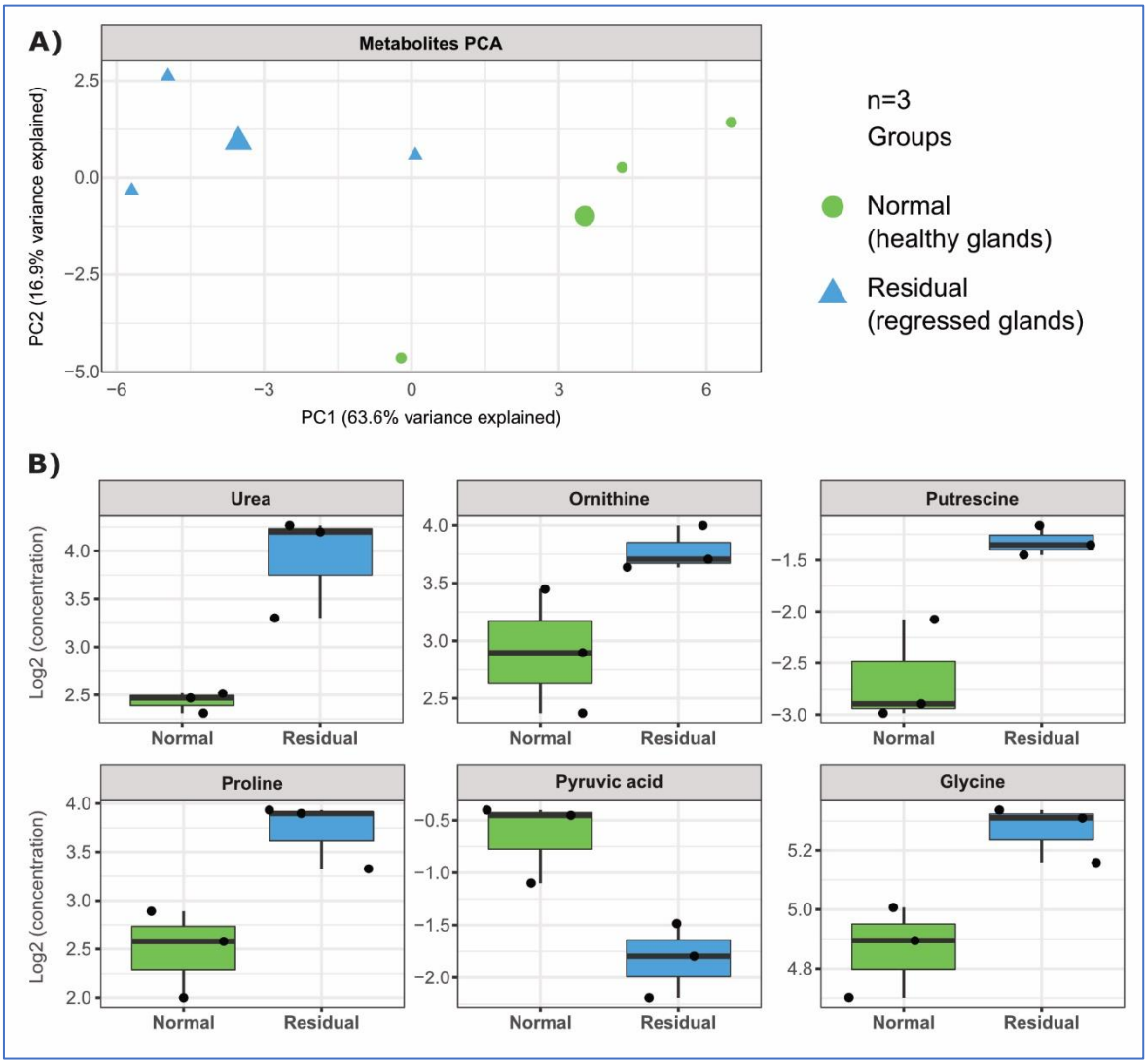

Figure 35: Extracellular metabolic profiles from regressed and normal mammary glands measured by GCMS. A) PCA plot; B) Selected significantly changed metabolites (LIMMA, p<0,05). The concentration is measured as log2 (AUC).

We observed a trend in the increase of extracellular lactate and decrease of extracellular glucose in the regressed glands compared to normal, which would be consistent with our findings from 3D cultures, but the sample variation in these two metabolites was quite large in this experiment (Appendix, Figure 54). Metabolites that were not in agreement with the ones measured in 3D cultures were glutamine and aspartate that were higher

extracellularly in the residual samples (regressed samples, Appendix, Figure 54) as opposed to *in vitro* measurements where they were lower compared to the normal samples (see Chapter 3.4.3, Figure 29). Other metabolites (that were not measured *in vitro*) found to be significantly higher when residual (regressed tissue) was compared to the normal included: 2 keto-glutaric acid, 4-hydroxyproline, asparagine, taurine and uracil (Appendix, Figure 54).

Next, we wanted to validate upregulated targets (see Chapter 3.4.4, Table 5). From this list, 8 genes were selected to be checked on protein level by IF staining performed on the formalin fixed, paraffin embedded (FFPE) tissues. The choice was driven by the availability of specific antibodies, known roles of these proteins and their potential significance as targets for interference. The following targets were further tested:

- HK2 that catalyzes the first step of glycolysis and is isoform commonly found overexpressed in many cancers;
- PDK1 that phosphorylates and inhibits pyruvate dehydrogenase (PDH), thus preventing conversion of pyruvate to acetyl-CoA;
- MCT4, a transporter that exports lactate from the cell;
- AK4 that directs metabolism towards aerobic glycolysis;
- LIN28A that acts as a glycolytic switch, favorizing aerobic glycolysis;
- ASS1 that catalyzes conversion of aspartate to arginosuccinate;
- ARG1 that catalyzes the last step of urea cycle, i.e. conversion of arginine into ornithine and urea;
- iNOS (NOS2) that catalyzes conversion of arginine to nitric oxide.

To get an overview of the whole section, slides were scanned using TissueFAXS. StrataQuest5, the corresponding program for the analysis, was used for the quantification of number/percentage of positive cells and the intensity of the signal in the selected ducts.

From the above-mentioned targets, HK2, AK4, LIN28A and iNOS showed very slight differences in their levels between the regressed and healthy tissue. However, strikingly significant differences were observed when the levels of ARG1 were compared (Figure 36).

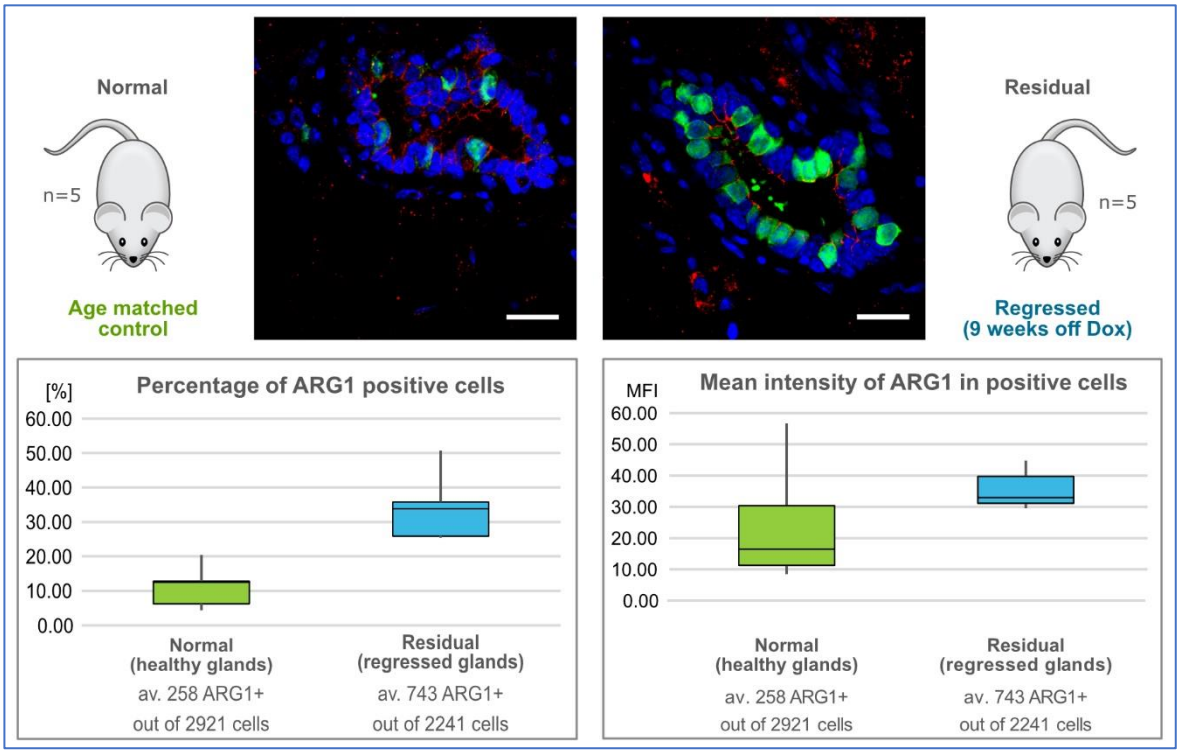

Figure 36: Arginase1 verification on the tissues of animals 9 weeks off doxycycline (regressed mammary glands) compared to healthy age-matched controls. Top panel: IF stains – E-cadherin in red, ARG1 in green; nuclei stained with DAPI (blue). Scale bar: 50 μm. Below: Percentage (left) and mean intensity (right) of ARG1 positive cells. Abbreviations are explained in the List of Abbreviations.

Even though ARG1 was present in the healthy tissue as well, it was obviously more abundant in the regressed glands, as determined by quantification: percentage of ARG1 positive cells was significantly higher (unpaired homoscedastic one-tailed t-test, $p<0,001$) in the group of residual cells. The trend towards increased mean intensity of ARG1 was seen in the regressed ducts.

IF staining of HK2 did not reveal prominent differences between the regressed and normal glands, but we did observe a tendency towards its slightly enhanced levels in the regressed glands (Figure 37). The same trend towards slightly higher levels in the regressed glands was observed in IF stains for LIN28A, AK4 and iNOS (Appendix, Figure 55). Differences were not seen in the levels of PDK1 (Appendix, Figure 56A). Slight decrease of ASS1 was observed in the regressed glands (Appendix, Figure 56B). Unfortunately, it was not possible to detect the chosen MCT4 antibody on the tissue sections.

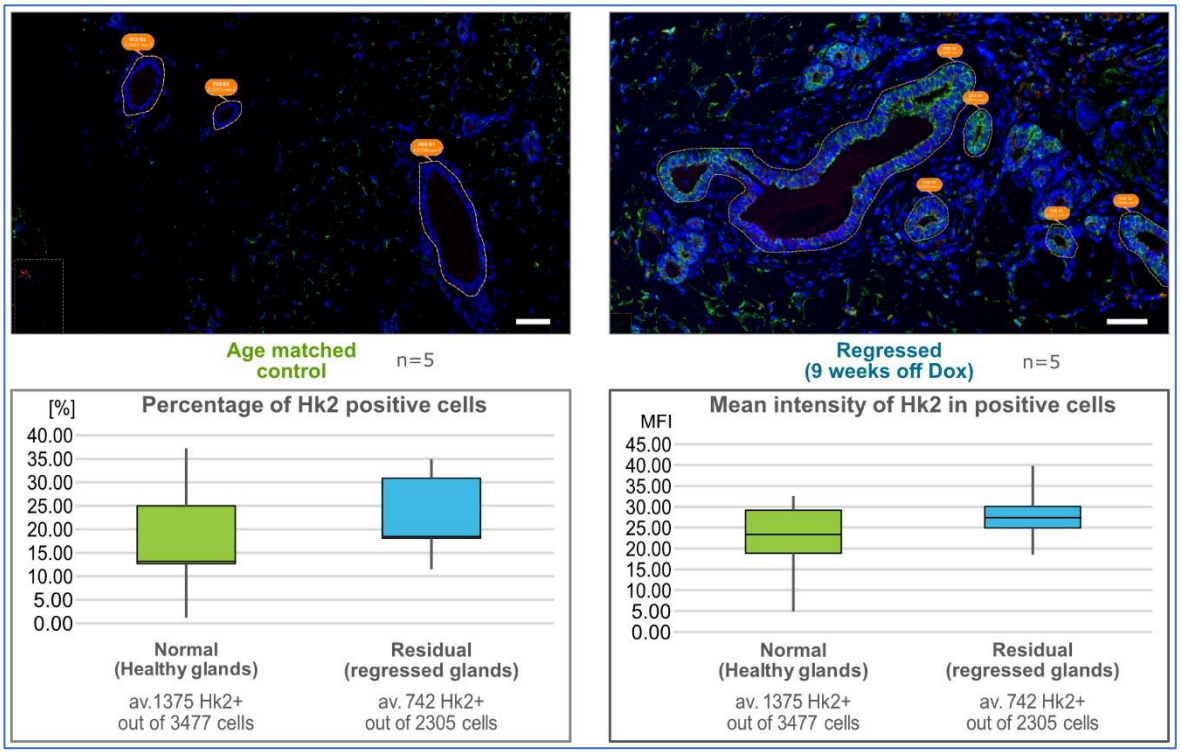

Figure 37: IF staining of HK2. CDH1 (E-cadherin) in red, HK2 in green; nuclei stained with DAPI (blue). Some examples of regions of interests (ROIs) are shown. Scale bar: 50 µm. Below: Percentage (left) and mean intensity (right) of HK2 positive cells. Abbreviations are explained in the List of Abbreviations.

Since NOS2 was found strongly upregulated on the transcriptional level in the residual samples *in vitro* (see Chapter 3.4.4, Table 5) and slightly, but not significantly upregulated on the protein level *in vivo* (Appendix, Figure 55C), we additionally performed NOS assay to check its specific enzymatic activity. We measured significantly increased (unpaired homoscedastic one tailed t-test, $p<0,03$) NOS activity in the regressed samples compared to the age-matched controls (Figure 38A).

As enzymes involved in glycolysis showed upregulation on the transcriptional level, starting from HK2 that catalyzes its initial step to LDHA which catalyzes end reaction converting pyruvate to lactate (see Chapter 3.4.4, Table 5), we wanted to verify increased glycolytic flux towards lactate production in the residual cells. Thus, we performed flux analysis using labeled ^{13}C glucose. This experiment confirmed statistically higher (Wilcoxon, Mann-Whitney test; $p<0,05$) conversion to lactate in the regressed mammary glands (Figure 38B).

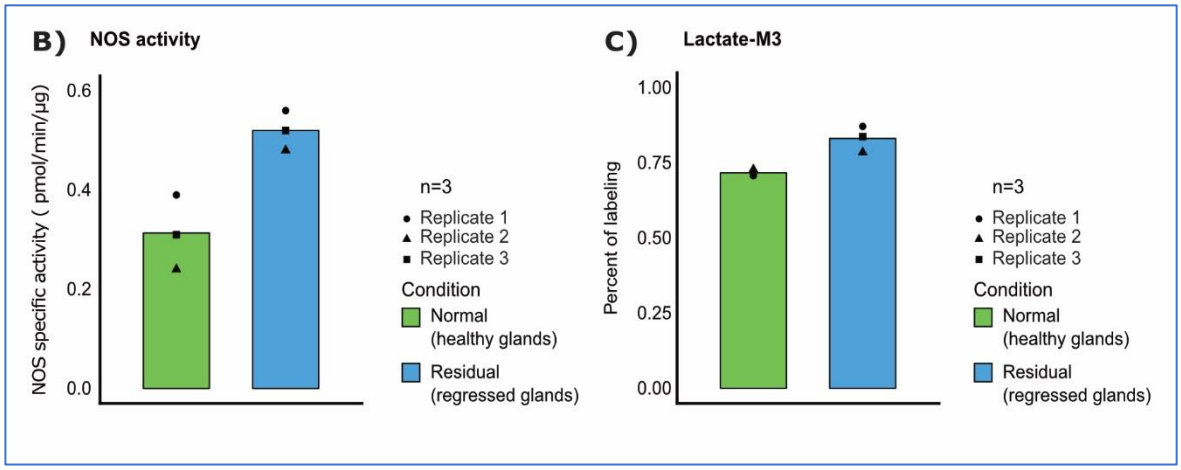

Figure 38: In vivo intracellular metabolic verifications on the regressed and healthy mammary glands. A) NOS enzymatic assay; B) Levels of lactate labeling by ^{13}C glucose

3.6 CORRELATION OF THE MAIN FINDINGS WITH HUMAN DATASETS

After verifying the most prominent findings hinting at alterations in glycolysis and urea cycle/arginine metabolism *in vivo*, we further explored their relevance in the human breast cancer. This was done through the comparison of microarray datasets obtained from 21 patient samples (132) collected before the treatment (fine needle aspiration, tumor cells) and after application of neoadjuvant systemic therapy (resected after 4-6 months, i.e. post-treatment, residual) with the normal tissues from 9 healthy women (133). After data preparation, its filtering and exclusion of normal-like subtype due to high biological variation and poor diagnostic criteria, PCA was performed (Figure 39) and it showed clustering that reflected the pattern previously observed *in vitro*, from our RNA-seq data (see Chapter 3.3, Figure 17B). Pre-treatment (tumor, shown in squares), post-treatment (residual, shown in triangles) and healthy (normal, shown in circles) samples clustered separately.

Furthermore, groups that were different among each other with respect to the tumor type (indicated by different symbols) and treatment, still clustered together. The first principal component explained 14,6% of variance and the second 11,4%.

Clustering of transcriptomic signature obtained from our mouse model with specific gene expression profiles from the patient samples revealed its closeness to basal HER2 positive sample (communication with Katharina Zirngibl). We next looked into differentially expressed genes between post-treatment (residual) and normal samples,

involved in glycolysis and urea cycle, to check if the targets we identified in mice would be also differentially expressed in the publicly available human sample datasets of MRD.

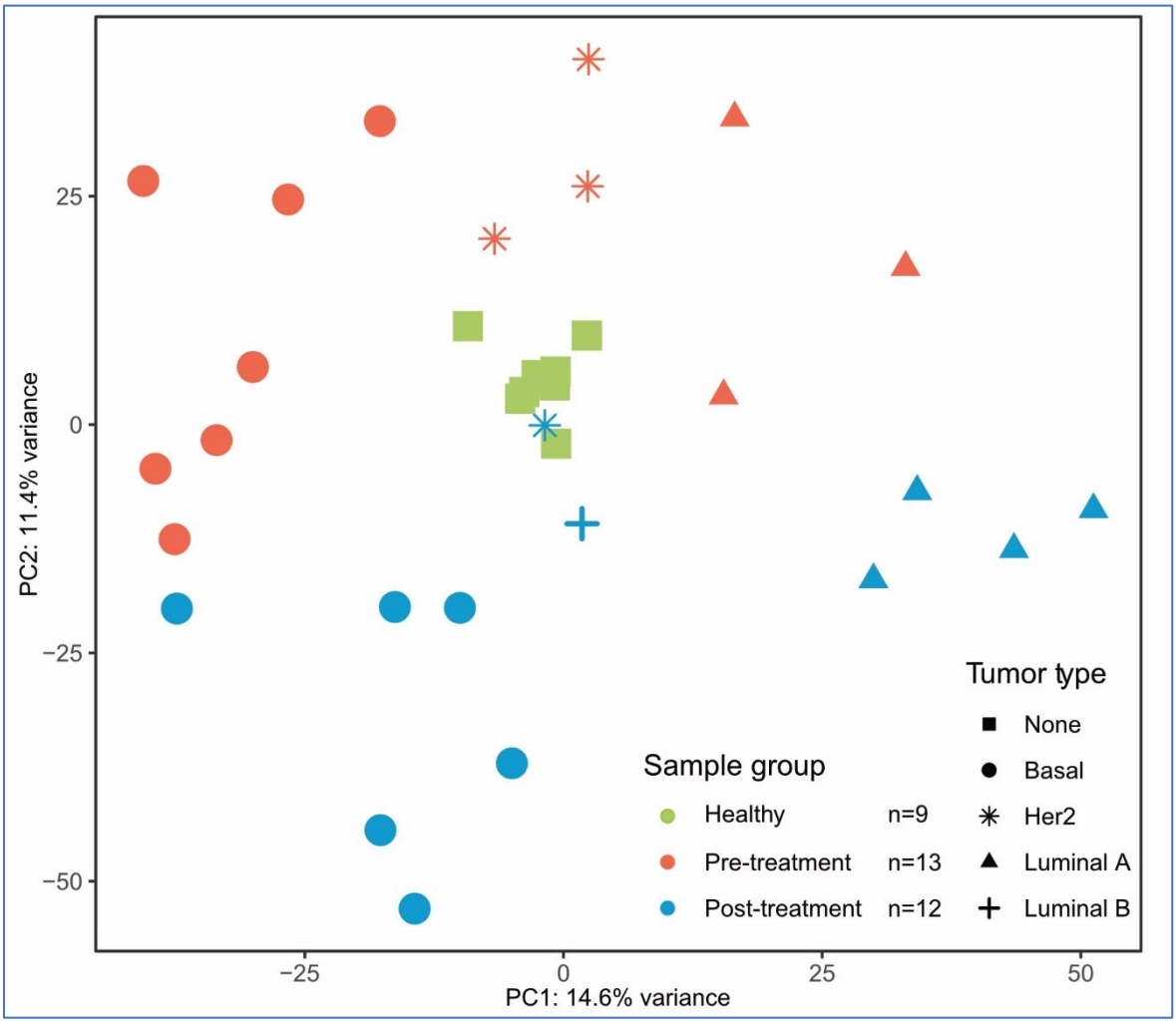

Figure 39: PCA performed on microarray datasets from breast cancer patients (pre- and posttreatment) and healthy women. Abbreviations are explained in the List of Abbreviations.

Strikingly, the following genes encoding enzymes involved in urea cycle and nitric oxide (NO) synthesis upregulated in our mouse model were also significantly (p_{adj}<0,1) upregulated in the post-treatment sample of this subtype compared to the controls (Figure 40, log2 FC shown in colors): ASS (6.3.4.5), ARG (3.5.3.1) and NOS (1.14.1339).

Besides overexpression of HK (2.7.1.1), PFK (2.7.1.11) and LDH (1.1.1.27), other glycolytic enzymes found upregulated in mouse were not found significantly changed

(Figure 41). Still, increased gene expression of these genes indicated changes in glycolysis and upregulation of LDH in particular hinted at redirection of glycolytic flux towards lactate production.

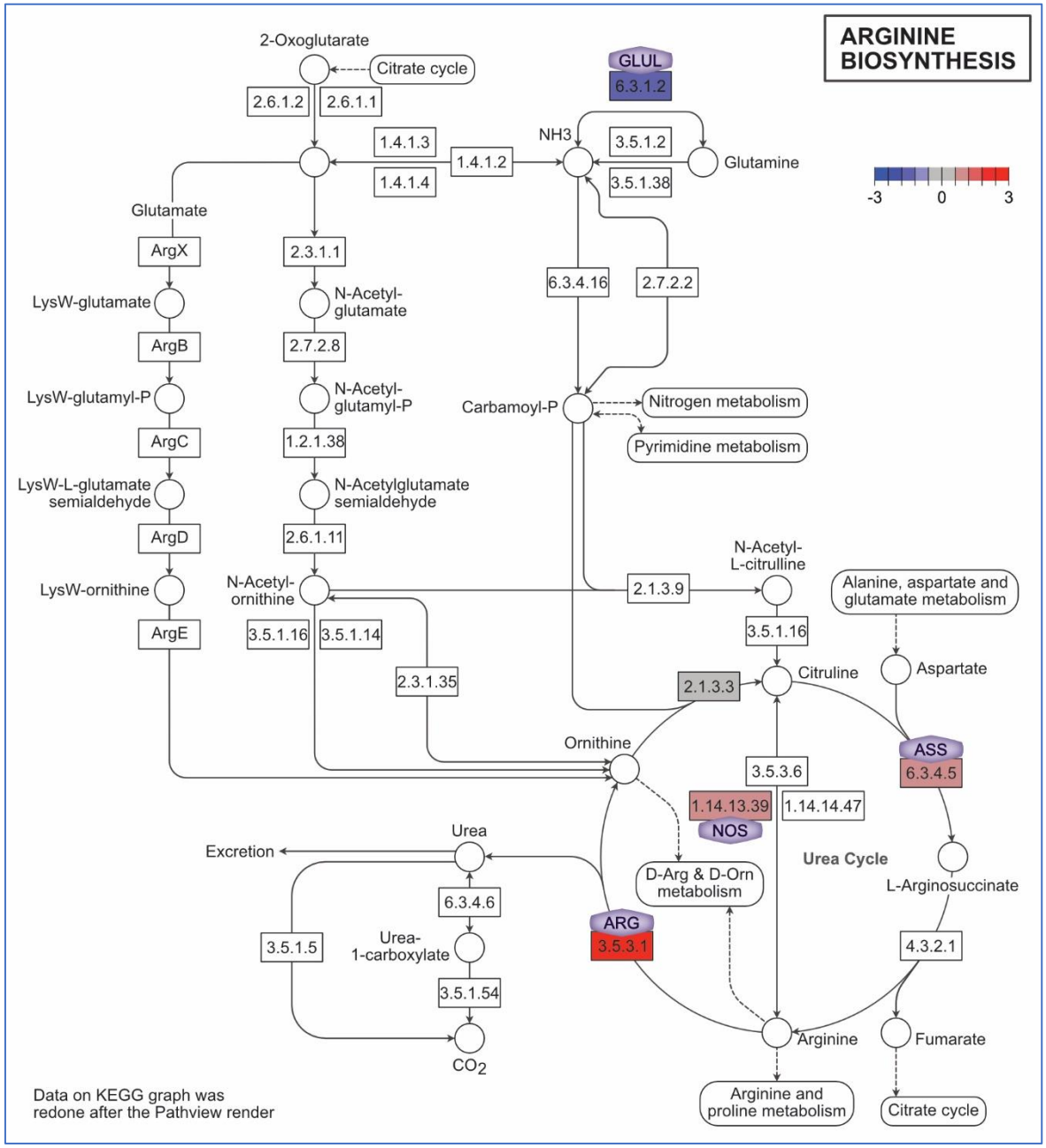

Figure 40: Altered gene expression encoding enzymes involved in arginine metabolism (urea cycle, NO synthesis) in basal HER2 positive breast cancer subtype sample. Pathway map was adapted from KEGG pathway database (134)*. Abbreviations are explained in the List of Abbreviations.*

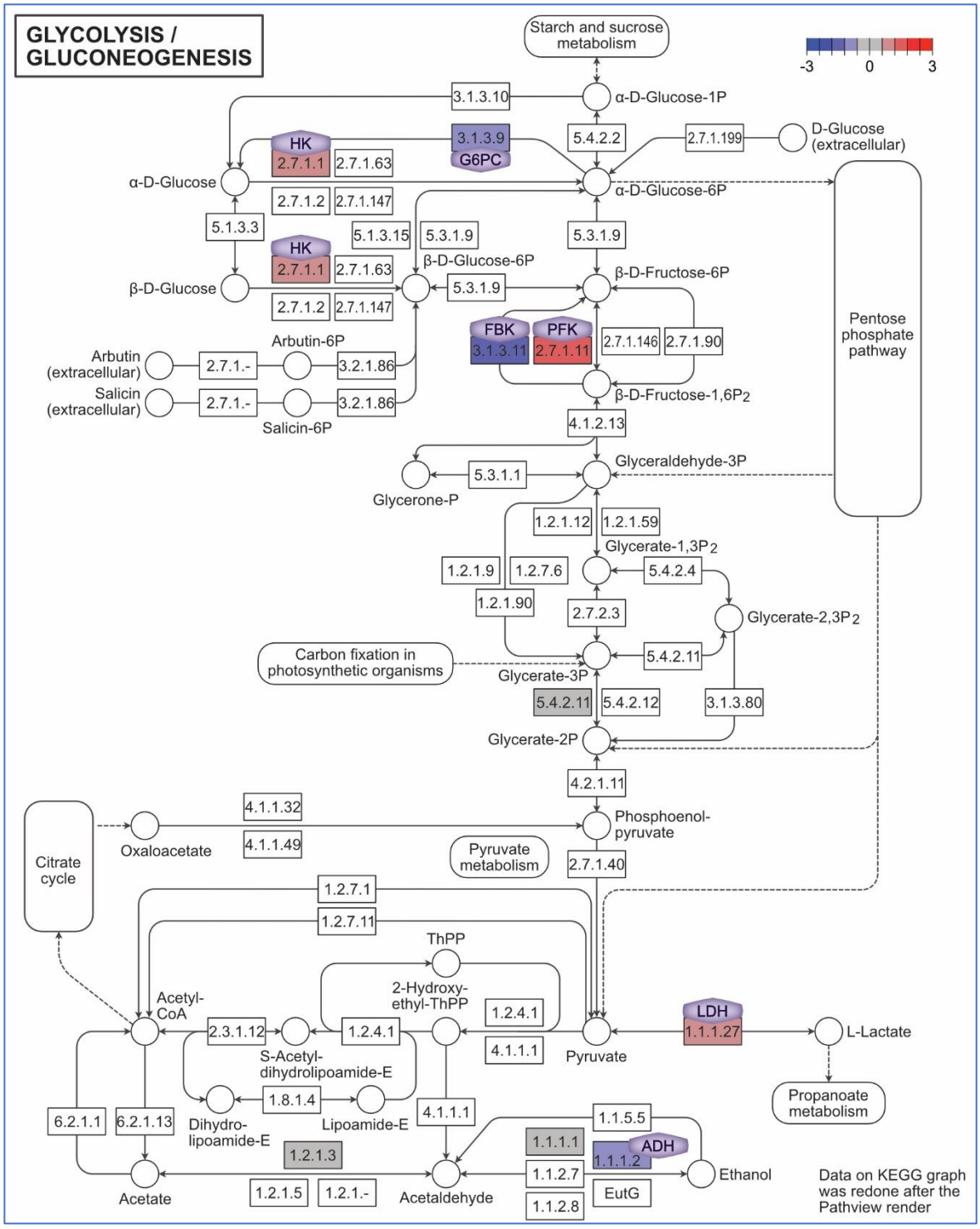

Figure 41: Altered gene expression encoding enzymes involved in glycolysis in basal HER2 positive breast cancer subtype sample. Pathway map was adapted from KEGG pathway database (134). *Abbreviations are explained in the List of Abbreviations.*

However, we could not draw the conclusions about this particular subtype (basal HER2 positive) as it was represented with only one sample in the dataset. Nevertheless, the next closest subtype, basal HER2 negative, showed upregulation of ASS (6.3.4.5) (Appendix, Figure 57) and alterations in glycolysis (Appendix, Figure 58) as well, having overexpressed genes encoding GAPDH (1.2.1.12), ENO (4.2.1.11) and LDH (1.1.1.27). As altered glycolysis was not seen in luminal A and luminal B subtypes, it could be that this feature was more specific for basal HER2 positive and negative breast cancer subtypes.

Overall, human samples of HER2 negative and positive subtypes showed higher levels of enzymes involved in urea cycle/arginine metabolism: ASS, ARG, NOS. The glycolytic genes found overexpressed in these subtypes were: HK, PFK and LDH.

We identified glycolysis and urea cycle as the most prominent altered pathways in the mouse residual cells (see Chapter 3.4.4), which we further verified *in vivo* through metabolic measurements on the mouse regressed (compared to mouse normal) tissue, measurement of flux towards lactate, IF stains of selected proteins and enzymatic NOS assay (see Chapter 3.5). Measured metabolites and flux hinted towards alterations in these pathways and IF stains and enzymatic assays verified the most interesting targets – ARG1, HK2 and iNOS being the most important ones, as they were verified on several levels (all three transcriptionally *in vitro*; all three on protein levels *in vivo* and NOS through enzymatic activity). Furthermore, the relevance of these pathways and targets was verified in human datasets of MRD which, for basal HER2 negative and positive subtypes showed changes in glycolysis and urea cycle. In this regard, basal HER2 positive sample was particularly interesting as it showed the upregulation of ASS, ARG, NOS, HK and LDH that we identified, using our mouse model, as potentially important targets of MRD.

3.7 LIPIDOMICS ANALYSES REVEAL CHANGES IN LIPID COMPOSITION IN RESIDUAL POPULATION

Seeing the changes in the transcriptome and metabolic profiles of residual cells, we were interested if this would hold true for another class of compounds – lipids. As lipid metabolism clearly emerged from predictions based on RNA-Seq and array data as a characteristic of residual population in our previous study (119), we employed lipidomics to validate these findings on metabolic level and to expand our picture of minimal residual disease.

Metabolic method that we established for harvesting metabolites from 3D cultures proved successful for the subsequent lipidomics analyses. Extraction of lipids was done using Bleigh-Dyer protocol (chloroform-methanol-37% HCl). Membrane- and protein-bound, as well as free lipids were detected by "shotgun lipidomics" using Qtrap 6500+ coupled with NanoMate. In total, 227 lipids were measured from intracellular samples.

PCA on lipidomics data showed distinctive profile of residual population (shown in blue), from normal (shown in green) and tumor (shown in red), similarly as was seen on the transcriptomic level (Figure 42A). The first principal component explained 51% of variance and the second 22%.

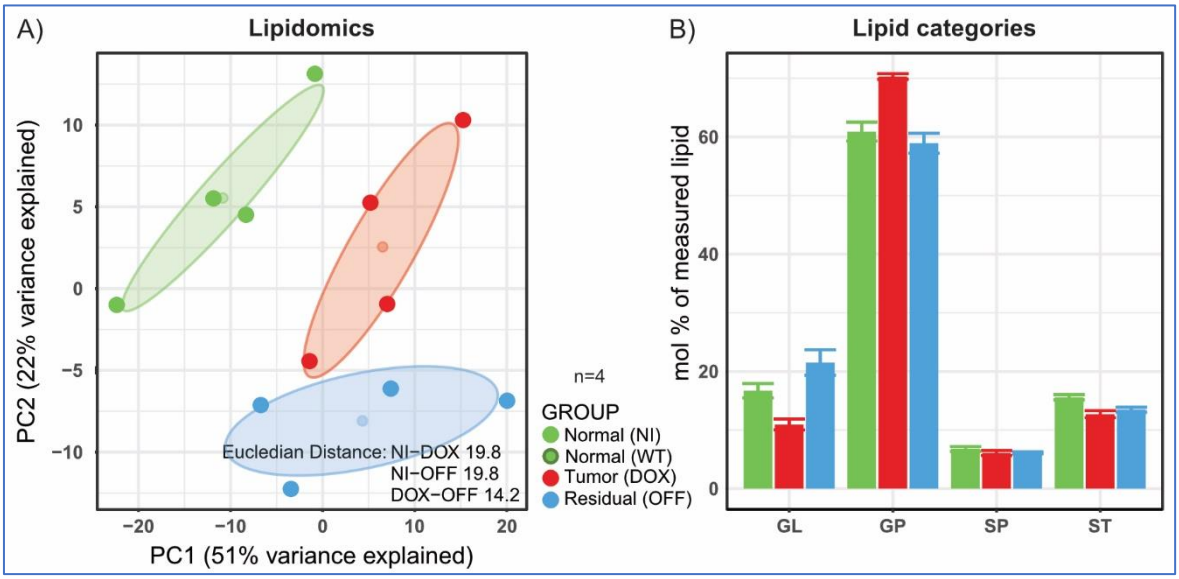

Figure 42: Differences in lipid composition between normal (NI), tumor (DOX) and residual (OFF) populations. A) PCA plot based on measured lipids; B) Measured lipid categories. Error bars represent Mean ± 1SE. Abbreviations are explained in the List of Abbreviations.

Based on these data, centroids were calculated, showing the center of each sample group on the plot. Euclidean distance between the centroids showed that the tumor and residual samples were still closer (DOX-OFF 14,2) than normal and residual (NI-OFF 19,8). Confidence interval of a multivariate normal distribution is represented by the ellipses (CI = 0,95).

Figure 42B shows measured lipid categories: glycerolipids (GL), glycerophospholipids (GP), sphingolipids (SP) and sterol lipids (ST). Statistically significant differences (t-test; $p<0,05$) between the residual and normal populations were seen in slightly lower mol percentage of sterol lipids in the residual cells. When compared to the tumor, residual cells showed significantly higher amount of glycerolipids and lower amount of glycerophospholipids. Comparison of tumor cells to normal showed the lower molar percentage of sterol lipids and glycerolipids and higher molar percentage of glycerophospholipids in the tumor cells.

The measured lipid classes are shown in the Figure 43A: phosphatidylcholine (PC), lysophosphatidylcholine (LPC), phosphatidylethanolamine (PE), phosphatidylserine (PS), phosphatidylinositol (PI), phosphatidylglycerol (PG), phosphatidic acid (PA), diacylglycerol (DAG), ceramide (Cer), sphingomyelin (SM), hexosylceramide (HexCer), dihexosylceramide (Hex2Cer), cholesterol (Chol), cholesteryl ester (CE) and triacylglycerol (TAG). The prefix "O" indicates that there is an alkyl ether substituent in the marked species; "P" stands for plasmalogen substitute. A statistically significant (marked with asterisk) increase was observed in PG species in the residual cells when compared to normal. Significantly lower levels of LPC, PS, PG O-, PA, HexCer, Hex2Cer and cholesterol were seen in the residual population compared to normal. When compared to the control, tumor cells also exhibited significantly lower levels of PS, PA, HexCer, Hex2Cer and cholesterol, but higher TAG and PC, PE, PI and Cer (Appendix, Figure 59). In comparison to the tumor state, residual cells had lower levels of PC, PE, PI, PG O- and Cer and higher levels of PG and TAG (Appendix, Figure 60).

When looking across all the lipid classes at the level of lipid species, we observed many statistically significant changes in abundance of certain lipid species. This was true also for the lipid species within the classes which did not show overall changes in abundance. For instance, we observed a shift from short to long chain fatty acids in sphingomyelin species (Figure 43B), which could be indicative of some changes in membrane composition and its stiffness/rigidity. This was also seen in the lipid species within ceramides. Another interesting change is the depletion of arachidonic acid species in phosphatidylethanolamines class (Appendix, Figure 61).

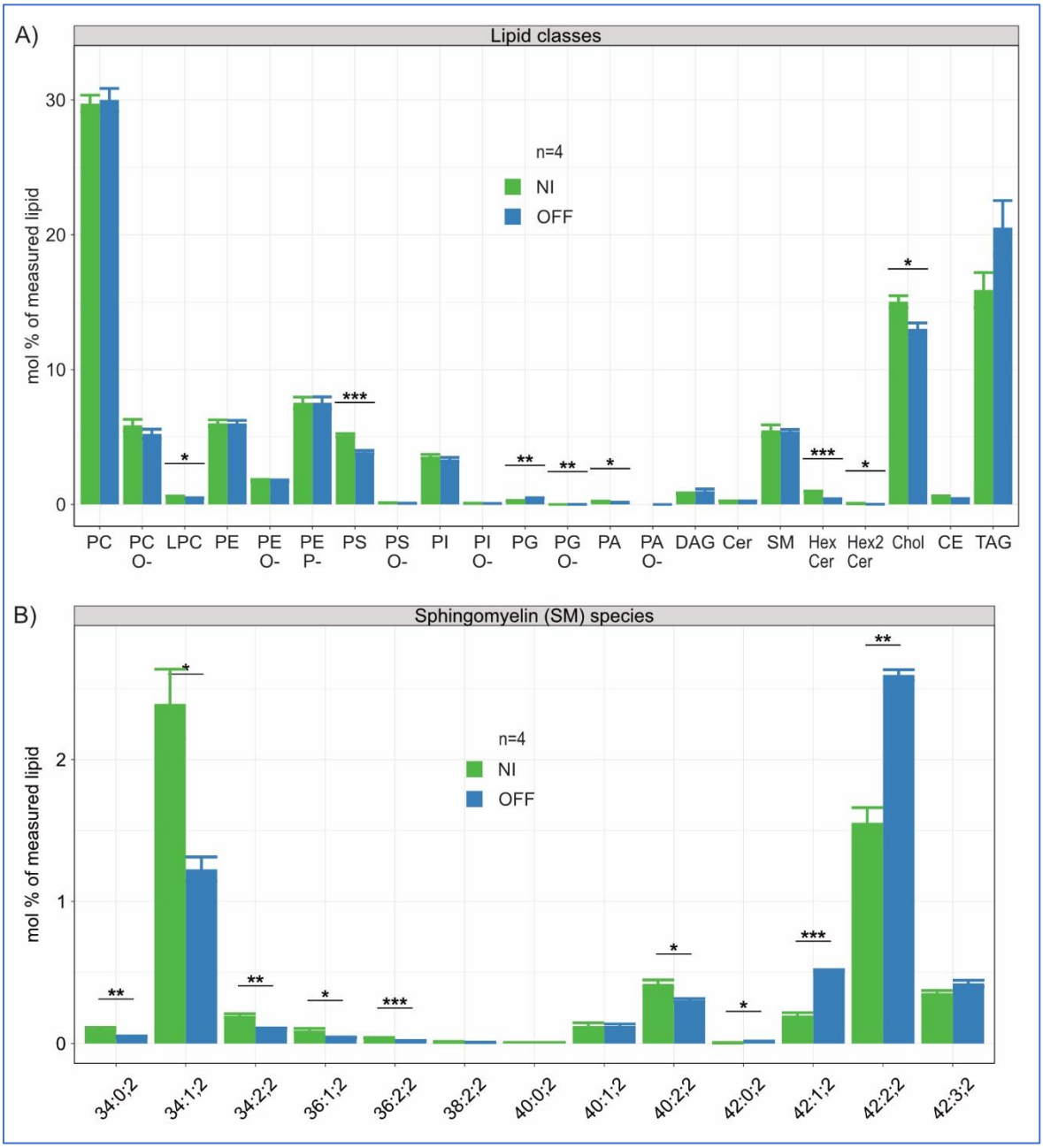

Figure 43: Changes in lipid composition between residual and normal population. A) Measured lipid classes; B) Shift from short to long chained sphingomyelin (SM) species in the residual cells. Error bars represent Mean ± 1SE. Abbreviations are explained in the List of Abbreviations.

In order to further verify the most prominent lipid features and potentially interesting lipid biomarkers in residual population, we attempted to set up and employ imaging mass spectrometry (IMS). IMS is an exciting new tool for obtaining information on spatial

distribution of metabolites within a sample. This particular method was chosen as it would allow to discriminate between the signal that comes from mammary gland epithelial structures and the surrounding adipose and stromal tissue.

Mammary glands were embedded in gelatin (as this material was compatible with IMS), snap-frozen and cut into sections that were dried in vacuum desiccator and covered with matrix that would assist the ionization. As the laser was moving across the sample, it collected the mass spectra, whereby ion intensities were plotted against the relative position of the analytes from the sample. Analysis was done using bioinformatic framework for FDR-controlled metabolite annotations for high-resolution IMS (135). Workflow is summarized in Figure 44A. Specific mammary gland composition, i.e. high percentage of fat (70-80%) and incompatibility of freezing temperatures for surrounding adipose tissue, ducts and embedding media proved to be a great challenge for cryo-sectioning. This led to the suboptimal morphology of the samples (Figure 44B, upper panel) that could affect the final picture of metabolites' spatial distribution. Thus, great efforts were invested into preparing the sample for subsequent analyses. This was improved through use of certain fixatives and cryoprotectants that significantly preserved the morphology of the ducts (Figure 44B, lower panel). The drawback to this was that the sample treatment negatively affected metabolite detection compared to the freshly frozen samples; the number of annotated metabolites at FDR 0,2 was around three times lower in the treated samples (Appendix, Figure 62A,B). However, for the purpose of targeted metabolomics, method has the potential to be further optimized through fine-tuning of other parameters (choice of matrix, mode, laser power etc.). As we were mostly focused on detecting lipid metabolites such as the ones potentially corresponding to stearic acid and phosphatidylserine (shown in Figure 44C) this method could indeed be a great target approach to validate lipid biomarkers *in vivo*. However, detection of other classes of compounds is also possible (Appendix, Figure 62C), opening possibility for spatial imaging of other potentially interesting targets.

Lipidomic analyses demonstrated changes in lipid profiles of the residual cells. Future verification of these findings *in vivo* could yield novel biomarkers and provide new insights into the role of lipids in MRD. Since normal and regressed mammary gland phenotypically look very similar, with appropriate standards, normalization and relative quantification it would be possible to compare the abundance and spatial distribution of desired molecules between the two populations using IMS.

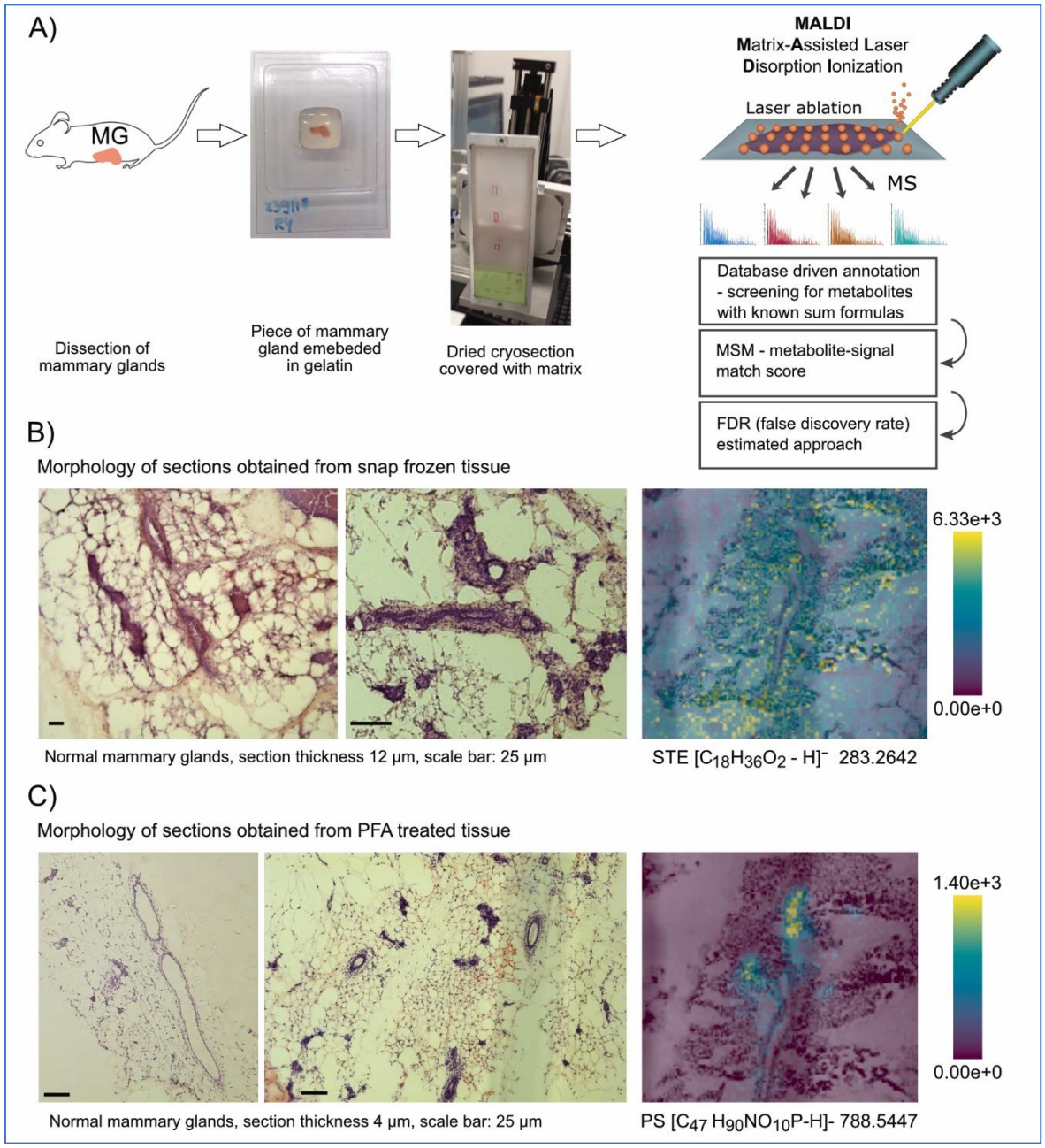

Figure 44: Setting-up spatial metabolomics - imaging mass spectrometry for in vivo verification of potential lipid biomarkers. A) Workflow; B) Comparison of morphology of snap-frozen and PFA-treated tissue; C) Examples showing detection of metabolites corresponding to stearic acid and phosphatidylserine (PS) in the regressed gland. Abbreviations are given in the List of Abbreviations.

3.8 METABOLIC MEMORY IN RESIDUAL CELL POPULATION – KEEPING THE MAIN METABOLIC FEATURES OF TUMOR CELLS

Since prominent features of metabolism characteristic for tumor cells were preserved in the residual population also *in vivo* even 9 weeks after oncogenes deactivation, we referred to them as "metabolic memory". But when is this memory established and how is it kept?

We hypothesize that there is a threshold, an early point in neoplastic transformation where cells still have the flexibility to return to the normal state. From our initial experiments of deactivating oncogenes earlier (i.e. mimicking the treatment in earlier stages), we noticed that there were differences in the response of the cells to doxycycline withdrawal when they had been kept on doxycycline for shorter time (60 hours) than usual 5 days: it seemed that they were not dying, but rather simply reintegrating into the existing rim. Thus, we hypothesize that the cells that were shortly exposed to oncogene action had the plasticity and capacity to restore their normal phenotype. To explore this further, we decided to follow the dynamics of oncogenesis (from initiation over progression to full blown tumor structures) and regression in detail.

3.8.1 Phenotypes during oncogenesis and tumor regression

In order to track the dynamics of oncogenesis and regression in our synchronously behaving acinar population in 3D culture, phenotypes were closely followed employing IF stains and confocal imaging. First, RNA was collected and extracted over 18 timepoints for each of the 4 biological replicates (Figure 45A). Then, another experiment was conducted over 17 timepoints for each of the 3 biological replicates, whereby the cells had shorter (60 hours) exposure to the oncogenes action (Figure 45B). Seeing the preliminary results of comparison between the two experiments, we decided to include even earlier timepoint, i.e. even shorter (36 hours) exposure to the oncogenes action, which would correspond to even earlier "treatment" (Figure 45C). This experiment was done over 15 timepoints for each of the 3 biological replicates.

We termed the kinetics of regression as:

- late regression (from 5 days on doxycycline up to 7 days off),
- mid regression (from 60 hours on doxycycline up to 7 days off), and
- early regression (from 36 hours on doxycycline up to 7 days off)

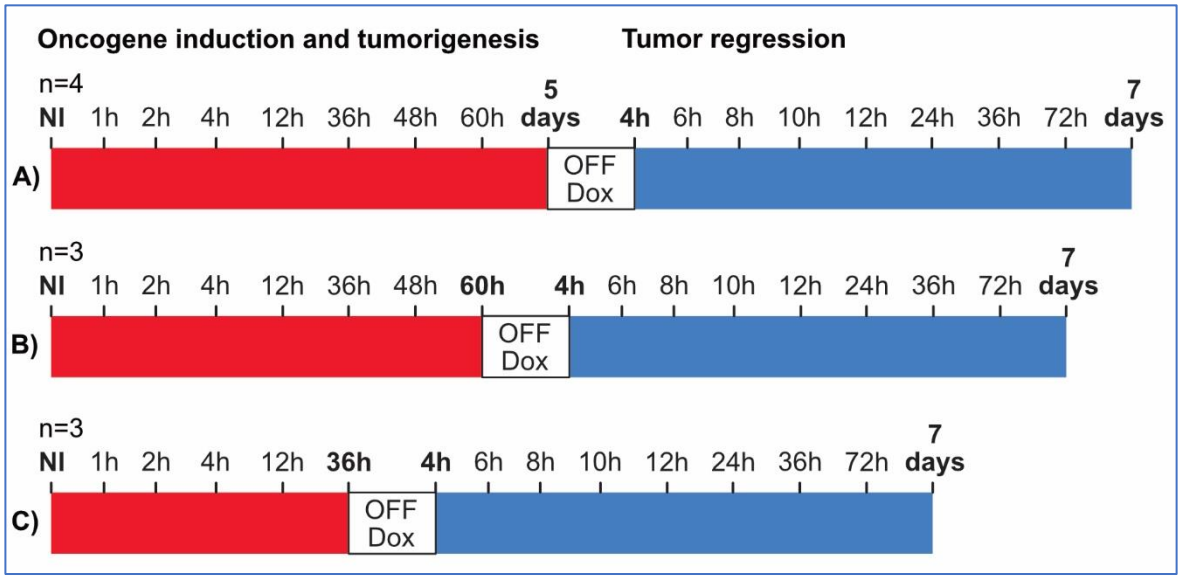

Figure 45: Experimental design to follow kinetics of tumorigenesis and regression. Phenotypes were recorded and RNA collected over shown timepoints during tumorigenesis (red) and regression (blue). A) Exposure to oncogenes action for 5 days and "late" regression; B) Exposure to oncogenes action for 60 hours and "mid" regression; C) Exposure to oncogenes action for 36 hours and "early" regression

Imaging of the induced structures in comparison to the healthy ones (Figure 46A) showed that the structures exposed to oncogenic action for 36 hours were still limited to a single rim (Figure 46B, upper panel). This changed during tumor progression – at 60 hours they started to fill up the lumen (Figure 46B, middle) and at 5 days they corresponded to a full-blown tumor (Figure 46B, lower panel).

Upon oncogene deactivation (doxycycline withdrawal), the structures followed the same regression pattern (Figure 46C) and 7 days after they all looked phenotypically normal (Figure 46D).

IF stains for polarity markers: alpha-6-integrin (shown in red), ZO-1 (shown in green) and GM130 (shown in magenta) showed loss of polarity in the structures going through the tumor progression (Figure 46B, middle and lower panel) and at the initial phases of the regression (Figure 46C, middle and lower panel). The regression to the single rim and repolarization of the structures was achieved at the end of the regression (Figure 46D, middle and lower panel). Structures exposed shortly to the oncogenes' expression did not seem to go through phenotypic changes, resembling to the normal structures in all the phases (Figure 46B, C, D, upper panel).

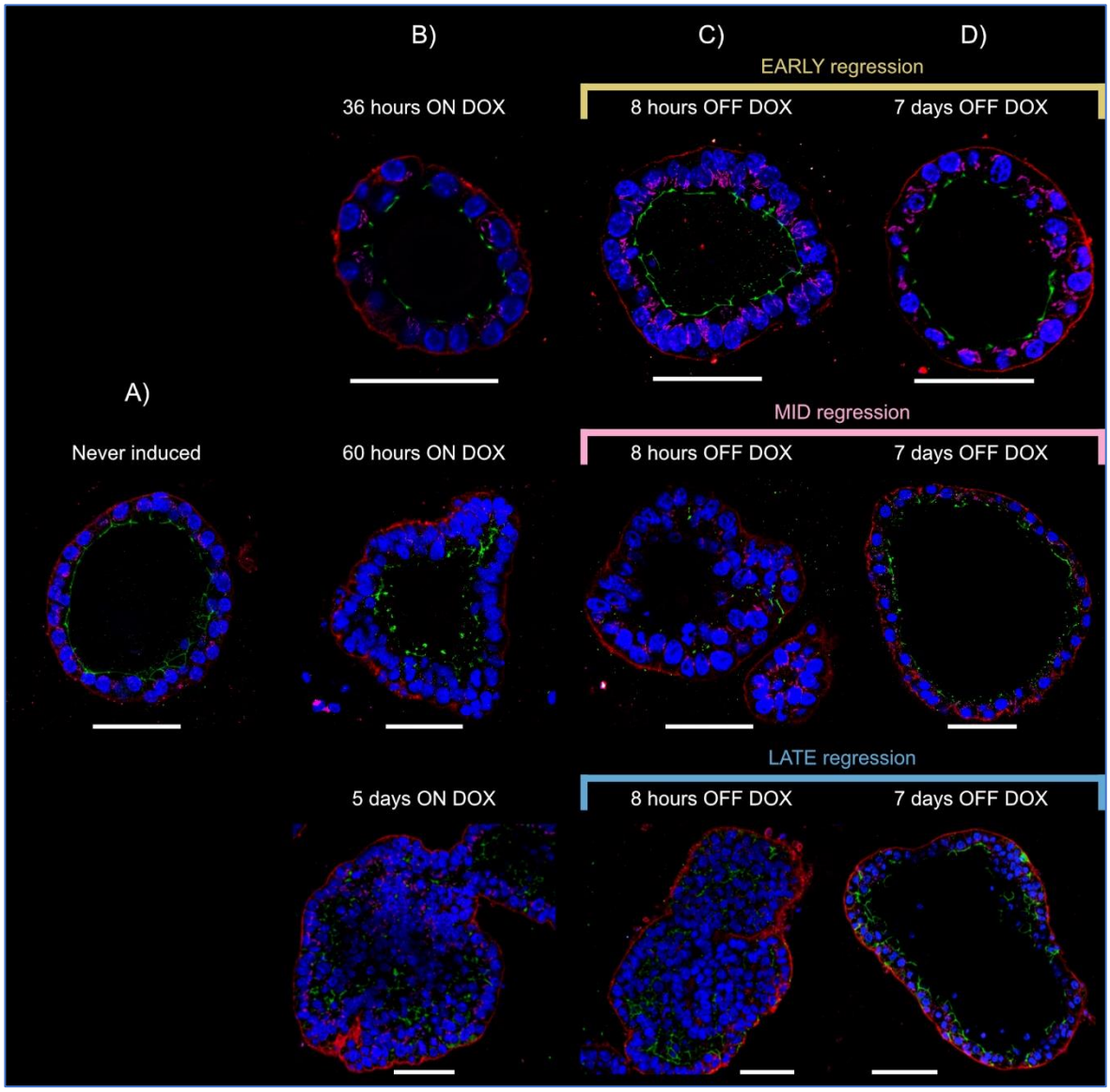

Figure 46: Phenotypes of the structures at the most critical time-points during tumor initiation, progression and regression. Polarity markers – α6-integrin (red), ZO-1 (green), GM130 (magenta); nuclei stained by DAPI (blue). A) Normal structure (never induced); B) Structures during tumor initiation and progression; C) Structures during initial phases of tumor regression; D) Structures at the end of regression. Scale bar: 50 μm.

Next, we performed IF staining for following the expression of human *MYC* oncogene (Figure 47, shown in green). Even though the structures looked phenotypically healthy, identical to normal at 36 hours on doxycycline, human *MYC* was expressed (Figure 47B, upper panel) as further during tumor progression (Figure 47, middle and lower panel).

Upon oncogene inactivation, human *MYC* expression was gone by 6-8 hours after and structures looked the same in early, mid and late regression (Figure 47C, D).

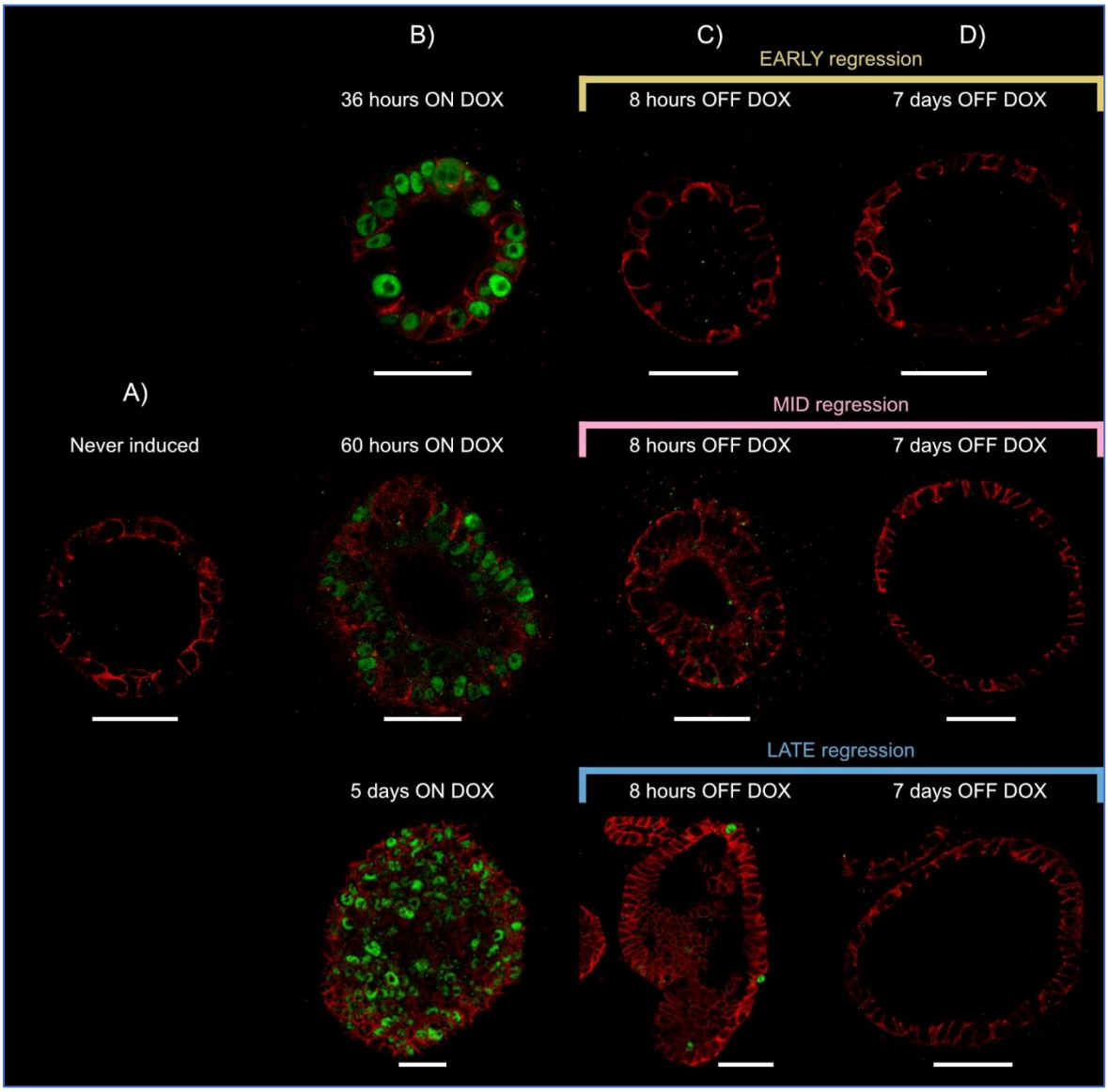

Figure 47: MYC expression in the structures at the most critical time-points during tumor initiation, progression and regression; E-cadherin (red); human MYC (green). A) Normal structure (never induced); B) Structures during tumor initiation and progression; C) Structures during initial phases of tumor regression; D) Structures at the end of regression. Scale bar: 50 µm.

Nevertheless, the stain for apoptosis using Caspase3 marker of cell-death showed that structures in the early and mid-regression did not seem to be dying in the response to oncogene withdrawal as was the case in the late-regression (Figure 48). Caspase3 was

absent in the healthy structure (Figure 48A). It was present in a few cells during the tumor progression (Figure 48B, middle and lower panel). While majority of the cells were dying during in the initial phases of late regression (Figure 48, lower panel) this was not obvious for the structures in early and mid-regression (Figure 48C, upper and middle panel).

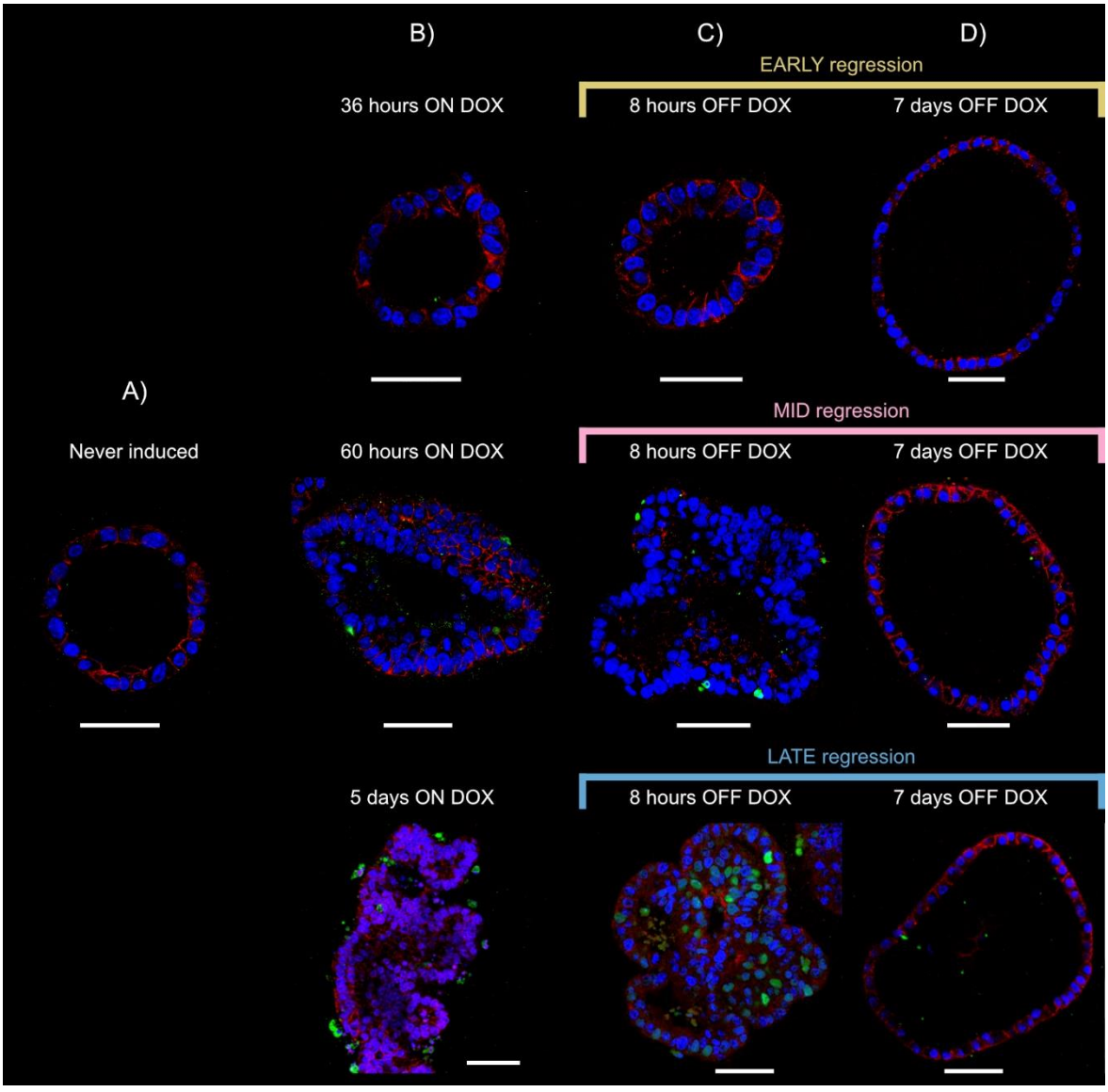

Figure 48: Caspase3-dependent apoptosis in the structures at the most critical time-points during tumor initiation, progression and regression; E-cadherin (red), Caspase3 (green). A) Normal structure (never induced); B) Structures during tumor initiation and progression; C) Structures during initial phases of tumor regression; D) Structures at the end of regression. Scale bar: 50 µm.

3.8.2 Transcriptomic profiles during oncogenesis and tumor regression

The observations on the level of phenotypes were further followed by RNA-sequencing. PCA plot from transcriptomics data (Figure 49) illustrates the kinetics of the changes: green dots represent the normal, never induced cells; red dots and red arrow indicate timepoints during tumor initiation and progression; blue dots and blue arrow indicate timepoints in the late regression; pink dots and pink arrow show timepoints during mid-regression; finally, yellow dots and yellow arrow indicate timepoints during early regression. From the PCA plot it was visible that the residual cells from late and mid regression had a slightly different trajectory in the beginning (corresponding to the different time of oncogene inactivation and their different response), but later on clustered together, not returning to the normal state, still being profoundly changed.

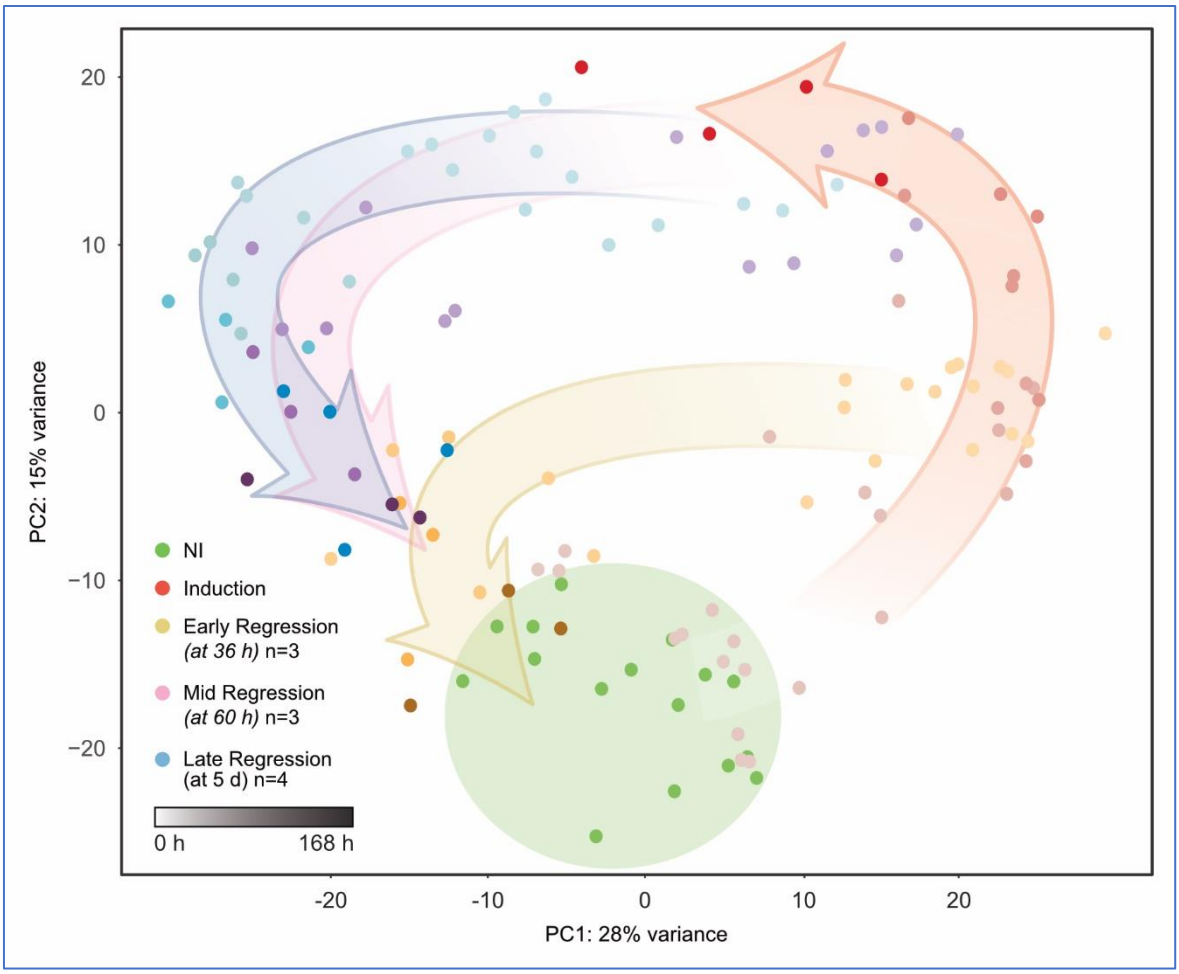

Figure 49: PCA showing clustering of cell populations at many time-points during tumor initiation, progression and late, mid and early regression.

Strikingly, in the early regression trajectory, cells once exposed to the oncogene action managed to follow back to the normal state. We identified this as the "the point of no return", upon which major events that drive tumor progression happened; a threshold that, once being crossed, changed the cells in such a profound manner, probably shaping "oncogenic memory" that would be kept even though the driver oncogene was no longer expressed.

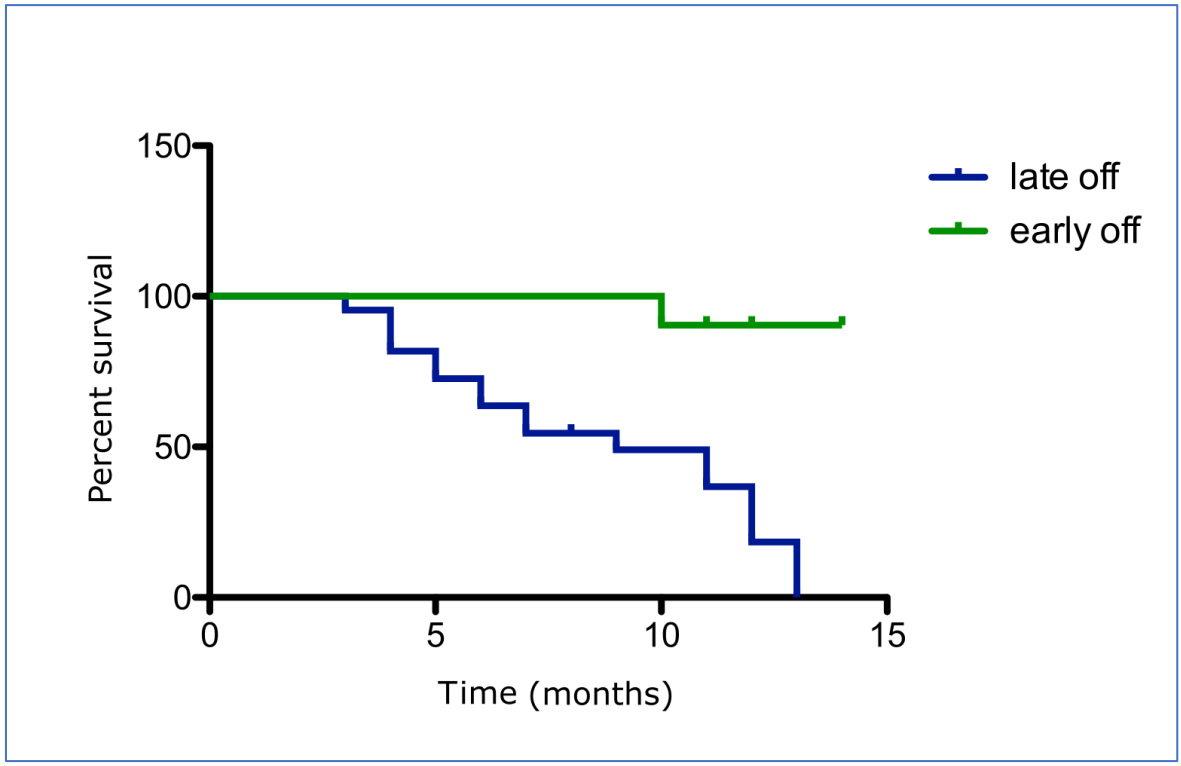

Figure 50: Kaplan-Meier survival curve for the mice with shorter (5-10 days; "early off" group) and longer (4-6 weeks; "late off" group) exposure to the oncogene action.

Our findings were reflected in the observations from the *in vivo* experiments. We compared relapse-free survival between animals exposed to oncogene action for short time (5-10 days; corresponding to our early regression endpoints *in vitro*), named "early off") and standard, long time named "late-off" (4-6 weeks, corresponding to our late regression endpoints *in vitro*). The mice exposed for short time to the oncogene had a significantly (Log-Rank test, $p<0,0001$) higher relapse-free survival (less of relapses compared to the other group).

Chapter 4

DISCUSSION

4.1 A THREE-DIMENSIONAL CELL CULTURE SYSTEM AND INDUCIBLE MOUSE MODELS AS A CORRELATE OF MINIMAL RESIDUAL DISEASE

Cell cultures have been of a great importance in exploring, uncovering and understanding intricate cellular processes, shedding the light on many biological questions. *In vitro* studies of metabolism have provided new insights into how its reprogramming helps cell growth and proliferation, which nutrients are extensively utilized and which pathways are altered. Three-dimensional culture systems provided an additional step forward, as they represented a transition between commonly used classical 2D cultures and multicellular structures grown in more appropriate 3D organization which mimick organ architecture and function.

The system we employ to culture primary luminal epithelial cells harvested from the mammary glands of *TetO-MYC/TetO-Neu/MMTV-rtTA* mice in 3D, provided the following: matrigel that represents exogenous extracellular matrix (ECM) supported the growth of primary mammary cells in 3D, physiological conditions, allowing them communication, correct signaling, cell-to-cell- and cell-to-matrix interactions, more faithful diffusion of nutrients and proper cellular adhesion. All these features together enable polarized cell-cell communication and gene expression patterns which would otherwise be absent or altered in 2D cultures (136). Compared to *in vivo* experiments, which represent the current "gold standard" with respect to modeling complex physiological cell behavior, experiments using 3D cultures enabled more controlled conditions, better reproducibility and require a shorter time to be conducted, while preserving the hallmarks of tissue architecture and phenotypic readout.

Employing an inducible system (described in Chapter 3.1) that allowed us to express oncogenes specifically in the epithelium of a mature mammary gland (mimicking oncogenesis in adulthood) and switching them off (specifically, as they would be aimed at by the targeted therapy), brought us closer to the human patient situation. Selection of primary epithelial cells enabled working with this relatively homogenous, clean and synchronous system, in the absence of many confounding factors coming from the complexity and heterogeneity encountered *in vivo*. These primary cells are not genotypically or phenotypically changed by immortalization (137).

Having an extract of basement membrane (Matrigel) with collagen has been crucial to propagate proliferating single cells, which would later dispose their own ECM components. Some 36 hours after seeding single cells, colonies of 6-8 cells were formed. Their growth and division in the following days resulted in the formation of a hollow ball-

shaped acinus (spheroid). Lumen was formed by epithelial cells, which are organized in a single layer. This acinus corresponded to the TDLU in the mammary gland. These cells exhibited contact inhibition and were polarized, which meant that the specific proteins and lipids would be asymmetrically distributed, giving apical-basal polarity (138). When cells are polarized, specific molecules are found on the surface of the apical domain (oriented towards the lumen). Along with this, complexes are also present here that allow for the exchange of nutrients and other molecules. The basolateral part is composed of cadherins which enable interaction with neighboring cells and the ECM, attaching the cells to the basement membrane (138).

Studying oncogenesis and regression driven by two oncogenes – c-MYC and Her2/Neu – reduced the tumor complexity and allowed us to observe the tumor initiation, progression and the phenomenon of addiction to the driver oncogenes. In addition, it allowed us to study the characteristics of the cells that escape from oncogene addiction. Upon oncogene induction, cells in these acini proliferated in a disorganized fashion, filling up the lumen in the absence of contact with the basement membrane, which reflected its nature as carcinoma. As the cells uncontrollably divided, their polarity was also gradually lost. GM130, which is a marker for Golgi, was not anymore restricted to its apical location within the cells; ZO-1, a marker for intercellular tight junctions which separate membrane components between apical and basal domains, was majorly reduced; alpha-6-integrin, a transmembrane protein responsible for anchoring the basal part and interpreting the signals from the microenvironment was not anymore restricted to basal domain.

However, upon oncogene withdrawal, the majority of the cells died but a certain subset survived. In another inducible system *TetO-cMYC/TetO-Kras/MMTV-rtTA* it was shown that a large fraction of the cells without a matrix attachment lost their mitochondrial potential and died (118). Even though the mechanism behind the survival of these cells remains unknown, it is likely that re-establishing the contact with the basement membrane is also crucial; similar was found in MCF10A cells where proapoptotic protein Bim is inhibited by integrin β1- and EGFR signaling, preventing anoikis (118).

Surviving cells reorganized and regressed back to the single rim. These residual structures exhibited a hollow lumen, phenotypically looking very similar to the controls. However, microarray data (119) and RNA-Seq analysis revealed that these cells were transcriptionally different from both tumor- and control populations (described in Chapter 3.3). This residual population is dormant, which was previously demonstrated by low Ki67 (a proliferation marker) stain positivity (119) and also reflected in the downregulated pathways that reflect cell division processes.

Moreover, spontaneous *in vitro* relapses, in the absence of doxycycline, were detected when organoids derived from residual structures were reseeded, which was not the case for the cells coming from the normal, healthy structures (119). Thus, this system represents a good correlate of MRD that enabled us to characterize it in depth, employing a variety of methods – immunofluorescence and confocal microscopy for following the changes in phenotype, RNA collection for obtaining transcriptomic profiles, metabolites harvesting for revealing its metabolic landscape and DNA harvest for assessing methylome profiles and nature of preserved metabolic memory (approach described in Chapter 3.2). Moreover, our 3D cultures faithfully recapitulate *in vivo* dynamics.

Although the 3D culture models represent an immensely valuable tool for studying MRD we are also aware of its limitations. Acini in 3D are isolated structures whereas in the mammary glands they are branching, connected with the ducts and under the influence of many physiological processes such as hormonal changes and tissue remodeling during various stages of the estrous cycle (137). This is also a monoculture model, which lacks other types of cells, primarily the outer myoepithelial layer. While in the 3D model luminal epithelial cells are in direct contact with the basement membrane, *in vivo* they would be mostly in contact with myoepithelial layer, connected to the basement membrane only through punctual contacts (137).

Immune responses, other types of cells in the microenvironment and the systemic effects of the organism as a whole are also missing. Moreover, metabolic choices of cancer cells are heavily influenced by the available nutrients and thus might not reflect the real scenario, as there is a discordance between the abundance, ratio and presence of different nutrients in the cell culture media and in the body (139). Although better than in 2D, nutrient diffusion in 3D models might not be a faithful recapitulation of diffusion *in vivo*. Additional complexity brought in *in vivo* conditions comes from the dynamic interaction of tumor cells with their environment, their normal counterparts, other cell types and tissues and the effect that the whole-body metabolic changes have on the cancer (139).

Despite certain constraints, 3D cultures still represent a great tool for research; they have been able to faithfully reflect many key biological phenomena found *in vivo* (137), especially in terms of increased nutrient uptake, flux through glycolysis, protein turnover and lipid catabolism (139). 3D cultures continue to be utilized in numerous scientific and pharmacological studies, tissue engineering and regenerative medicine. Limitations can be overcome through subsequent validation of the prominent findings in murine models *in vivo*, which is an approach we took.

In vivo validation of the key findings was performed through experiments on inducible mice models which were fed with doxycycline supplemented food to express the oncogenes. Inducible mouse models significantly improved research in the field of cancer, harboring key oncogenes identified in human tumors and enabling their fast induction in the tissues of interest as well as their rapid de-induction. This possibility to switch the desired oncogenes on and off gives them a great advantage over constitutive models. Characterization of our 3D models through many time-points during tumorigenesis and regression, showed that the *MYC* oncogene was induced as early as one hour after doxycycline administration and was no longer detectable 6-8 hours upon doxycycline withdrawal. The mouse mammary tumor virus long terminal repeat (MMTV) sequence enabled tightly regulated and homogeneous expression of oncogenes specifically in the mammary gland tissues (140). One aspect that still needs to be improved is that the oncogenes are induced in the majority of cells, which does not reflect the real state where neoplastic transformation starts in one cell. Some new murine models where this issue is being addressed are being developed (141).

Our previous research showed that the oncogene activation in *TetO-cMYC/TetO-Neu/MMTV-rtTA* mice led to the rapid onset of invasive mammary carcinomas (119). When animals developed full-blown tumors with a large tumor burden (10% of body weight maximum expansion), they were put on the normal diet, thus silencing the oncogene expression and allowing for tumor regression to a nonpalpable state. Complete regression happens due to the strong oncogene dependence. However, tumors reoccur after a long dormancy period, usually six months after oncogene deactivation (119).

Recurrences were also previously demonstrated in another MMTV-rtTA regulated mouse model with the same activated Neu transgene, where its expression remained undetected in relapses, indicating the independence of a residual population on Neu overexpression (121). In another mouse model with inducible *MYC*, the authors showed that the re-activation of the tumor initiating oncogene happened through somatic mutation in the MMTV-rtTA that rendered it independent of doxycycline activation (142). Since this subset of surviving cells is known to subsequently drive tumor relapses, it allowed us to verify findings from our 3D culture system also *in vivo*, on the population that more faithfully represented residual cells; these cells serve as a correlate of MRD.

Furthermore, we were interested how the findings from the mouse models correlated to the human situation.

Mouse mammary carcinomas are heterogeneous, with around 17 distinct subtypes characterized based on their specific transcriptomic profiles (143). They harbor different

driver mutations, but some also show intra-model variation (143). Besides existing subtypes in mice and humans, defined based on their histopathology and molecular profiles, identification of novel biomarkers keeps revealing new subgroups, with many distinct but also overlapping features. When compared to the human breast cancers according to their transcriptomic profiles, murine models with *MYC* amplification were linked to human basal-like breast cancer (143).

Although MMTV-Neu mouse models surprisingly resembled more to the luminal A tumors (143), another study also showed their clustering to a proportion of HER+ human cancers (144). Residual cell population from our mouse models clustered closely with the basal subtype, sharing some of the key altered pathways such as glycolysis or the urea cycle/arginine metabolism in basal Her2+/- samples. This was not surprising, given the fact that the oncogenesis in our system was driven by *MYC* and Neu oncogenes and considering their previously established link with human breast cancer subtypes.

Mouse mammary glands exhibit a certain difference from human breast architecture. In terms of their structure, nulliparous mammary glands are simpler and their lobules develop and expand only during pregnancy – after involution, the glands remodel back, resembling a nulliparous state (145). While in human breast lobules can have a different estrogen or progesterone status, in mouse glands, large regions of estrogen positive and progesterone positive luminal cells are rare (145). Another difference is observed in the stroma, which is predominantly composed of adipose tissue in mice, containing only small amounts of fibrous tissue that is more dominant in the human breast (146). While primary metastatic sites of breast cancer cells in humans are the bone, brain and lungs, cancer cells in mice metastasize primarily to the lungs, which likely comes from certain differences in predominant vascular routes (141).

Nonetheless, mouse and human mammary glands still display many similarities. Their basic structural and functional characteristics were found to be very similar: luminal epithelial cells and the myoepithelial layer, expressing specific markers, lay on the basement membrane that separates them from stromal components. Besides correlations found in this study, other analyses of mouse and human transcriptomes showed conservation of many genes and pathways, validating the utilization of mouse models in studying human breast cancer (147).

Tetracycline-regulated mouse models are widely used. Besides enabling fast and reversible oncogene expression, they are commercially available, easy to use and have minimal toxicity (140). Upon de-induction, tumors regress to a nonpalpable state, reflecting the oncogene addiction that is observed in many human cancers in the clinics,

as well as the escape from therapy, survival and regression, which was also demonstrated in our mouse model (119). This feature makes it especially valuable as a tool to study MRD, with potential novel markers that could be further verified in samples/databases from human patients and potentially translated a clinical setting.

4.2 A MULTI-OMICS APPROACH FOR STUDYING MINIMAL RESIDUAL DISEASE

Development of high-throughput methods and advanced bioinformatic tools have offered exciting possibilities to explore many different aspects of cancer. Recent technological advancements have resulted in the rapid generation of large, multivariant and complex datasets followed by the emergence of different -omics disciplines (aiming for the analysis of "all constituents considered collectively", i.e. "ome") (148). Besides established fields like genomics, transcriptomics, proteomics and metabolomics, new subdisciplines such as lipidomics, fluxomics, secretomics and pharmacogenomics keep unfolding, leading to more specific datasets and growing complexity. Interpretation and integration of the –omics data thus requires engagement and collaboration of different experts from their respective areas.

As no single –omics approach can give us a full picture and understanding of the complicated biological system as a whole, thus employing a multi-omics approach yields an overall systemic grasp of key processes underlying such complexity. Linking multi-omics datasets together is greatly helped by the existence of publicly available, curated databases which allow mapping the findings from various –omics disciplines back to annotated genes that have been greatly characterized and described, along with their biological function.

A need for integrative approaches is imposed by the fact that the cell/organism is an integrated system and should be observed as such.

Changes in cancer occur on many different levels: genetic, transcriptional, protein, metabolic and interactions of the whole molecular network. This makes cancer a highly complex disease, with multiple different molecular systems at play. Therefore, this requires a holistic view and a multi-omics approach (149). This is why efforts to understand cancer include engagement of different –omics technologies that are widely applied in the fundamental research of cancer. Similarly, this is now also applied in a clinical setting, from the discovery of differentially expressed genes important in tumorigenesis to a range of biomarkers and drug targets (149).

DISCUSSION

Residual cells in MRD are also entities that needs to be looked as a system. In attempt to reveal its potential vulnerabilities and uncover its elusive features, we looked at its transcriptome, metabolome, lipidome and methylome. We took the same approach in characterizing normal and tumor cells for the comparison. This resulted in generating the following datasets for residual (OFF), tumor (DOX) and normal (NI) cell populations:

- RNA-sequencing data,
- metabolic profiles from untargeted metabolomics (ions and tentative annotations) acquired through LCMS,
- metabolic profiles (precisely identified metabolites) acquired through GCMS,
- lipid profiles (lipid categories, classes and species) acquired through shotgun lipidomics,
- DNA-methylation profiles acquired through bisulfite-free sequencing.

However, integrating –omics datasets is a challenging task as currently there are no routine and organized methods for integration, analysis and interpretation of multi-omics data. The integration is mainly achieved through interdisciplinary approaches and collaboration between biologists and bioinformaticians. While gene expression profiles provide easier interpretation, as transcriptomics established itself as one of the oldest and well-explored –omics disciplines, understanding the metabolome is more difficult. Challenges in interpretation and analysis of metabolic datasets are posed by the fact that metabolomics is still a relatively new discipline, with a lack of annotated metabolites in publicly available databases that would enable mapping them and also an absence of tools that provide a direct link between each metabolite and transcript. Integrating these big datasets is a challenging work, but also rewarding, as it identifies interesting and invaluable pieces of puzzle in uncovering many biological questions.

To get an idea of the global profiles of residual and tumor states, as well their relation to each other and the controls in terms of differences or similarities, we employed Principal Component Analysis and hierarchical clustering. Pathways characteristic for MRD and primary tumor in comparison to the normal state, were revealed through enrichment analyses. We looked at the datasets conceptually, analyzing the transcriptome and metabolome individually, then comparing the obtained information and interpreting them together for a system overview. While the value of this approach comes from getting the details from different levels of information, there is a risk that some associations might be missed if the different datasets are not analyzed together

(150). This can be overcome by employing statistical analysis of multi-omics data, i.e. using methods that look for statistical associations between components of different –omics datasets (150).

We extended our analyses to examining lipid profiles in the residual cells, with the goal of gaining additional knowledge about the system. Namely, lipids have role in many cellular processes and are also implicated in major diseases. Although lipidomics emerges as a valuable approach, especially with the development of high-throughput platforms, challenges in interpreting lipidomic datasets remain, as it is one of the youngest disciplines still in development. Further advancement of bioinformatic tools for interpretation of this specific datasets holds a promise for better understanding of many biological processes involving lipid molecules.

Looking at the metabolome brought us closer to biochemistry of the cell, while transcriptome and metabolite predictions provided a more detailed overview, giving us a full picture of gene expression changes and potential targets in MRD. In addition, lipidomic analyses extended our knowledge on the characteristics of the residual cells, while assessing DNA methylation profiles will help us in preliminary testing of our hypothesis regarding the molecular mechanism driving metabolic memory. Overall, exploring MRD from different angles was a powerful combination that provided us with complementary information and a global map of changes in our biological system. Employing this approach allowed us to obtain a comprehensive metabolic landscape of MRD and identify metabolic vulnerabilities of potential significance in eradication or prevention of recurrent tumors.

4.3 THE METABOLIC LANDSCAPE OF THE RESIDUAL CELL POPULATION

Residual cells that remain in the body in undetectable levels, representing MRD, have the potential to eventually give rise to tumor recurrence which can be lethal. Even though it is clear that understanding their biology and revealing their vulnerabilities is of utmost importance, difficulties to obtain and follow them limit our knowledge. Thus, to characterize this elusive cell population, we employed inducible mouse models of breast cancer and primary 3D cultures of mammary organoids, which allowed us to study the residual cells upon oncogene inactivation. As we knew that these residual cells were able to drive relapses (119) and our microarray data suggested that altered metabolism could be at play, we hypothesized that metabolic reprogramming established during oncogenesis could be carried over into residual cells, leaving them with metabolic features that set them apart from normal mammary epithelial cells. This would possibly

prime them with a tumorigenic potential and lead to subsequent recurrences, but could be also used as a target that could be selectively hit. We showed that, despite the phenotypic similarity to the normal cell population, the residual cells exhibited a distinct transcriptional and lipid profile that distinguishes them from normal, but also from the tumor cells (described in Chapter 3.3 and 3.7). Strikingly, their metabolic profiles were similar to tumors, despite apparent normal, repolarized and non-proliferative phenotypes, and the absence of oncogene expression (described in Chapter 3.4). We identified increased flux through glycolysis, the urea cycle and increase NOS activity as the main metabolic hallmarks present in the tumor and preserved in the residual cells, also verified *in vivo*. Furthermore, we found that these alterations seem to be a feature of a basal-like HER2 negative and positive breast cancer subtypes, based on the comparison of microarray datasets from the residual cells of the patients undergone neoadjuvant treatment and normal tissue from the healthy women (described in Chapter 3.6).

4.3.1 Despite phenotypic similarities, the residual population is different from its normal counterpart in its transcriptional and lipid profile

In order to characterize the residual structures starting from their appearance to their gene expression profiles, we employed immunofluorescence and RNA-sequencing. This showed us that the residual cell population was phenotypically very similar to the control: structures regressed back to the monolayered acinus, repolarizing and establishing the contact with the basement membrane. Cells were non-proliferative, in the state of dormancy, as was shown by low Ki67 positivity (119) and downregulated cell division processes seen from the GO Enrichment terms. Moreover, the driver oncogenes, *MYC* and Her2/Neu were inactivated, as was demonstrated at the MYC protein level and through gene expression profiles for both of the oncogenes.

Despite the phenotypic similarity to the normal structures, transcriptionally, residual cells exhibited a distinct profile, that distinguished them from both the tumor and normal populations. In addition, PCA based on lipidomic data obtained from normal, tumor and residual cells, revealed differences in their lipid profiles. The clustering of these three groups was strikingly similar to the clustering observed on the transcriptomic level.

GO enrichment analysis based on transcriptomics data showed that residual cells differ from normal cells mostly in the downregulation of cell division and cell cycle, signaling, response to stimuli and cell communication, as well as the upregulation of cell adhesion, regulation of localization, locomotion, activation of cell surface receptor signaling- and

immune-response related pathways, glycolysis, nucleoside metabolism, nucleotide phosphorylation and metabolic processes related to ATP generation.

The upregulation of locomotion, cell adhesion and epithelial cell differentiation in the residual population probably reflects the movements of the cells during tumor regression, their reintegration into the single rim, repolarization and restoration of contacts with the basement membrane (extracellular matrix). Residual cells apparently need to re-establish these contacts for their survival, as they lost oncogenic activity that could otherwise prevent anoikis.

Interestingly, despite this reintegration into a phenotypically normal-looking rim, cell communication was still downregulated, as in tumor cells. Absence of cell-cell communication or its dysregulation is also observed in many human cancer cells and is indicative of aberrations in intracellular signaling pathways (151) and possible alterations in the composition of the cell membrane. Indeed, GO enrichment of downregulated processes indicating cellular compartment revealed changes impinging on the plasma membrane and cell periphery contacts. Thus, further exploring changed or aberrant communication patterns in the residual cell layer would be needed, as it could be important in the later progression of MRD to tumor recurrence.

Surprisingly, certain metabolic processes were also upregulated in the residual cells, indicating higher flux through glycolysis, purine metabolism, nucleotide phosphorylation and ATP generation. While glycolysis is usually upregulated in the proliferating cells mostly for the generation of intermediates for biosynthesis, residual cells are non-proliferative so it is not likely to remain enhanced for this particular reason. Purine metabolism is also important for the production of building blocks, but also generation of co-factors which promote cell survival, which could be its potential role in the residual cells (152). As research in acute lymphoblastic leukemia shows that purine biosynthesis plays an important role in drug resistance and relapse (153), it would be interesting to explore it further in the context of MRD in breast cancer. Nucleotide phosphorylation could be potentially important in signaling events that would enable the survival or adaptation of residual cells in response to the deactivation of oncogene signaling. Increased ATP metabolism could be a consequence of enhanced flux through glycolysis.

Overall, altered biological processes obtained from RNA-Seq data of residual and tumor cells compared to the normal cells, pointed out to some key differences between the residual and tumor population. Moreover, residual cell populations, compared to tumor cells, display large phenotypic differences: while tumor structures have a filled lumen as a result of uncontrolled proliferation, residual structures have their lumen cleared upon

regression. Additionally, they are repolarized, dormant and in contact with the basement membrane. Besides this, the most important distinction is the lack of oncogene expression in the residual cells after oncogene inactivation. This is consistent with the *TetO-MYC/TetO-Kras^{G12D}/MMTV-rtTA* model (118). The rapid regression of the tumor reflects the phenomenon of oncogene dependence, described in mouse models, but also seen in the clinical setting, where interference with the single driver oncogene represents the basis of targeted therapies, such as is the case with EGFR inhibition in NSCLC (17). Survival of the residual cells means that they were able to escape oncogene addiction, establishing pro-survival signals despite the absence of oncogene signaling. Regression to a single rim and the establishment of communication with the basement membrane could be crucial for survival, as a study on the *TetO-MYC/TetO-Kras^{G12D}/MMTV-rtTA* system shows that most of the internal cells lose their mitochondrial potential, which leads to caspase-3-dependent apoptosis; this, however, does not seem to affect cells that are directly attached to the matrix (118).

Although residual structures regress back to a single rim and seemingly restore their polarity, as seen from the immunofluorescent stains and transcriptionally upregulated pathways related to cell-adhesion, lipidomics data suggest that this restoration might not be fully back to the normal state, as their lipid– and, thus, possibly membrane composition is different. At the level of lipid species, many differences were observed, even within classes that were not altered between the residual, normal and tumor cells. Particularly interesting features include a shift from short to long chained fatty acids in sphingomyelin and ceramide species and depletion of arachidonic acid species in the phosphatidylethanolamine class. In addition, genes encoding some of the enzymes involved in these processes showed high overexpression levels in comparison to the residual cells. It could be that the similar pattern in transcriptomic and lipidomic PCA plots reflects alterations in specific genes that could lead to changes in the lipid composition. The observed alterations in the length of fatty acids might indicate changes in membrane fluidity.

4.3.2 Despite phenotypic differences and absence of oncogene signaling, residual population shows similarity to the tumor in its metabolic profile

Strikingly, despite obvious differences between tumor and residual cells (in their phenotype, proliferative state and oncogene expression), untargeted, broad metabolomics (Q-Exactive MS) demonstrated a similarity in their metabolic profiles. These two populations showed clear clustering from the normal cells based on the hierarchical bi-clustering of detected intracellular ions. Potential effects that doxycycline could have had on this clustering was excluded, as the wild type controls (non-inducible

cells, lacking the rtTA inducer, treated by doxycycline) clustered together with the normal, never induced cells. Metabolic enrichment analysis based on intracellular ions revealed significant enrichment in sugar and amino-acid metabolism.

The separation of the three states (normal, tumor and residual) based on hierarchical bi-clustering was not clear in case of measured extracellular ions, probably due to the fact that many non-significant metabolites in the media were detected which did not necessarily reflect true biology of the samples and contributed to the extracellular noise. This was also reflected in the metabolic enrichment, as there were no significantly changed pathways.

Our observation of the metabolic similarity between residual and tumor cells was further confirmed by GCMS, which is generally referred to as a more precise method. PCA plotting and the heatmap based on 54 measured intracellular metabolites showed co-clustering of the residual cell population with the tumor, reflecting some kind of metabolic memory in the residual cells. Similar results were observed from the PCA plot and a heatmap done on the 30 measured extracellular metabolites, where the residual population was closer to the tumor than to normal cells, but still exhibited a distinct profile. This could be a result of other metabolic differences, alternative metabolic routes and fuels, and the proliferative status of the cell.

To our knowledge, this is the first study that showed overall metabolic similarities between tumor and residual cells, hinting to some kind of metabolic memory. Data from experiments *in vivo* back up the idea that this is not simply a transient phenomenon. Regressed glands, 9 weeks after oncogene inactivation, displayed differences in comparison with the healthy glands from age-matched control animals. In this context it is important to mention that the cells in the mammary gland undergo hormonal stimulus induced proliferation and remodeling every 5 days (duration of the hormonal cycle in a female mouse). Therefore, the metabolic similarity of residual disease to the tumor state was preserved despite several cycles of tissue remodeling *in vivo*.

Taken together, we propose that the metabolic changes, arising from the initial oncogenic signaling, might establish another layer of regulation that could even surpass the initial oncogene addiction, rendering the cell dependent on the metabolic network.

4.3.3 Metabolic changes in the residual cells

Untargeted metabolomics revealed metabolic differences between residual, normal and tumor cells. While some ions could not be annotated and many dipeptides found, the

most significantly changed intracellular ions in residual cells were annotated to the following metabolites: leucine, proline, succinic acid, methionine and AMP found in higher levels and 5,10-methenyltetrahydrofolic acid found in lower levels in residual (as well as tumor) cells compared to the control. Increased levels of proline, succinic acid and methionine might indicate epigenetic changes. Proline can serve as an alternative energy source in the times of stressful conditions as its specific degradation generates ATP and ROS that can lead to epigenetic reprogramming (154). Besides being an intermediate metabolism of TCA cycle, succinate emerges as an interesting oncometabolite as, when found in pathologically higher amount, can drive tumorigenesis and epigenetic changes (155).

Extracellular levels of lactic acid were elevated in the residual and tumor cells, indicating a glycolytic flux towards lactate production instead of pyruvate. Low levels of arginine and high levels of ornithine in residual and tumor cells hint towards alterations in arginine metabolism and the urea cycle. In addition, among the significantly changed metabolites we also found ribose 1,5-bisphosphate, one of the pentose phosphate pathway metabolites, which was, interestingly, much lower in the residual cells compared to both tumor and normal, indicating potential relevance of this pathway in the residual cells. Indeed, microarray data from our residual cells pointed to the upregulation of this pathway (119) and reliance of residual cells on the enhanced PPP was also observed in epithelial ovarian CSCs (156). Since residual cells in our system show elevated levels of ROS (119), the significance of upregulated PPP could be due to the generation of NADPH which is needed for coping with the oxidative stress.

The advantage of employing untargeted metabolomics was that we could obtain a wide picture and global profiles as there were thousands of measured ions. However, the limitation was that these metabolites could not be precisely identified as the method generates tentative annotations. Thus, measurements obtained from untargeted metabolomics hold a potential for exploring these changes in the future, with the high-throughput platforms and growing datasets that will enable better metabolite identification. This could eventually reveal novel metabolic biomarkers, as we identified many dramatic changes in the ions that could not be annotated. Putative annotations made interpretation of these data difficult, but this is where additional measurements by the GCMS and the transcriptome profiles obtained from RNA-Seq data could help gain a better understanding of the observed metabolic changes.

Trying to put the altered metabolite levels into perspective by exploring the literature and conceptually looking at the potentially relevant genes from RNA-Seq data revealed that in some cases, the observed metabolic change could be linked with the specific gene

changes in both residual and tumor cells. For instance, a higher amount of adenosine monophosphate (AMP) was found in both residual and tumor cells as compared to the normal. A study exploring potential metabolic biomarkers using cancer cell lines identified AMP as one of the six most prominent, significantly upregulated metabolites in cancer cells (157). On a transcriptional level, we observed alterations in AMP metabolic processes, especially in the expression of Ak4, Ak3 and 5'-nucleotidase genes (Nt5e) in both tumor and residual cells. However, in other cases, the change in a specific metabolite observed both in tumor and residual cells could be linked to the upregulated genes only in the tumor cells. Naturally, there are many limitations in this approach, as it is impossible to have a complete overview of many genes that could be driving changes in many alternative pathways leading to the same result. While this is the most obvious explanation, it is also interesting to remember the differences in transcriptomic profiles between the two populations, even when the PCA was based only on the metabolic genes. While on the transcriptional level, residual cells exhibit a unique profile that differentiates them from the tumor and control, on the metabolic level, they are closer to the tumor. Several reasons could be the reason for this observation. Most transcriptomic differences could probably be attributed to the activation of the *MYC* and Neu oncogenes (*MYC* alone exerts 10,6 times higher gene expression) and their downstream genes. The extent of the gene expression changes in the tumor is thus much greater than in the residual cells, which lack the oncogene signaling. The expression of around 5000 genes was changed exclusively in the tumor, and doing the PCA only on this subset of genes results in the same clustering as compared to when all the genes are taken into account (from communication with Katharina Zirngibl, PhD candidate, Patil group, EMBL, University of Heidelberg), which provides quite a strong argument for this explanation. Furthermore, residual cells restore normal expression of portion of genes after the driver signal is gone. For example, the residual structures are able to repolarize, re-establish signaling, and re-establish contact with the basement membrane and ECM. Therefore, reintegrating in the surviving rim and restoring expression of genes regulating these processes could be largely responsible for clustering of residual cells away from tumor in the PCA comparison. To test this assumption, we plan to do the PCA on the subset of such genes. Some of the genes were changed only in the residual cells (communication with Katharina Zirngibl), setting them apart from both tumor and normal. This signature is probably a reflection of engaging alternative pathways and survival adaptations or stress following silencing of the oncogenes and reflecting the resulting "oncogenic shock" (oncogene withdrawal). Another explanation for the distinctive transcriptional, yet similar metabolic profiles of the tumor and residual cells, could come from the fact that, among the genes changed in both the residual and tumor cells, expression levels were much more dramatically altered in the tumor for

metabolic signaling pathways. However, even less pronounced changes in the transcriptome of the residual cells could lead to similar changes in the enzymes that bring forward an altered metabolic network, compared to the normal cells. As a result, residual and tumor cells would display similarly dysregulated metabolic networks that could arise from the portion of genes that were epigenetically changed during tumorgenesis and remained differentially expressed, despite the absence of oncogenes, thus exerting some kind of a "memory". When the PCA analysis is performed taking the genes that are found differentially expressed both in tumor and residual cells, the residual population clusters closer to the tumor, but still independently (communication with Katharina Zirngibl). This shows that the residual population, while still exhibiting a specific transcriptional profile, does preserve the subset of changes established in the transcriptome from the tumor state. This could contribute to the similarity in some main metabolic features. Furthermore, the established metabolic network could drive further changes and represent another layer of regulation.

The heatmap of 54 measured intracellular metabolites by GCMS showed that tumor and residual cells formed an obvious cluster distinguishing themselves from the healthy controls (never induces and wild-type cells treated with doxycycline). Around one third of detected intracellular metabolites were changed in the residual and tumor population compared to the normal cells. The Over-Representation Analysis on these changed metabolites revealed that the observed changes could reflect alterations mostly in sugar- and amino acid metabolism, but also in metabolism of certain lipids. From the 30 detected extracellular metabolites, half of them were significantly changed between the three populations – here, residual cells exhibited a more distinctive profile with metabolite levels somewhere in between tumor and normal. This is probably a result of the previously discussed differences about the levels of the changed enzymes, inactivation of the oncogene signaling and thus decreased uptake of certain metabolites such as glutamine, which are not needed to that extent as there is no need to support proliferation.

Although GCMS measurements hinted at the most probable metabolic alterations, the limitation of this approach is that these metabolite levels are just a "shot" in time, making it difficult to assume the possible fluxes and dynamic changes. Still, looking at the levels of the key metabolites intra- and extracellularly, gave us ideas about the consumed and secreted metabolites: levels of extracellular glucose were much lower in residual and tumor cells compared to the controls and levels of intracellular glucose were also lower, indicating its higher uptake and utilization. Levels of lactate were high both intra- and extracellularly in tumor and residual cells, highlighting its production from glycolysis and

excretion from the cell. The same was observed for proline, ornithine and urea. Putrescine, a polyamine that is normally produced in small amount in the normal cells, was found to be lower intracellularly, but in higher amount extracellularly in the tumor and residual cells, which probably reflects its increased export.

Overall, these metabolic changes indicated that the alterations in the residual cells mainly happen in the glycolysis and urea cycle. To explore this hypothesis further, we combined transcriptomic and metabolic datasets with metabolite predictions that generated an overview of dysregulated processes, allowing the comparison of residual cell features with the normal and tumor cells.

4.3.4 Data integration reveals enhanced glycolysis, urea cycle and NOS2 activity as the main metabolic features preserved in the residual cells

Overlay of the KEGG reactions and metabolic predictions inferred from RNA-Seq data with the actual measured metabolites revealed metabolic changes in the tumor cells in compared to normal reflected metabolic features typical of cancer: enhanced glycolytic flux towards lactate, upregulated TCA cycle, alterations in fatty- and amino acid metabolism. Upregulated glycolysis provides tumor cells with the intermediates for biosynthesis. Enhanced TCA cycle ensures that these cells meet their energy demands. Decreased levels of palmitate and stearate could be the result of the upregulated fatty acid oxidation as alternative source of energy. Increased levels of many amino acids reflect anabolic metabolism. The upregulation of urea cycle was a surprising, novel finding. In addition, changes in glutathione and methionine metabolism and SAM to SAH conversion could indicate epigenetic changes.

The same overlay done on the residual cells revealed that, despite inactivation of the oncogene and their non-proliferative state, residual cells still preserve enhanced glycolysis and urea cycle. In addition, they exhibit metabolic alterations in TCA cycle, amino acid-, lipid and glutathione metabolism in comparison to the normal cells.

Increased uptake of glucose, upregulated flux through glycolytic pathway and higher excretion of lactate found both in tumor and residual cells highlight the Warburg effect (aerobic glycolysis). Aerobic glycolysis is important and not surprising feature of many tumor cells – it fuels their growth and proliferation, as it provides many important intermediates for biosynthesis and maintenance of redox homeostasis (83,91). However, residual cells are non-proliferative, without these demands. What could then be a benefit and significance of Warburg effect in MRD and do all residual cells exhibit it?

DISCUSSION

Reports on residual cells' metabolism have been mainly based on research done on cancer stem cells (CSCs) that, besides being slow cycling and having self-renewing capacities, exhibit high resistance to therapy and thus are considered to be a main "reservoir" of MRD. The Warburg effect is not a universal feature of the residual cell population. Quite the contrary, CSCs appear to rely more on mitochondrial metabolism-consuming more O_2, displaying increased mitochondrial mass, membrane potential and high mitochondrial ROS levels (158). The breast CSC populations characterized by a high mitochondrial mass are more efficient in the formation of mammospheres, tumor initiation *in vivo* and resistance to paclitaxel (159). Increase in mitochondrial mass was also shown in primary human breast CSCs from metastatic breast cancer samples or a patient derived xenograft (PDX) (159). An active antioxidant response seems to keep the levels of ROS in control (158).

Highly active mitochondrial metabolism is also seen in lung cancer, glioblastoma (160) and colon cancer stem cells (161). Its importance for survival is demonstrated in the model of PDAC upon oncogene ablation. This residual dormant cell population displayed dependence on OXPHOS and autophagy that enabled their survival in an unfavorable microenvironment (162). Cells with the ability to metastasize often have to rely on other sources of energy, mitochondrial activity and peroxide signaling (163).

While dependence on OXPHOS seems to be characteristic for CSC, other studies report that the Warburg effect and reliance on glycolysis is seen in CSC population in breast, lung and ovarian cancer, osteosarcoma, glioblastoma (158), nasopharyngeal carcinoma radioresistant cells and hepatocellular carcinoma TICs (160). These contradictory findings, sometimes reported even in same tumor types, could be explained by different microenvironmental and experimental conditions, cells' differentiation status, emergence of different CSCs subpopulations (epithelial- or mesenchymal-like) and overall tumor heterogeneity that would be reflected in engaging different metabolic solutions (158).

Several studies indicate the reason why CSC population could favor glycolysis. In CSCs derived from human breast cancer tissues increased activity of key glycolytic enzymes is needed to maintain stemness phenotype and cellular redox homeostasis (164), as mitochondrial ROS could affect stemness through inhibition of β–catenin signaling (165). In stem-like spheroids obtained from ovarian cancer, glycolysis plays a role in resistance to apoptosis, providing survival in very hypoxic environment, making the tumor aggressive (166). In stem-like U87 glioblastoma cells, where mitochondrial metabolism is suppressed, glycolysis represents a compensatory mechanism (167). In our system, the favoring of glycolysis over mitochondrial respiration could be a way to keep redox

homeostasis, since our previous study has shown that the residual cell population copes with the oxidative stress as high levels of ROS come as a product of enhanced lipid metabolism (119). Another reason could be rooted in the need of residual cells to rearrange and reestablish focal adhesions and contact with ECM, as a recent study found glycolysis to be necessary for the cellular movement in breast and prostate cancer cells (168). Moreover, glycolytic intermediates are not just necessary for proliferation but also for generation of other important molecules such as NADPH, as seen in quiescent fibroblasts. Even though enhanced glycolysis was previously not considered to be a feature of quiescent cells, the latter study demonstrates that this is not always necessarily the case (169).

Dysregulated urea cycle is another interesting pathway that emerged in both tumor and residual cells. This pathway is used for conversion of toxic ammonia (waste from protein degradation) to urea and is mainly active in the liver. The role of the urea cycle in MRD is unknown and only recently started to emerge as being dysregulated in cancer cells, where it is linked with anabolic processes through providing nitrogen (170). In the residual cells, enhanced urea cycle could be a cell's way to deal with the excess products that might be a consequence of changed abnormal signaling, a leftover from tumor state. Interestingly, another enzyme involved in arginine metabolism, NOS2, catalyzing synthesis of nitric oxide (NO), was highly upregulated on transcriptomic level. The role of NOS2 in cancer is still unclear. In breast cancer lacking ER, NOS2 is found to be frequently upregulated and associated with poor prognosis (171). However, its role in MRD remains unknown. As NO is involved in many processes, it could modulate the tumor microenvironment, promoting tumor progression.

Upregulated glycolysis and urea cycle, as well as Nos2 overexpression were preserved even 7 days after oncogene expression was switched off, hinting to a long-term effect rather than a reflection of a temporary state in the cell's rewiring back to normal. This was confirmed by *in vivo* experiments performed on the mammary glands that were taken 9 weeks after inactivation of the oncogenes.

Knowledge on metabolic properties of residual cells mainly comes from studies on CSCs. While they certainly provide important insights into characteristics and potential vulnerabilities of MRD, there are uncertainties revolving around CSC concept: a need for more specific markers, some contradictory reports on their phenotype and the results from xenograft transplant assays, as well as open questions about their significance in clinical practice (112). Since specific markers for CSCs are still not well defined, the existence of other residual subpopulations that have not been studied yet, is likely; the specific environmental conditions and different selective pressure could also drive the

evolution of clones with not necessarily stem-like features or stem cell origin that capable of surviving the treatment. In this regard, our research on the residual cell population obtained upon inactivation of MYC and Neu/Her2 oncogenes relevant in human breast cancer is unique and provides additional knowledge on MRD.

In summary, enhanced glycolysis and urea cycle/arginine metabolism, found in tumor cells, seem to be preserved pathways in the residual cells and upregulated in comparison with normal cells. Therefore, the overexpressed genes involving key enzymes in these processes would be worthwhile exploring as potentially interesting therapeutic targets.

4.3.5 Correlation with human datasets of MRD reveals potential relevance of these findings in the patients' settings

To understand the significance of upregulated glycolysis and urea cycle, as well as enhanced NOS2 activity in MRD in human breast cancer, two different microarray datasets (publicly available) were analyzed. The datasets included: tumor (pre-treatment), residual (post-treatment) and normal (from healthy women) tissue samples. PCA showed that they clustered in three separate groups reflecting the general pattern from our *in vitro* data. We also saw that samples from different tumor types clustered together, despite differences in their molecular biology and chosen treatments. This indicates that the residual cell population as a whole, irrespective of the heterogeneity within, possesses distinct features on the transcriptional level. It could be that the residual cells are indeed intrinsically different, smaller subpopulation, highlighting the problem of tumor heterogeneity and posing the need for finding better therapeutic combinations, aimed not just at the bulk of the tumor, but eradication of all the cancer cells within. Another possibility is that the residual cell population encompasses cells that were able to evolve under the selective pressure of the applied therapy, adjusting and surviving the harsh conditions, lingering in the body in the dormant state. Overall, the clustering of the analyzed human samples showed that the residual cells were different form the primary tumor, yet still with tumorigenic potential as they were also not normal cells. However, it would be interesting to see if the metabolic profiles of MRD would exhibit the similarity to the tumor. If so, this would highlight the importance of designing good and selective metabolic drugs.

Comparison of the differentially expressed genes between the residual (post-treatment) basal subtype and normal samples revealed the upregulation of several genes involved in glycolysis and urea cycle/arginine metabolism. The basal subtype group was chosen because it showed the closest clustering to the mouse dataset. This is in agreement with the studies correlating mouse and human breast tumors which show that mouse

mammary tumors with *MYC* overexpression are close to the human basal-like breast cancer subtype (143). In our analysis, mouse dataset was even closer to the particular basal Her2+ sample, which was not surprising since both *MYC* and Her2 were activated during tumorigenesis. In basal Her2+ sample we identified *ASS1*, *ARG1*, *NOS*, *HK2* and *LDH* as upregulated genes matching the mouse genes involved in glycolysis and urea cycle/arginine metabolism.

From the mouse genes involved in glycolysis and urea cycle that we identified as potential targets, upregulation of *ASS1*, *ARG1*, *NOS*, *HK2* and *LDH* was found in this subtype. Unfortunately, basal Her2+ group was represented with only one sample in the dataset, which poses the need for evaluating these findings on higher number of samples. However, these alterations, although to a lesser extent (*ASS1*, *GAPDH*, *ENO* and *LDH* upregulated) were found in basal Her2 negative samples. It could be that these features are indeed specific for this subtype, as altered glycolysis was not found in luminal A and B samples. This is consistent with the findings in tumors where glycolysis seems to be the feature of basal-like and HER2-enriched tumors (172).

4.4 METABOLIC TARGETS IN MINIMAL RESIDUAL DISEASE

Inter- and intratumor heterogeneity represents a major hurdle in cancer treatment. Heterogeneity stems from different levels – gene mutations, different patterns of gene expression and different interactions with the microenvironment. While metabolic heterogeneity also exists, looking for the vulnerabilities on the metabolic level might help to reduce this immense complexity to a smaller number of metabolic solutions and metabolic nodes at the intersections of major metabolic pathways crucial for the survival of cancer cells.

Treating the MRD on the metabolic level could eventually lead to prevention or delay of tumor recurrences, as some studies show success in interfering with metabolic pathways in the residual population:

- metformin leads to loss of stemness traits in CSCs dependent on mitochondrial metabolism (158);
- inhibition of EMT and loss of stem-like properties in breast cancer cells is achieved through application of a wide spectrum antibiotic, doxycycline (173);
- another novel synthetic-lethal approach through combined therapy with doxycycline and vitamin C showed efficacy in targeting CSCs (174);

- drugs aimed at glycolytic enzymes and transporters could be applied in treatment of residual cells with glycolytic profiles (175);
- utilization of ROS scavengers lowered the DNA damage and application of mifepristone halted hormone-driven proliferation leading to a delay of tumor relapse in a mouse model (119).

Exploring metabolic landscape of MRD in breast cancer model helped us identify several potential vulnerabilities of the residual cells (summarized in Figure 51), mainly enhanced glycolysis, urea cycle and NOS2 activity.

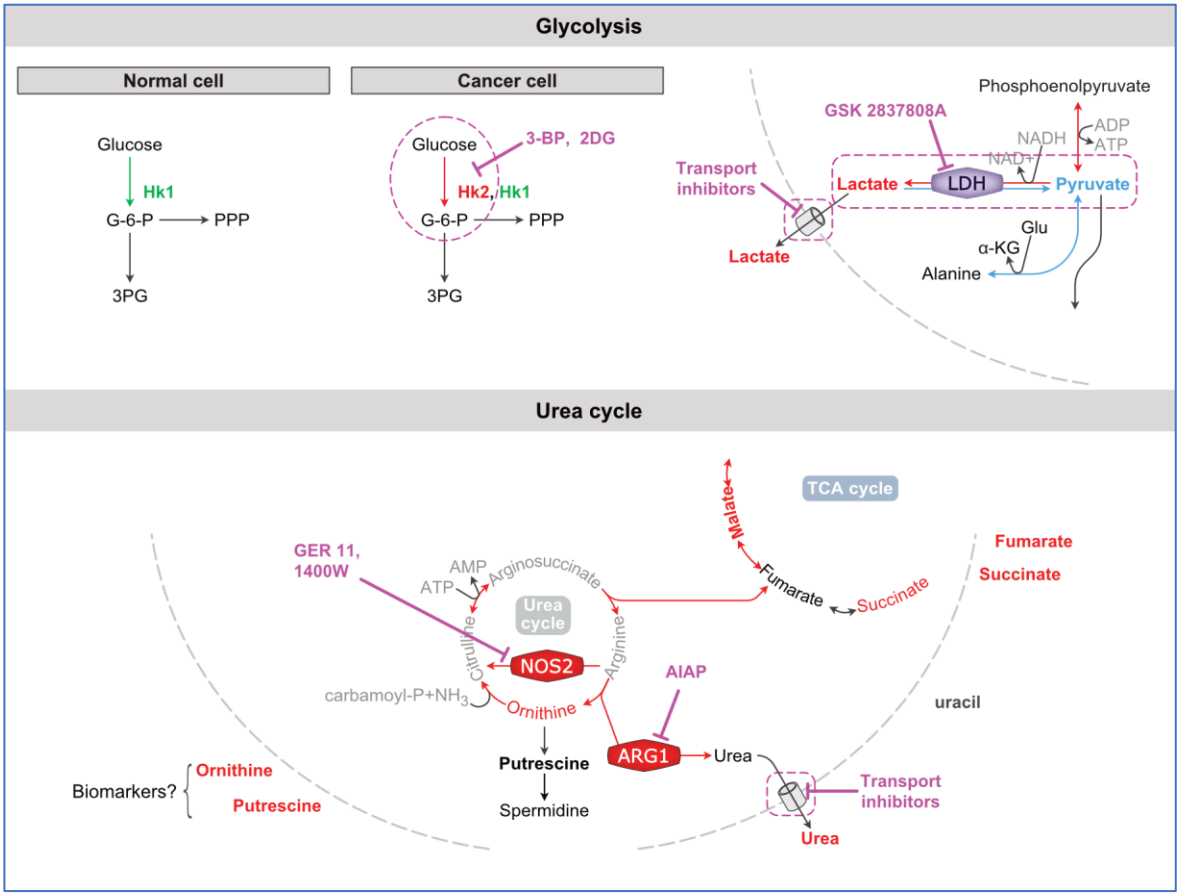

Figure 51: Potential targets in MRD

The identified metabolic features, characteristic for MRD, were also partly mirrored within the tumor cells, but different and distinct from metabolism of the normal cells. Therefore, this opens the possibility for selectively hitting residual and tumor cells with the metabolic therapy. Since increased glycolysis is also a feature of some normal cell types, the challenge in finding the proper therapeutic window remains. With this regard,

targeting HK2 could be one interesting option since this isoform is characteristic for cancer cells and expressed in only a few normal cell types) and inhibitors for HK2 already exist. Interfering with the glycolysis at its end reaction with LDH inhibitors or inhibitors of lactate transport could be another therapeutic option. Finally, we believe that the metabolites of urea cycle could be potentially significant as biomarkers (such as putrescine which is already used as biomarker for some cancers) for monitoring the MRD with the development of very sensitive methods as well as the potential targets, especially since it is not found in many other cell types. Still, finding a therapeutic window that would spare the liver as a main organ where urea cycle is very active, will pose a further challenge.

4.5 IS THERE A METABOLIC MEMORY IN THE RESIDUAL CELLS?

Since residual cells are so profoundly changed, having a distinct transcriptional and lipid profile and metabolic hallmarks similar to tumor cells, we refer to this phenomenon as "metabolic memory". This is substantiated by the fact that we profiled these cells 7 days after oncogenes were inactivated in our 3D cultures and 9 weeks following oncogene inactivation *in vivo*. Moreover, in our previous study, the key features of deregulated lipid metabolism were also verified in mouse tissues of the regressed gland long after oncogene signaling and the patients' samples after neoadjuvant treatment (119). Concept of metabolic memory is used to described a phenomenon observed in diabetes – prolonged exposure to high glucose levels leads to diabetic complications even after its control has been successfully achieved by therapy (176), while early and aggressive treatment reduces this risk. Similarly, it could be that the prolonged exposure to the oncogenic metabolic changes establishes the memory, preventing the cells to return to their normal, physiological state even in the absence of oncogene signaling.

As metabolic reprogramming takes place during tumorigenesis, we hypothesized that there is a threshold, a "point of no return", when these changes leave a profound effect on the cellular network. While majority of cells in our 3D culture die upon oncogene withdrawal, this was not the case with the cells that were exposed to oncogene action for a shorter time (60 hours instead of usual 5 days). Seeing this different response, we thought that these cells would be able to return to the normal state. We followed the phenotypes using IF and collected the RNA from 8 timepoints during tumorigenesis and 9 timepoints during tumor regression, which enabled us to closely follow these processes. Initial analysis of the first three replicates showed us that, despite initial differences in the response to oncogene inactivation, 7 days after oncogene inactivation cells that were exposed to oncogenes for 60 hours in the end clustered together with the

ones from the parallel experiment (longer exposure, 5 days). Seeing this made us introduce another experimental condition, where the oncogenes would be switched off even earlier (after 36 hours). We termed the three conditions as late, mid and early regression (described in Chapter 3.8). In the early regression we observed the same response to oncogene withdrawal as in the mid regression (no apparent cell death, just reintegration in the existing rim). PCA plot based on the RNA-Seq data obtained from the many timepoints during oncogenesis and regression demonstrated that the early regression cells seem to eventually return to the normal state (after being exposed to oncogene action for 36 hours). Thus, we identified this as the "point of no return" after which metabolic memory is exerted.

This opened several questions. First, what are the major metabolic events in the tumor progression and what could be the metabolic processes important in their survival? The comparison of the kinetics of oncogenesis leading to 36 hours, 60 hours and 5 days, with enrichment of metabolic pathways (analysis in progress), could potentially identify processes possibly important in leaving the "legacy", a memory that would change these cells so profoundly, preventing its return to the normal state. In the same way, zooming in the differences between the three regression trajectories (early, mid, late) could reveal features or targets crucial in their survival.

On the other hand, looking at the endpoints (residual cells 7 days after oncogene withdrawal only) in comparison to the normal cells and one another, would help detect changes important in establishing the future tumorigenic potential of residual cells; their ability to subsequently drive tumor recurrence. This could help filter out the noise that could be coming from other confounding processes causing changes with respect to the normal cells.

Another relevant question remaining to be answered is if the "early regression" cells really restore their metabolism, as indicated in their transcriptomic profiles. To test this, we measured their metabolic profiles *in vitro* and *in vivo*. Analysis of this data is in progress and will provide an important clue about the oncogene exposure effect and establishment of metabolic memory.

Nevertheless, we have first indications from our mouse colony: Kaplan-Meier curve showed that the mice that had been exposed shorter to the oncogene action (5-10 days versus 4-6 weeks) had significantly less relapse cases, indicating that these cells indeed have the ability to rewire to their normal state, establishing normal metabolic features. Overall, this highlights the need for better diagnostic tools or biomarkers able to detect breast tumors in the earlier stages of development.

DISCUSSION

Probably the most pressing question is: which molecular entity is the driver of metabolic memory? As metabolites play a part in epigenetic processes, we hypothesize that epigenetic changes are behind the phenomenon of metabolic memory. Interestingly, many genes encoding enzymes involved in epigenetic processes (such as DNA methyltransferases) were upregulated upon oncogene activation. This directed us to obtain the DNA methylation profiles from residual, tumor and normal cells. In addition, mechanisms explored in diabetic metabolic memory point to the DNA methylation (177). We obtained DNA methylation profiles by employing a new method of bisulfite-free sequencing. Data analysis is in progress. Comparison of the residual cells with the normal cells could show the extent of methylation changes, while the comparison with tumor cells could indicate if the specific methylation patterns were already established in the tumor state and preserved in residual, providing a first proof for epigenetic mechanism behind oncogenic memory.

4.6 FUTURE PERSPECTIVES

To answer the questions raised in Chapter 4.5, the following experiments were performed and data was collected (summarized in Table 6). The results of the bioinformatic and GCMS analysis were not available at the point of this thesis submission.

Table 6: Summary of the experiments in progress

Data collection	**Experimental data**	**Perspective**	**Data analysis**
Oct 2016	Mid-regression (60h-7 days) kinetics: RNA Seq and IF images	GO terms and metabolic enrichment to reveal important metabolic pathways changing along the kinetics of mid-regression	Bioinformatic analysis ongoing
Oct 2016 - Nov 2017	Early regression (36h-7 days) kinetics - RNA Seq and IF images	GO terms and metabolic enrichment to reveal important metabolic pathways changing along the kinetics of early-regression	Bioinformatic analysis ongoing
Oct 2016 - Nov 2017	5d-60h-36h regression trajectories - RNA Seq	comparison of late-, mid- and early regression trajectories to filter out the most important metabolic changes	Bioinformatic analysis ongoing
Oct 2016 - Nov 2017	RNA Seq + GCMS + untargeted MS + lipidomics	building a model based on RNA Seq data with measured metabolic constraints to predict the fluxes	Bioinformatic analysis ongoing
Oct 2016 - Nov 2017	RNA Seq + GCMS + untargeted MS + lipidomics	statistical integration of all the available datasets using MOFA; further integration at the pathway-level; Metscape analysis	Bioinformatic analysis ongoing
Nov 2017 - Oct 2018	GCMS measurements	obtaining metabolic picture of early regression endpoint/testing if there is a metabolic signature in *vitro*	GCMS analysis ongoing
Nov 2017 - Jun 2018	DNA methylome - bisulfite free - sequencing	testing the concept of metabolic memory through analysis of DNA methylome	Bioinformatic analysis ongoing
Oct 2017 - Oct 2018	GCMS, ^{13}C glucose labelling, NOS assay	obtaining metabolic picture of early regression endpoint/testing if there is a metabolic signature in *vivo*	GCMS analysis ongoing

DISCUSSION

The analysis of the collected data will help us to better understand the metabolic changes during tumor progression and their potential significance in the regression and survival of the residual cells. Transcriptomic and metabolomic data integration and overlay of the profiles of residual cells that were exposed to oncogene action for the different time periods, will point to processes that could be of relevance in tumor recurrence. Furthermore, metabolic profiles of the residual cells that were exposed to oncogene signaling for a short time, obtained from *in vitro* and *in vivo* experiments, will provide an answer about whether the transcriptomic differences observed are also reflected on a metabolic level, compared with other time points. Comparison of the three regression trajectories could highlight the pathways important for survival of these cells upon oncogene silencing and their potential relevance in escaping oncogene addiction. A model based on the RNA-sequencing data, with the constraints from metabolic measurements, will enable the flux predictions and improve our understanding of the metabolic picture in MRD. Further statistical integration of all the available datasets will bring another level of validation and could help to reveal other potentially interesting targets. DNA methylome analysis will shed a light on our hypothesis of epigenetic changes behind the observed metabolic memory.

To further understand the role of enhanced glycolysis, urea cycle and NOS2 activity in MRD, we plan to perform the experiments where we would interfere with these features using the inhibitors (summarized in Table 7).

The readout of these experiments would be:

1) untargeted and targeted metabolomics analyses,

2) viability assays,

3) and cell death assays.

Metabolomics analyses could reveal possible alterations of metabolic profiles in these cells as a result of interference with the targets; whether they would be able to switch to the alternative metabolic pathways and, if so, which routes they would engage.

Viability and cell death assays would be used to reflect if there is an addiction of the residual cells to the observed processes and if targeting these pathways would eradicate the residual cells. After the *in vitro* experiments, we would eventually test the successful inhibitors *in vivo*.

Table 7: Planned experiments for interference with the observed targets in MRD

Planned experiments	Inhibitors	Readout
Interference with glycolysis	GSK28378088 3-bromopyruvate acriflavine	1. untargeted metabolomics / GCMS 2. viability assays 3. cell death assays
Interference with the urea cycle	AIAP ABH Ammonium salt	
Interference with NOS2	GER 11 Piperidine	

Chapter 5

CONCLUSIONS

CONCLUSIONS

Using a multi-omics approach, we have characterized the refractory, residual cell population that was left following oncogene inactivation in the breast tumor cells. To model this cellular state, we employed a tractable mouse model and primary mammary cells derived from those mice, which were grown in a 3D cell culture system.

Despite displaying a seemingly normal phenotype, residual cell populations exhibited a distinct transcriptional profile, in comparison to normal cells. Nevertheless, changes on the transcriptomic level also differed from the tumor state, indicating a unique nature of the residual cells.

The lipidomic profile of residual cells was also unique, hinting at overall changes in the cell phenotype, which were also reflected in the lipid composition.

The overall metabolic profile of residual cells bore resemblance to the tumor, in contrast to the differences observed on the transcriptional and lipid level. Residual cells retained some metabolic features that were also found pronounced in the tumor state and therefore seemed to be preserved long after oncogene inactivation. This reflects the profound and vast effect of the changes during tumor progression and alterations in the metabolic network, implying the existence of "metabolic memory". This phenomenon, particularly its potential driver, needs to be further understood. We hypothesize that these features are established through an altered epigenetic landscape.

Enhanced glycolysis, urea cycle and NOS2 activity were the most prominent features preserved in the residual cells. Moreover, these traits could be particularly important in the MRD of basal HER2 positive and HER2 negative breast cancer subtypes. The observed metabolic differences between the residual and normal cells should be further explored and could provide an interesting insight into metabolic nodes that serve as potential interference points against MRD. Interference of these metabolic nodes could potentially be used in a therapeutic setting.

The kinetic analysis of oncogenesis and the respective phases that led to tumor regression indicated that the length of exposure to oncogene action could have a significant effect. It could equip the residual cells with a mutagenic potential, which could develop into a tumorigenic potential over a longer time period, enabling the cells to form recurrent tumors. Consequently, the detailed molecular profiles that we obtained using the tractable 3D culture system of MRD formation might have clinical relevance in the future. One possible application might be their use in further classifying tumor staging. This could be valuable because of the importance of early breast cancer detection, and the need for more sensitive diagnostic procedures.

Chapter 6

MATERIALS AND METHODS

6.1 ANIMALS

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. 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.

The mouse strain used was *TetO-cMYC/TetO-Neu/MMTV-rtTA* (121) in FVB background. Genotypes of the animals were determined by PCR on genomic DNA from the tail tissue, obtained by tail-digestion in 75 µl of digestion buffer (NaOH 25 mM + EDTA 0,2 mM) at 98°C followed by addition of 75 µl Tris-HCl (40 mM, pH 5,5) and centrifugation at 4000 rpm for 3 min, whereby 1 µl of the supernatant was used in the PCR reaction. Reagents used for PCR are summarized in the Table 8 with primer sequences provided in Table 9 and PCR programs in Table 10. 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, A9539-500G) gel with Ethidium bromide solution in a final concentration of 0,5 µg/ml (Sigma, E1510-10ML). The products were visualized using Quantum-Capt1 documentation system + instrument (Peqlab).

Table 8: PCR components for genotyping

Reaction component	**Final concentration**
Primer F (Forward)	0,25 pMol/µl
Primer R (Reverse)	0,25 pMol/µl
dNTP mix	200 µM
Enzyme (Taq polimerase by EMBL)	1 U/20 µl
Dream Taq Buffer (ThermoFisher Scientific, B71)	1x

Table 9: Primer sequences for genotyping of TetO-Myc/TetO-Neu/MMTV-rtTA animals

Primer	**Forward sequence**	**Reverse sequence**
TetO-*MYC*	TAG TGA ACC GTC AGA TCG CCT G	TTT GAT GAA GGT CTC GTC GTC C
TetO-Neu	GACTCTCTCTCCTGCGAAGAATGG	CCTCACATTGCCAAAAGACGG
CMV-rtTA	GTGAAGTGGGTCCGCGTACAG	GTACTCGTCAATTCCAAGGGCATCG

Table 10: PCR programs for TetO-Myc/TetO-Neu/MMTV-rtTA

	TetO-*MYC*			TetO-Neu			MMTV-rtTA	
Step	**T**	**T**		**T**	**t**		**T**	**t**
1	94°C	3 min		95°C	1 min		94°C	5 min
2	94°C	10 s		95°C	15 s		94°C	30 s
3	62°C	30 s		64°C	15 s		57°C	30 s
4	68°C	1:30 min		72°C	1:30 min		72°C	30 s
5	94°C	10 s		95°C	15 s		72°C	5 min
6	60°C	30 s		61°C	15 s		10°C	hold
7	68°C	2 min		72°C	1:30 min			
8	72°C	10 min		95°C	15 s			
9	4°C	Hold		58°C	15 s			
10				72°C	1:30 min			
11				95°C	15 s			
12				55°C	15 s			
13				72°C	1:30 min			
14				72°C	10 min			
15				10°C	hold			

6.1.1 *In vitro* experiments

Primary mammary epithelial cells were obtained through harvest of mammary glands from 8 weeks old virgin females, whereby mammary gland tissue was digested and epithelial cells were selected and cultured in 3D.

6.1.2 *In vivo* experiments

For induction of tumors, food pellets supplemented with doxycycline (625 mg/kg) was given to the mice. Animals were weekly monitored for the tumor detection and their overall health. Full blown tumors developed in the period of 4-6 weeks. When the tumor burden was too large (d=2 cm), animals were given food without doxycycline which resulted in the fast tumor regression to a non-palpable state. At the timepoint of the complete tumor regression (9 weeks after doxycycline withdrawal, i.e. oncogenes deactivation), mammary glands were harvested. Wild-type (non-inducible) siblings were given the same food in the exact same fashion. They represented age-matched controls and their mammary glands were collected at the same time and in the same manner. Before harvesting, animals were also checked for the phase of estrous cycle, by taking vaginal smears according to the modified protocol (178): 20-40 ul 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, V5265) washed in tap water and visualized using Leica Application Suite X and Leica DFC7000 T microscope (Leica Microsystems).

To assess the effect on shorter exposure to the oncogenes on metabolic profiles of the regressed glands, mice were given doxycycline supplemented food for a period of 10 days and then returned back to a normal diet. After 10-12 months, mammary glands of these animals and corresponding age-matched controls were harvested for the experiment.

For examining if there was a difference between shorter and longer exposure to the oncogenes action in the occurrence and time-period of tumor relapses, animals that were not taken for the experiment (with their age-matched controls) were weekly monitored for the emergence of tumor recurrence.

6.2 3D CELL CULTURE

Three-dimensional cell cultures were established according to the published protocol (118) with some modifications. Primary mammary epithelial cells were obtained from 8 weeks old virgin females of described mouse strains, through digestion of mammary glands in 5 mL of digestion media (Lonza/Amaxa DMEM/F12 1:1 Mixture with HEPES, L-Gln, BE12-719F), supplemented with HEPES to the final concentration of 25 mM, 150 U/mL Collagenase type 3 (Worthington, LS004183), 20 µg/mL Liberase Blendzyme 2 (Roche, 05401020001) and 5 ml of Penicillin/Streptomycin (Gibco Life Technologies, 15140-122). After digestion for 15-16 hours at 37°C in 5 % (vol/vol) CO_2 atmopshere, in loosely capped 50 mL polypropylene conical tubes, washing step with 45 mL of phosphate-buffered saline (PBS) was performed. Upon centrifugation at room temperature, 1000 rpm for 5 min, interphase between upper fat layer and cell pellet was removed and 5 mL of 0.25 % trypsin-EDTA (Invitrogen, 25200-056) was added. Suspension was incubated for 40 min at 37°C, 5 % CO_2 in loosely capped tubes, followed by the wash with 25 mL of STOP media (Lonza/Amaxa DMEM/F12 1:1 Mixture with HEPES, L-Gln, BE12-719F supplemented with HEPES to the final concentration of 25 mM and 10 % Tet System Approved Fetal Bovine Serum, Biowest, S181T) and treatment with 5-15 mg/mL DNase I (ThermoFisher, 18068015). After another centrifugation step at room temperature, 1000 rpm for 5 min, dissociated cells were resuspended in MEBM media (Lonza, Mammary Epithelial Cell Basal Medium CC-3151 with supplements from Mammary Epithelial Cell Medium BulletKit CC-3150) and plated onto collagen-coated plates (BD Biosciences, 356400) for selection of epithelial cells. Next day cells were

washed with PBS and the remaining ones treated with 500 µl of 0,25 % trypsin-EDTA until detachment. Trypsin was inactivated with 9 mL of STOP media (described above), followed by centrifugation step at room temperature, 1000 rpm for 5 min. Cell pellets were resuspended in PBS, counted, and mixed rapidly on ice with the prepared Matrigel-collagen mixture – Cultrex 3D Culture Matrix Basement Membrane Extract (Biozol, TRE-3445-005-01) and 1,5 mg/mL Cultrex 3D Collagen I rat tail (TEMA Ricerca, 3447-020-01). Mixed droplets in volume of 100 µl containing 12 500 primary mouse mammary epithelial cells were dispensed into flat bottom wells (Corning CellBIND 12 Well Clear Multiple Well Plates, 3336) or chambered cover glass slides (ThermoFisher Scientific, Nunc LabTek II Chambered Cover glass, 155379). After gel solidifying for 35-40 min at 37°C, 1.5 ml of MEBM serum-free media (supplemented with 2 mL of bovine pituitary extract, 0.5 mL of hEGF, 0.5 mL of hydrocortisone, 0.5 mL of GA-1000, 0.5 mL Insulin from Mammary Epithelial Cell Medium BulletKit CC-3150) was added to each well. Doxycycline (Sigma, Doxycycline hyclate, D9891) was titrated to lower concentration of 200 ng/ml. For metabolic analyses, from the start of experiment until collection, media was used in volume of 1 mL and changed every day at the precise times.

6.3 IMMUNOFLUORESCENCE

3D culture gels for immunofluorescence staining were fixed with 4 % paraformaldehyde (PFA) for 7-10 min and transferred to the IF deactivated clear glass screw neck vials (Waters, 186000989DV). They were washed three times with PBS and once in IF buffer (containing NaCl, Na_2HPO_4, NaN_3, BSA, TritonX-100, Tween-20; pH 7,4). Blocking was done using 1x IF buffer with 10 % goat serum (Jackson Immuno Research, 005-000-121) for 1,5 h. Incubation at 4°C overnight was done, with the primary antibodies diluted in primary block as described above and washed the next day in 1x IF buffer 3x, 15 min each. Incubation with secondary antibodies and 4', 6'-diamino-2-phenylindole (DAPI) was done in the primary block for 1 hour (dilution 1:1000) and followed by the washes with 1x IF buffer and 1x PBS, 3 times, 15 min each. Gels were mounted with Vectashield Anti-fade mounting medium (Vinci Biochem, VC-H-1500-L010) into LabTek II chamber slide (ThermoFisher Scientific, 50733) and imaged on a Leica SP5 confocal microscope using 63x water lens and LAS AF imaging software.

The following antibodies were used for the 3D cultures: alpha-6-integrin (BD Biosciences 25-0495-82, diluted 1:80), ZO-1 (Invitrogen 61-7300, diluted 1:500), GM-130 (BD Biosciences, 610823, diluted 1:100), E-cadherin (Invitrogen, 13-1900, diluted 1:200). Nuclei were stained with DAPI (ThermoScientific, 62248, diluted 1:1000). Anti-rabbit, anti-mouse, and anti-rat antibodies were purchased coupled with Alexa Fluor dyes from

Invitrogen (A21247, A11034, A11036). FFPE tissue sections were stained using the standard protocols for the following antibodies: HK2 (ThermoFisher Scientific, PA5-29326, diluted 1:250), PDK1 (Invitrogen, MA5-15797, diluted 1:200), MCT4 (Santa Cruz Biotechnology, sc-376140, diluted 1:50), AK4 (ThermoFisher Scientific, PA5-49333, diluted 1:25), LIN28A (LifeSpan BioSciences, LS-B14215, diluted 1:100), ASS1 (Abcam, ab170952, diluted 1:50), ARG1 (Novus, NBP1-32731, diluted 1:250) and iNOS (ThermoFisher Scientific, PA1-036, diluted 1:100).

Sections were mounted using ProLong Gold Antifade Mountant (ThermoFisher Scientific, P36930) and scanned using TissueFAXS Slides system (TissueGnostics). Quantification was done using StrataQuest Analysis Software (TissueGnostics).

6.4 IMMUNOHISTOCHEMISTRY

Dissected mammary glands were fixed in 10 % formalin solution, neutral buffered (Sigma, HT501128-4L) and subsequently washed with PBS. Tissue sections from FFPE samples were obtained by sectioning at 5 µm with RM 2135 microtome (Leica) and placing on the Superfrost slides (VWR, 630-0954). If frozen tissue was used for imaging mass spectrometry, it was cut at 4-18 µm with cryotome (Leica, CM3050 S). Staining by Hematoxylin QS (Vector, H-3404), Eosin 1 % Aqueous (RA Lamb, LAMB/100-D) was done following the standard protocol and samples were analyzed at LMD 7000 microscope (Leica). MT-CO1 antibody (Abcam, ab14705) 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, H-3300) for 30 min in a steamer and inactivation of endogenous hydrogen peroxidase activity with 10 % H_2O_2 solution (Sigma, H1009). After this, blocking was done using 10 % Normal goat serum (Jackson Immuno Research, 00500121) in 1x Phosphate Saline Buffer (PBS). Incubation with primary antibody (MT-CO1 diluted 1:500) was done in blocking buffer at 4°C overnight, after which the slides were washed 3x, 5 min each, using 1x PBS. Incubation with biotinylated antibody from Vectastain Elite ABC HRP Kit (Peroxidase, Rabbit IgG; Vector Laboratories, PK-6101) was for 30 min, followed by washing and incubation with Horse Radish Peroxidase (HRP) conjugated antibody and detection using DAB Peroxidase (HRP) Substrate Kit (Vector, SK-4100). Counter-staining was done using hematoxylin, after which the sections were dehydrated, mounted with DPX Mountant for histology (Sigma, 06522) and analyzed using LMD 7000 microscope (Leica) equipped with Leica CD310 digital camera and LASV3.7 (Leica) software.

6.5 QUANTITATIVE REAL-TIME PCR

Total RNA from the 3D culture gels was extracted using Trizol (Invitrogen, 15596026). Complementary DNA (cDNA) was prepared from 200 ng of RNA, using Super Script® III Reverse Transcriptase kit (Invitrogen, 18080093). Real-Time PCR reaction was done using SYBR Green PCR Master Mix (Applied Biosystems, 4309155), 2 ng of DNA per reaction and 10 µM primers. Primer sequences and PCR program is shown in Table 11. Runs were performed on a StepOne (Applied Biosystems) machine with technical triplicates. The fold changes in gene expression were calculated using the ΔΔCt method. Actin-beta (Actb) was used as a reference.

Table 11: Primer sequences and program for Real-Time qPCR

Primer	**Forward sequence**	**Reverse sequence**
Actb	ATGCCCTGAGGCTCTTTTCC	ACGGATGTCAACGTCACACT
TetO-*MYC*	GCGACTCTGAGGAGGAACAAGA	CCAGCAGAAGGTGATCCAGACT
TetO-Neu	CGTTTTGTGGTCATCCAGAACG	CTTCAGCGTCTACCAGGTCACC
Program steps	**Temperature (T)**	**Time (t)**
1	95°C	5 min
2	95°C	30 s
3	60°C	1 min
Repeating steps 2 and 3 for 40 cycles, followed by Melting curve analysis.		

6.6 WESTERN BLOT

3D culture gels were treated for 1,5 hours with collagenase and liberase for Matrigel disintegration. The disintegrated gels were collected in 15 ml Falcon tubes, shortly centrifuged at 1000 rpm for 5 min and washed in cold 1x PBS (two times; centrifuged at 1400 rpm for 4 min at 4oC), followed by pellet resuspension in protein lysis buffer (0,5 % sodium deoxycholate, 25 mm Hepes, pH 7,5; 150 mm NaCl, 1% Triton-X-100; 0,1 % SDS; 5 % glycerol, 2 mm EGTA) prepared with cOmplete protease inhibitor cocktail tablets (Roche, 11697498001) and PhosSTOP (Roche, 4906845001). After 20 min long incubation on ice, samples were centrifuged at 14 000 rpm for 15-30 min at 4°C. Supernatants were collected and protein concentration measured using Bradford assay and BSA standard curve. Protein lysates (volumes corresponding to 30-40 µg of protein per well) were mixed with 5 µl of Laemmli buffer (6x) and loaded onto wells for SDS-PAGE. Denaturation of protein for MT-CO1 was done at 60°C, in accordance with

manufacturer's instructions. Proteins were separated by SDS-PAGE and blotted onto PVDF membrane and stained with MT-CO1 antibody (diluted 1:500 in TBS-Tween 0,3%) at 4°C overnight. Beta-tubulin (Cell Signaling, #2146) was used as a reference antibody (in dilution 1:1000). The next day, membrane was washed with 0,3 % TBS-Tween and incubated with HRP-conjugated secondary antibody (Goat Anti-Mouse IgG, diluted 1:5000 in TBS-Tween 0,3%) for 1,5 hours at room temperature, followed by washes in 0,3 % TBS-Tween. The development was done using Amersham ECL Western Blotting Detection Reagent (GE Healthcare, RPN2106) and detection using ChemiDoc™ Touch Imaging System (BioRad). The images were processed in Image J.

6.7 TRANSCRIPTOMICS

6.7.1 RNA collection and extraction

RNA was harvested from a pool of two 3D gels per condition, using 900 µl of mirVana lysis buffer, and subsequently extracted using mirVana miRNA Isolation Kit, with phenol (Ambion, AM1560). After assessing RNA quality and concentration on Bioanalyzer (Agilent 2100, G2939BA), RNA was sequenced in the Genomics Core Facility at EMBL on Illumina NextSeq 500 platform, read length NextSeqHigh 75 bp.

6.7.2 RNA-Seq data analysis

After assessing the quality of the raw RNA sequencing reads by FastQC version 0.11.3 (179), adapter trimming using cutadapt version 1.9.1 (180) with default options providing the standard Illumina TrueSeq Index adapters was done. FaQCs version 1.34 (181) 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) (182) using Tophat2 version 2.0.10 (183) 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 (184) 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 (185). 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 cooks cutoff (186) 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 (187) to compute q-values. Genes with q-value<0,05 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 p_{adj}<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 [109] transformed transcript counts were utilized. For the calculation of the ellipses on the PCA plots the "stat ellipse" function from the R package ggplot2 was used [66].

6.7.3 Enrichment analysis

1925 of differentially expressed genes in residual cells (compared to never induced control; FDR-tool (188) q-value<0,05) and 5274 differentially expressed genes in tumor (compared to never induced, p_{adj}<0,1, Bonferroni correction for multiple testing) were taken for the GO enrichment analysis. GO enrichment analysis was performed using Fisher's exact test with a foreground of all respective differentially expressed genes and a background, which was composed of a unique set of 5 randomly picked genes per foreground gene exhibiting a similar expression mean over all samples. The chosen p-value cutoff was 0,01. Genes, whose transcript levels significantly changed (differentially expressed genes) were mapped to biochemical reactions according to the KEGG annotation of their encoding enzymes. In case there are multiple enzymes annotated to the same reaction, which have a conflicting sign of change, the sign of the gene with the strongest log2 fold change is depicted.

6.7.4 Reporter Metabolite Analysis

Reporter Metabolite Analysis was done using non-directional class, the distinct-directional class and the mixed-directional class, using improved model based on the HMR2 model (127). Reporter metabolite algorithm, the "piano" R package (128) with genes as keys and metabolites as values, was used to identify the metabolic reactions likely to be altered. The respective p-value/q-value and log2 FC of the differentially expressed genes were used to calculate p-values from a theoretical null distribution (10,000 permutations). Multiple testing adjustment was applied using the Benjamini-Hochberg procedure. The threshold for significance was p_{adj}<0,1, but maximally 5 % of the total list of tested metabolites. A pathway enrichment was calculated for gene sets of 1 gene per group or bigger.

6.7.5 Analysis of MRD datasets from patients

Microarray gene expression datasets from breast cancer patients pre- and post-treatment (132) and control breast tissue from healthy women (133) were downloaded from Gene Expression Omnibus (GEO) (189). Each dataset was first prepared independently – filtered for outliers and normalized. Background correction was done using "rma" function of the R package "oligo" filtering genes for minimal intensity filtering and gene annotation of probe set IDs with the removal of multiple mappings of transcript cluster identifiers. Two datasets were then combined and processed (normalized, checked for outliers, with genes with low intensities filtered). In order to address the batch effect of the joined data stemming from the two experimental settings which largely influenced the gene expression, the PC1 was removed from PCA. In addition, due to the poorly defined diagnostic category and high biological variability, the "normal-like" tumor subtype of the patients' dataset was removed.

6.8 METABOLOMICS

6.8.1 Collection and extraction of intracellular and extracellular metabolites

6.8.1.1 In vitro experiments

Before proceeding with intracellular metabolites collection, 150 µl of extracellular media was taken from the wells of each condition, snap-frozen and stored at -80°C until analysis. Organoid structures were then freed from Matrigel upon digestion for 1,5 h at 37°C with 3 µl of liberase and 3 µl of collagenase added to the media. 3 wells were pooled per condition/technical replicate, washed three times with PBS, centrifuged shortly (1000 rpm, 2 min, room temperature) and quenched by 200 µl cold (-80°C) HPLC-grade methanol (Biosolve Chimie, 136841). Validation of effective cell washing before quenching was done by comparison of metabolic profiles with 50 µl of the used MEBM growth medium of each well and the last washing solution of each sample (quenched with 100 µl of cold -80°C HPLC-grade methanol). Extraction of metabolites was done by 1:1 methanol-water protocol (190–192), with ribitol (Alfa Aesar, 488-81-3) as internal standard. For *Flow Injection Q-Exactive MS* extracellular metabolites were measured in extracellular media diluted 1:10.

6.8.1.2 *In vivo* experiments

For glucose labeling experiment, mammary glands were dissected, minced and digested for 2 hours at 37°C using collagenase and liberase enzymes, then were cultured for 8 hours at 37°C in 5 % (vol/vol) CO_2 atmopshere, in DMEM glucose- and pyruvate-free

media (ThermoFisher Scientific, 11966025) supplemented with 4,5 g/L labeled D-Glucose U-13C, 99% (Cambridge Isotope Laboratories, Inc., CLM-1396-1) and 2 mL of bovine pituitary extract, 0.5 mL of hEGF, 0.5 mL of hydrocortisone, 0.5 mL of GA-1000, 0.5 mL insulin from Mammary Epithelial Cell Medium BulletKit CC-3150. For non-labeled GCMS metabolomics experiment, mammary glands were dissected and cultured for 8 hours at 37°C in 5 % (vol/vol) CO_2 atmosphere, in DMEM, High Glucose (4,5 g/L glucose) GlutaMAX (Gibco, 10569044) supplemented with 2 mL of bovine pituitary extract, 0.5 mL of hEGF, 0.5 mL of hydrocortisone, 0.5 mL of GA-1000, 0.5 mL insulin from Mammary Epithelial Cell Medium BulletKit CC-3150. Extracellular metabolites were collected and snap-frozen in liquid nitrogen; for intracellular metabolites harvesting, cells were quickly washed two times in PBS and quenched with cold methanol. Metabolites were extracted using 1:1 methanol-water protocol (190–192), with ribitol (Alfa Aesar, 488-81-3) as internal standard.

6.8.2 Untargeted metabolomics by Flow Injection Q-Exactive MS

High-throughput discovery metabolomics was modified from the method by Fuhrer et al. (193). The LC/MS platform used consists of liquid chromatography system (ThermoFisher Scientific Ultimate 3000) with autosampler temperature set to 10°C which is coupled to a Q-Exactive Plus Fourier transform mass spectrometer (ThermoFisher Scientific) equipped with a heated electrospray ion source. The machine was operated in negative ionization mode, with isocratic flow rate of 150 µl/min of mobile phase (60:40 % (v/v) isopropanol:water buffered with 1mM ammonium fluoride at pH 9; with 10nM taurocholic acid and 20 nM homotaurine as lock masses). Recording of mass spectra was done in profile mode from 50 to 1000 m/z. Instrument was adjusted to the following settings: sheath gas 35 a.u.; aux gas 10 a.u.; aux gas heater 200°C; sweep gas 1 a.u.; spray voltage -3kV; capillary temperature 250°C; S-lens RF level 50 a.u.; resolution 70k @ 200 m/z; AGC target 3x106 ions, max. injection time 120 ms; acquisition duration 60s. An automated pipeline in R was used for spectral data processing. Tentative annotation of detected ions as metabolites was done using the Human Metabolome Database (HMDB) as a reference, assuming [M-H] and [M-2H] as ionization options and the exchange of one or two ^{13}C with the equivalent number of ^{13}C atoms with the method-inherent disability to distinguish between isomers.

6.8.3 GCMS analysis

Upon drying, metabolite extracts were derivatized to their (MeOx) TMS-derivatives: 1) with 50 uL of 20 mg/mL methoxyamine hydrochloride (Alfa Aesar, 593-56-6) solution in pyridine (SigmaAldrich, 437611) for 90 min at 40°C, 2) with 100 ul N-methyl-

trimethylsilyl-trifluoroacetamide (MSTFA) (Alfa Aesar, 24589-78-4), for 12 hours at room temperature (192,194). Upon derivatization (after 12 hours), metabolic profiles of samples were measured using a Shimadzu TQ8050 GCMS (triple quadrupole) system (Shimadzu Corp.), equipped with gas chromatograph with a 30 m x 0,25 mm x 0,25 µm DB-50 MS capillary column (Phenomenex, USA) and detector operated both in scan mode (recording in the range of 50-600 m/z) and MRM mode. Samples were normalized to ribitol as an internal standard and total metabolite levels (TML).

For experiment using labelled glucose, metabolites were dried and derivatized to their (MeOx) TMS-derivatives: 1) with 50 µL of 20 mg/mL methoxyamine hydrochloride (Alfa Aesar, 593-56-6) solution in pyridine (SigmaAldrich, 437611) for 90 min at 40°C, 2) with 100 µl N-tert-Butyldimethylsilyl-N-methyltrifluoroacetamide with 1 % tert-Butyldimethylchlorosilane (Sigma Aldrich, 00942) for 1 hour at 60°C and 12 hours at room temperature (192,194) and further processed as previously described.

Calculations of p-values were done using Linear Models for Microarray Data (limma) (195) and Benjamini-Hochberg adjustment method.

6.8.4 . Lipidomics analysis

Modified Bligh & Dyer protocol (196) was used for lipid extraction. Samples were analysed by direct infusion nano-ESI MS, using Qtrap 6500 + coupled with a NanoMate (Sciex). Samples were normalized to 100 % (mol %) of total lipid.

6.8.5 NOS enzymatic assay

Mammary glands were dissected and homogenized in NOS assay buffer and further processed following the Nitric Oxide Synthase Activity Assay kit (Abcam, ab211083) protocol for measuring enzymatic activity of nitric oxide synthase (NOS).

6.8.6 DNA methylome analysis

Following digestion of the Matrigel with collagenase and liberase, DNA was collected from the 3D cultures and extracted following Qiagen protocol for cultured cells (Qiagen Blood & Cell Culture DNA Mini kit, 13323). Bisulfite-free sequencing was performed at the Genomics Core Facility at EMBL Heidelberg.

Chapter 7

APPENDIX

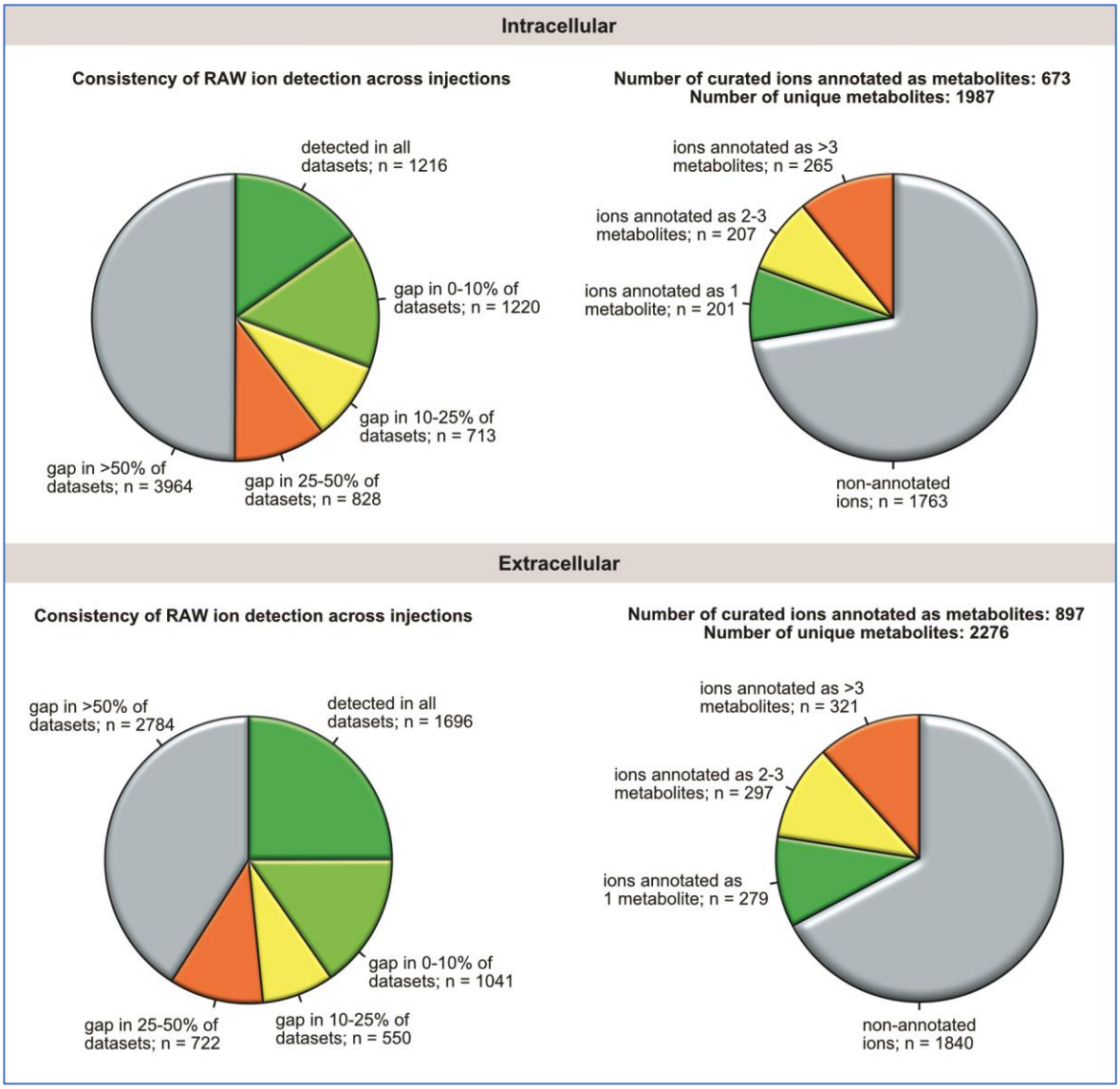

Figure 52: Ions detected by untargeted metabolomics Q-Exactive MS and data quality control. Top left: Consistency of intracellular raw ion detection across injections; Top right: Number of curated intracellular ions annotated as metabolites; Bottom left: Consistency of extracellular raw ion detection across injections; Bottom right: Number of curated intracellular ions annotated as metabolites

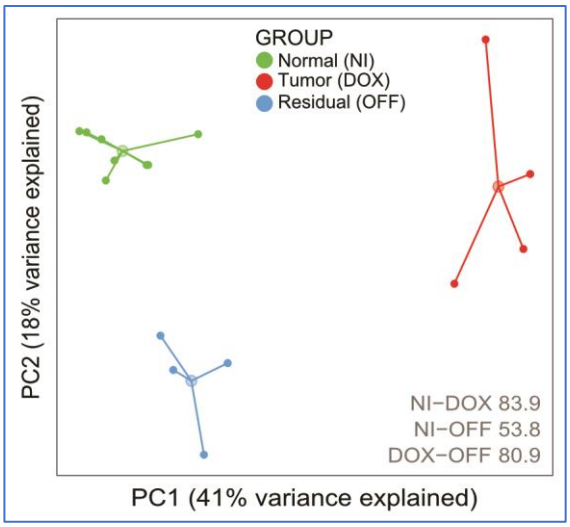

Figure 53: PCA on a subset of metabolic genes

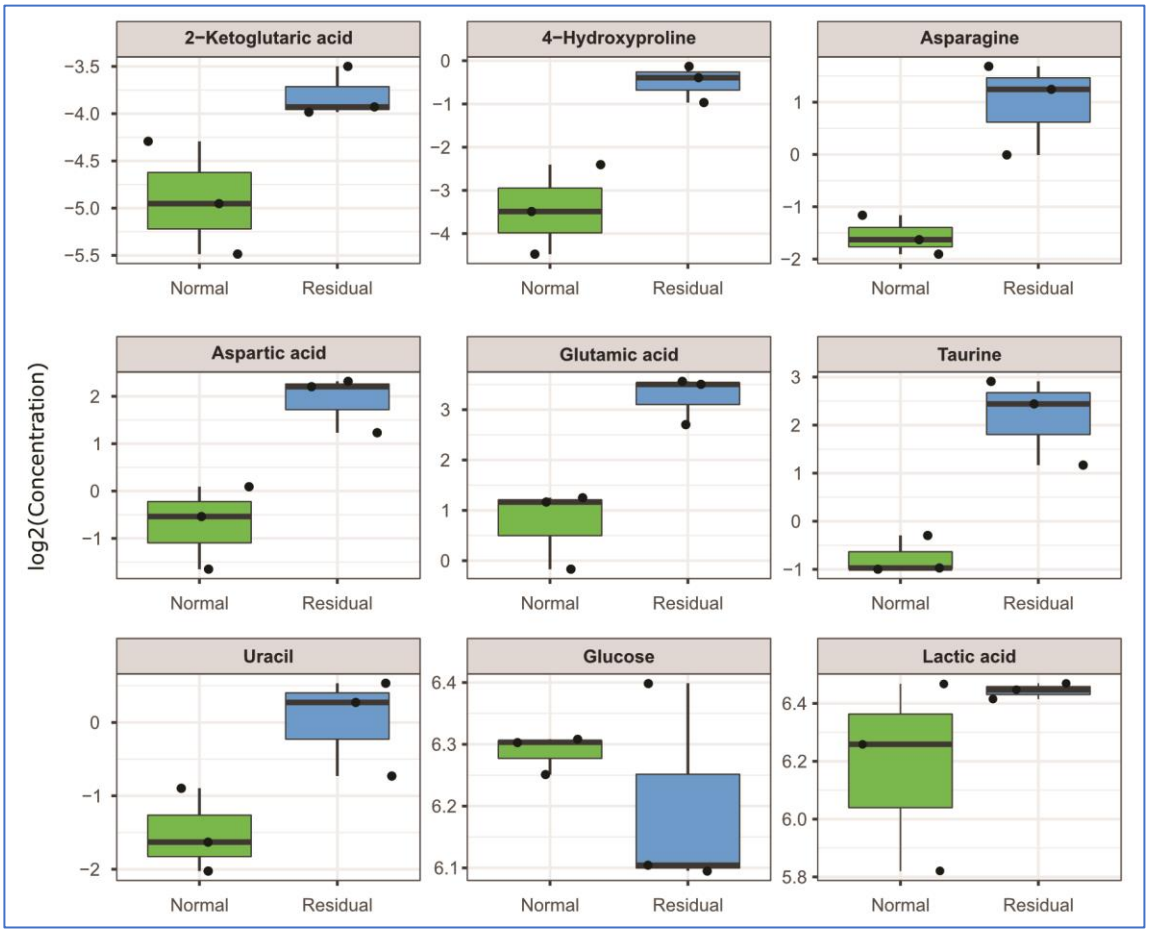

Figure 54: Other significantly changed metabolites measured in vivo

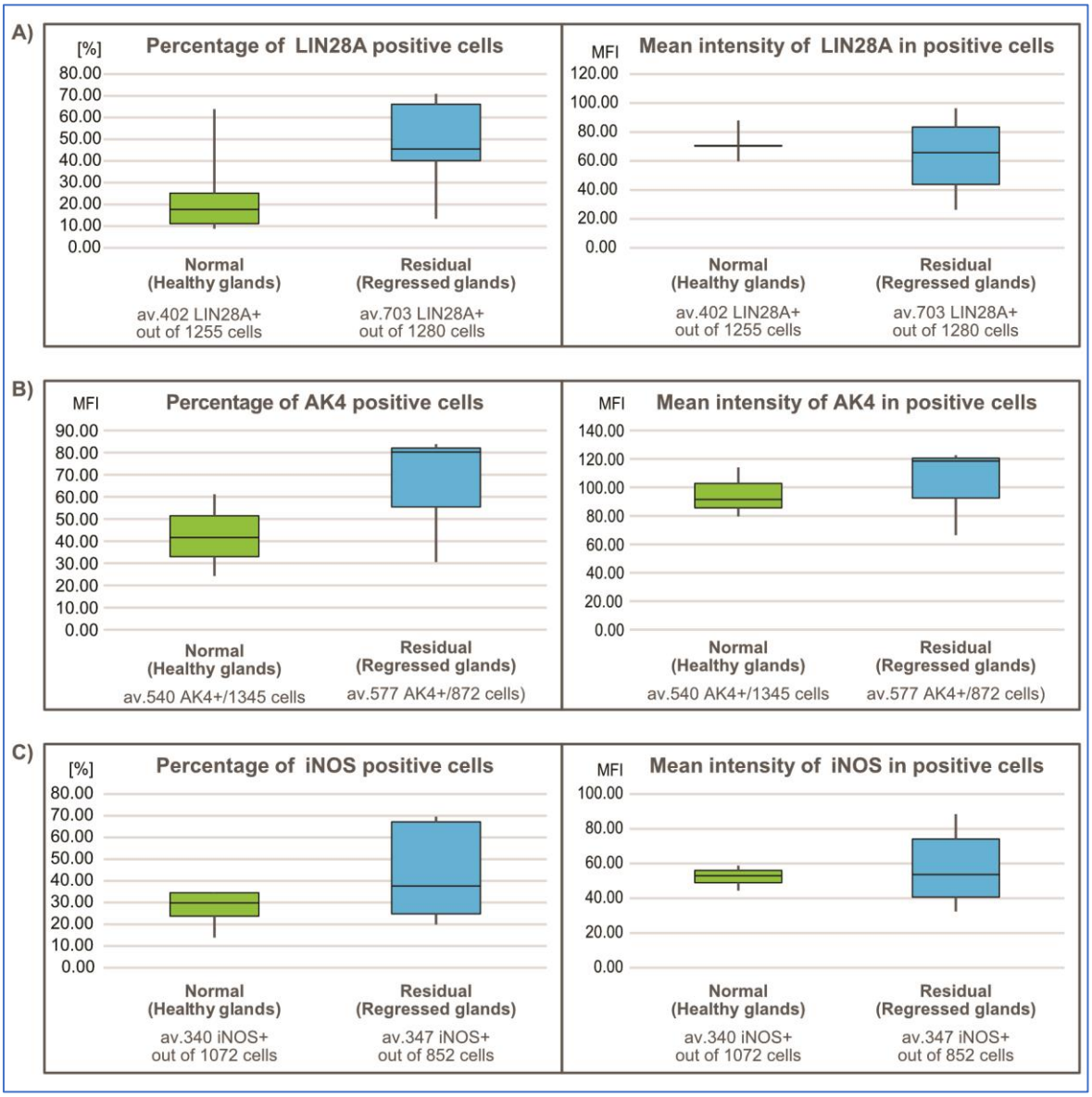

Figure 55: IF stains protein verification in the regressed mammary gland tissue and healthy tissues of the age-matched controls, for: A) Lin28A; B) AK4; C) NOS2.
On the left side are given percentages of positive cells for the respective proteins; on the right – mean intensity of fluorescence in positive cells.

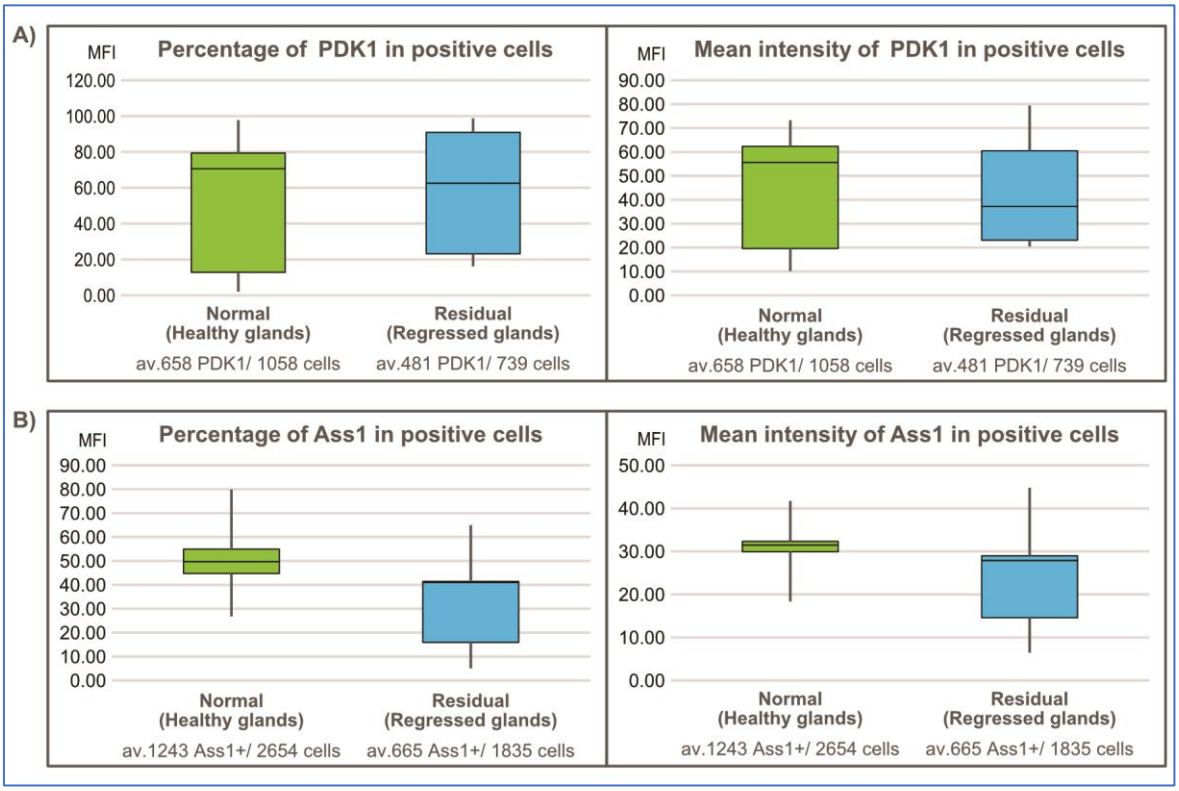

Figure 56: IF stains protein verification in the regressed mammary gland tissue and healthy tissues of the age-matched controls, for: A) PDK1; B) ASS1.
On the left side are given percentages of positive cells for the respective proteins; on the right – mean intensity of fluorescence in positive cells.

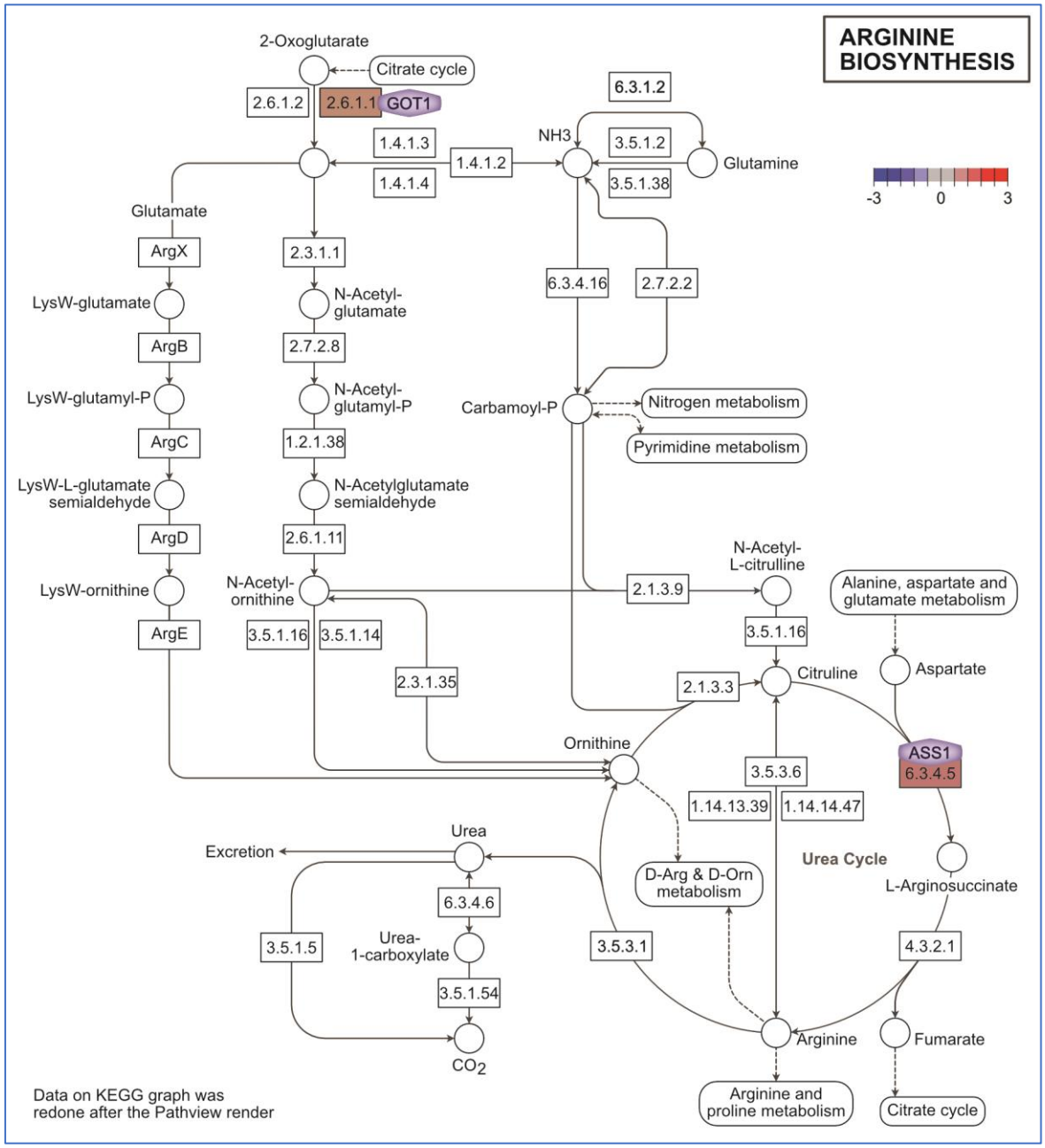

Figure 57: Altered gene expression encoding enzymes involved in arginine metabolism (urea cycle, NO synthesis) in basal HER2 negative breast cancer subtype sample. Pathway map was adapted from KEGG pathway database (17). Abbreviations are explained in the List of Abbreviations.

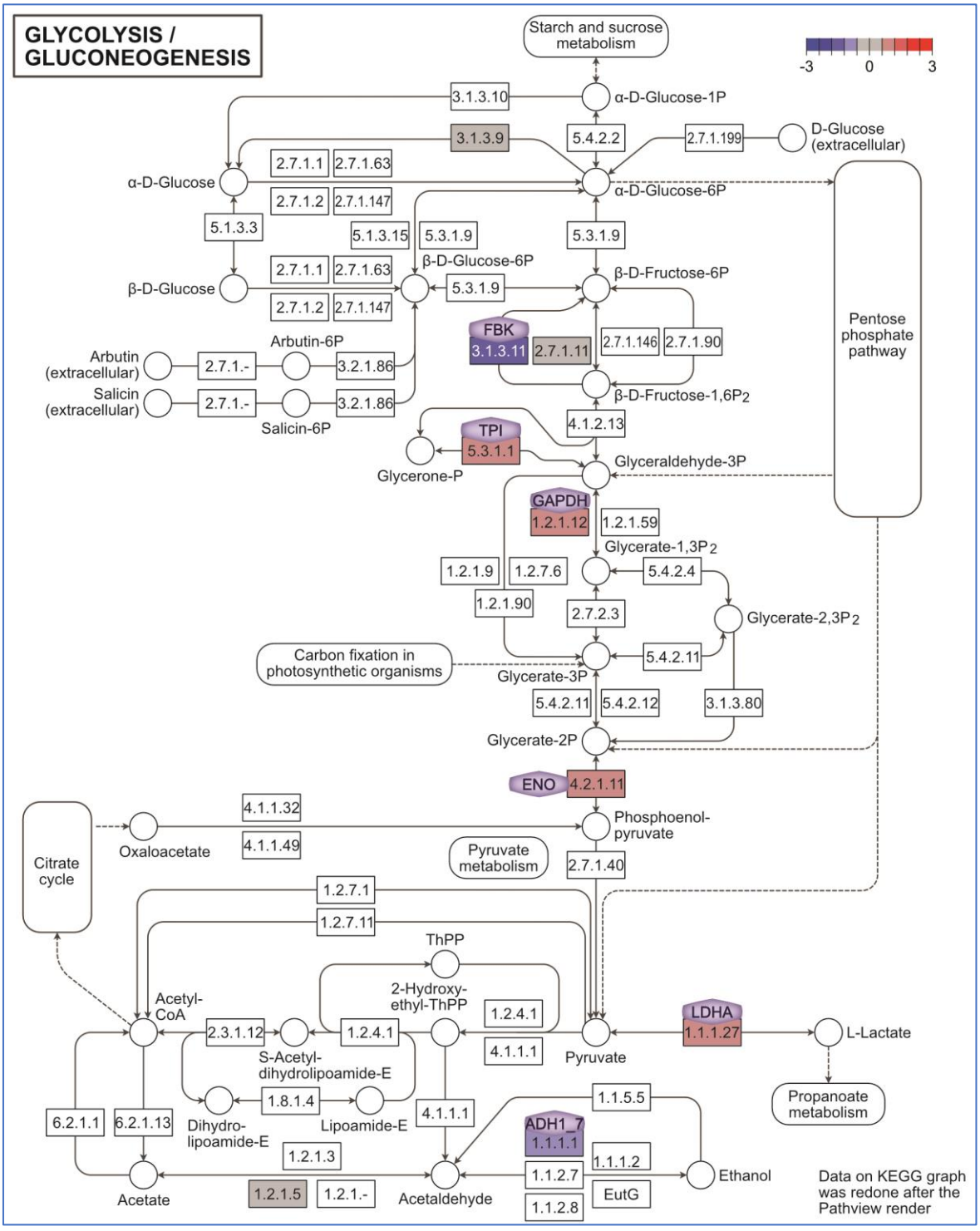

Figure 58: Altered gene expression encoding enzymes involved in glycolysis in basal HER2 negative breast cancer subtype sample. Pathway map was adapted from KEGG pathway database (17). Abbreviations are explained in the List of Abbreviations.

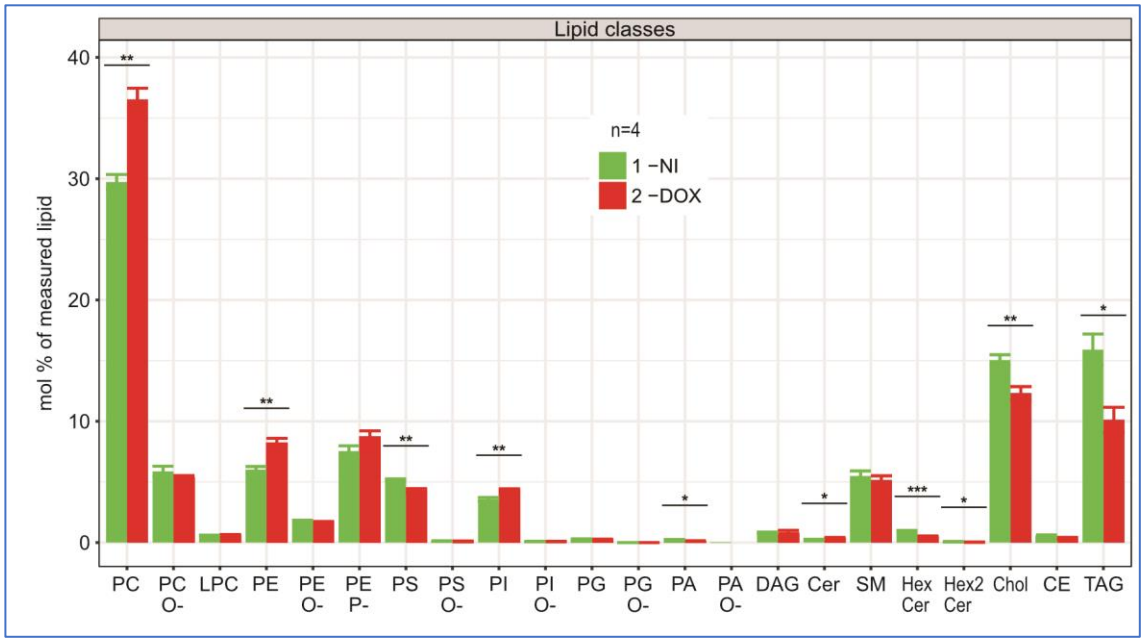

Figure 59: Changes in lipid composition in measured lipid classes between tumor (DOX) and normal (NI) populations. Error bars represent Mean ± 1SE. Abbreviations are explained in the List of Abbreviations.

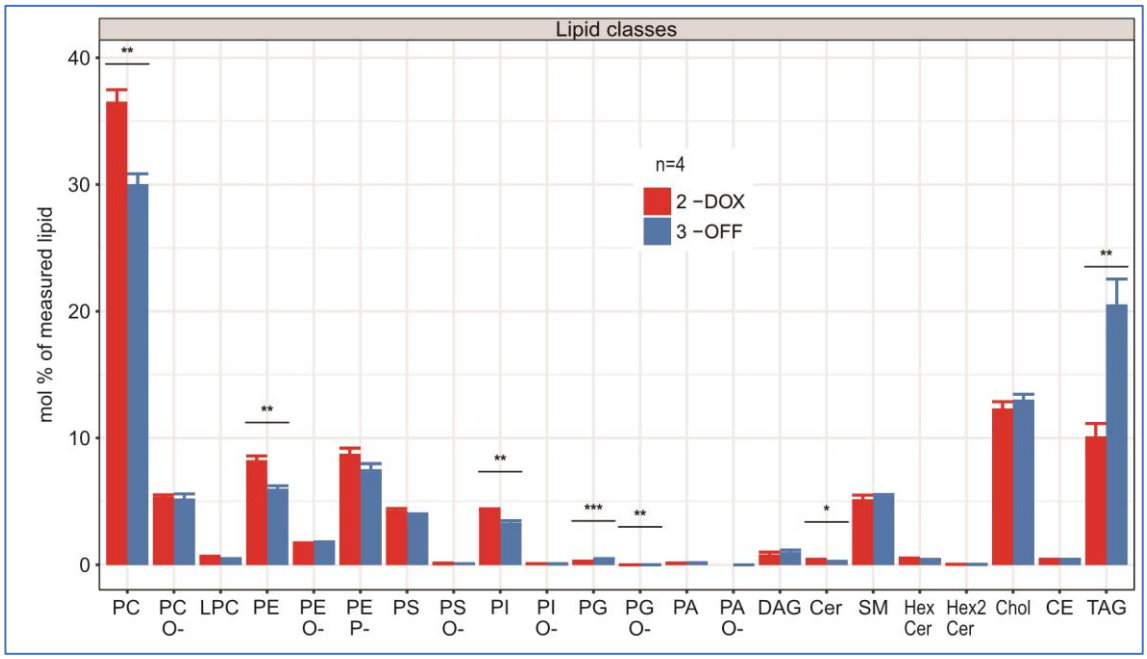

Figure 60: Changes in lipid composition in measured lipid classes between tumor (DOX) and residual (OFF) populations. Error bars represent Mean ± 1SE. Abbreviations are explained in the List of Abbreviations.

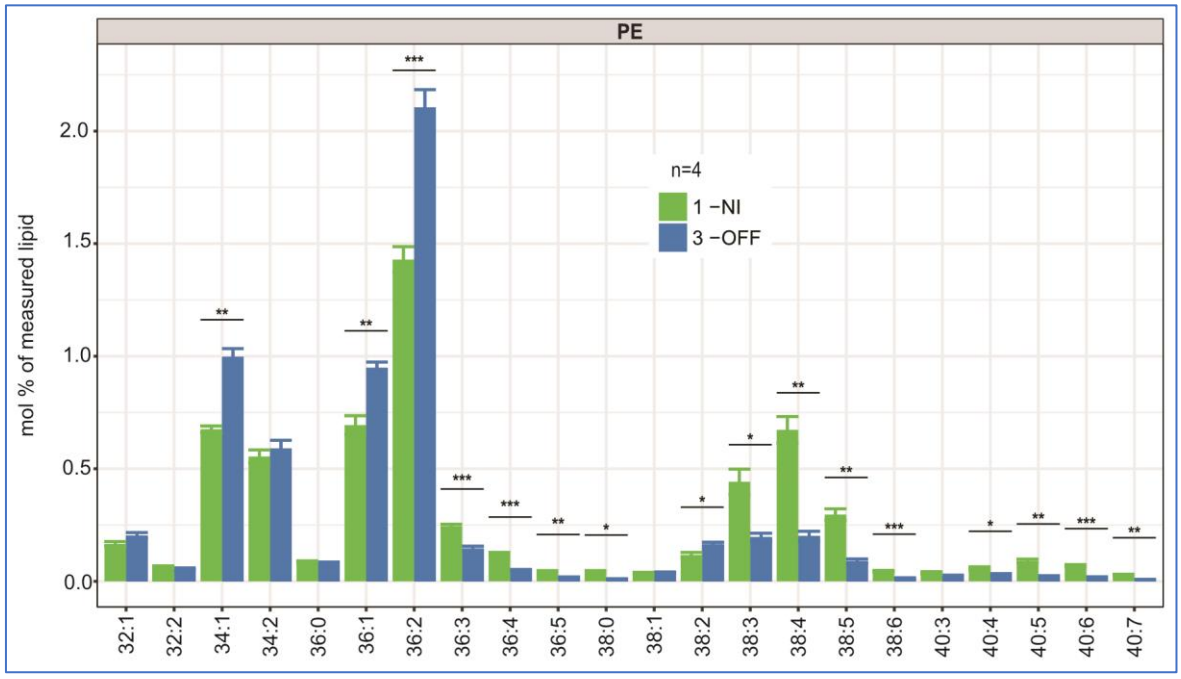

Figure 61: Changes in the lipid species within the class of PE between residual and normal cells. Error bars represent Mean ± 1SE. Abbreviations are explained in the List of Abbreviations.

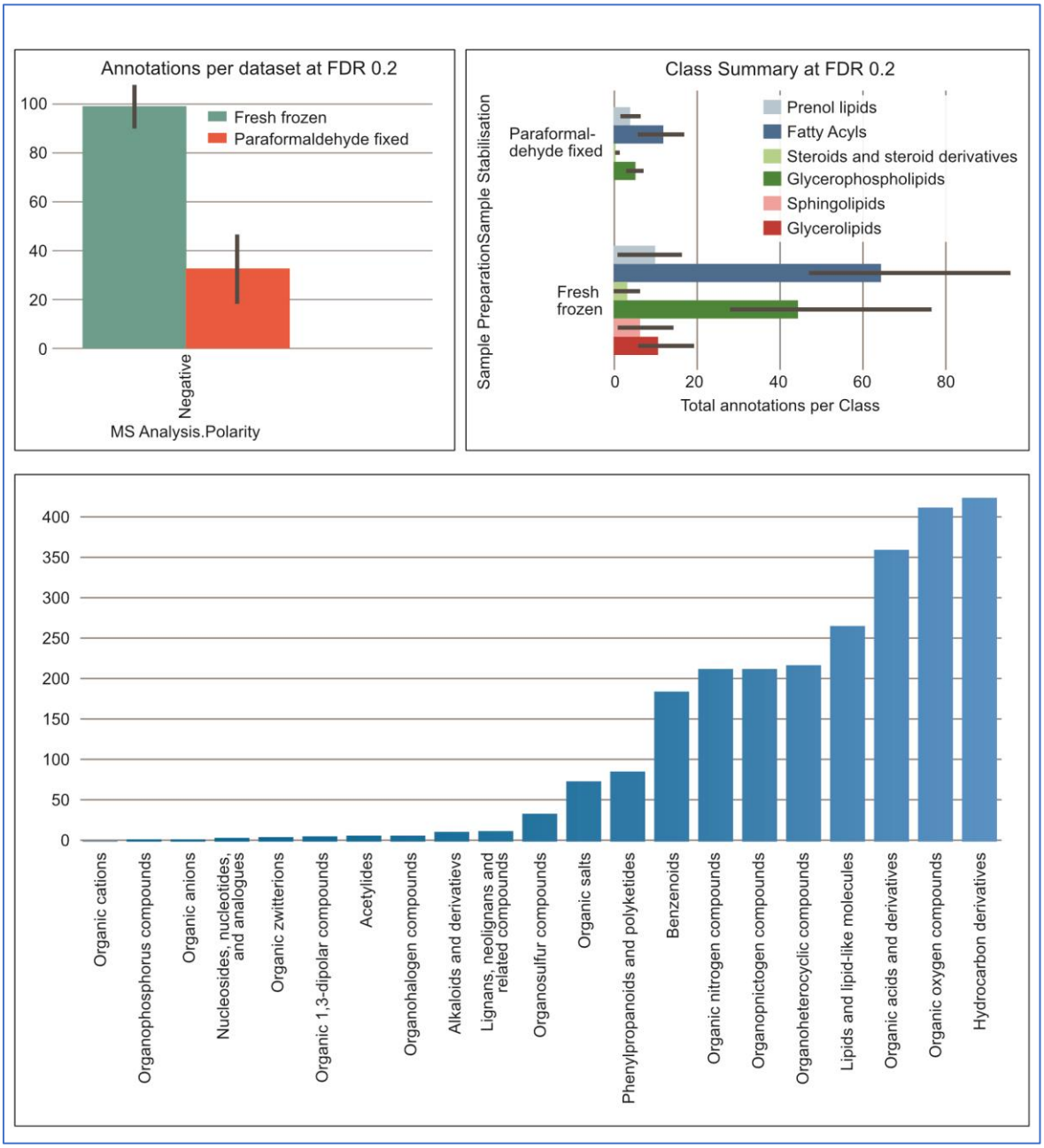

Figure 62: Metabolites detected by IMS. A) Comparison in number of annotations between the freshly frozen and PFA-treated samples, at FDR 0,2; B) Total number of annotations per class at FDR 0,2; C) Representation of all detected classes

Chapter 8

REFERENCES

REFERENCES

1. Hine RS, Martin E, editors. A Dictionary of Biology. Fifth. Oxford: Oxford University Press; 2004.

2. World Health Organization. Cancer - Key Statistics [Internet]. 2018 [cited 2018 Aug 28]. Available from: http://www.who.int/cancer/resources/keyfacts/en/

3. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136(5):E359–86.

4. Cancer.Net Editorial Board. What Is Personalized Cancer Medicine? [Internet]. Cancer.net Articles by American Society of Clinical Oncology (ASCO). 2012 [cited 2018 Nov 17]. Available from: https://www.cancer.net/navigating-cancer-care/how-cancer-treated/personalized-and-targeted-therapies/what-personalized-cancer-medicine

5. Jemal A, Ward E, Thun M. Declining death rates reflect progress against cancer. PLoS One. 2010;5(3):1–10.

6. Susan G. Komen Organization. Breast Anatomy - The structure of the breast [Internet]. 2018 [cited 2018 Sep 3]. Available from: https://ww5.komen.org/BreastCancer/TheBreast.html

7. Lynch PJ. Breast Anatomy Normal [Internet]. 2010 [cited 2018 Oct 25]. Available from: https://commons.wikimedia.org/wiki/File:Breast_anatomy_normal.jpg

8. NIH - National Cancer Institute; Terese Winslow. NCI Dictionary of Cancer Terms [Internet]. 2012 [cited 2018 Sep 3]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/lobular-carcinoma-in-situ

9. National Breast Cancer Foundation I. Breast Anatomy [Internet] [Internet]. 2016 [cited 2018 Aug 28]. Available from: https://www.nationalbreastcancer.org/breast-anatomy

10. Gudjonsson T, Adriance MC, Sternlicht MD, Petersen OW, Bissell MJ. Myoepithelial Cells: Their Origin and Function in Breast Morphogenesis and Neoplasia. J Mammary Gland Biol Neoplasia. 2005;10(3):261–72.

11. Tanner K. Regulation of the basement membrane by epithelia generated forces. Phys Biol. 2013;9(6):1–14.

12. John Hopkins Medicine Pathology. Types of Breast Cancer [Internet]. 2018 [cited 2018 Sep 3]. Available from: https://pathology.jhu.edu/breast/my-results/types-of-breast-cancer/

13. World Health Organization. Cancer - Breast Cancer [Internet]. 2018 [cited 2018 Aug 28]. Available from: http://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/

14. Xiao Y, Ma D, Ruan M, Zhao S, Liu XY, Jiang YZ, et al. Mixed invasive ductal and lobular carcinoma has distinct clinical features and predicts worse prognosis when stratified by estrogen receptor status. Sci Rep. 2017;7(1):1–9.

15. Hon JDC, Singh B, Sahin A, Du G, Wang J, Wang VY, et al. Breast cancer molecular subtypes: From TNBC to QNBC. Am J Cancer Res. 2016;6(9):1864–72.

16. Breast Cancer Org. Male Breast Cancer [Internet]. 2018 [cited 2018 Sep 28]. Available from: https://www.breastcancer.org/symptoms/types/male_bc

17. Weinstein IB, Joe A. Oncogene addiction. Cancer Res. 2008;68(9):3077–80.

18. Yersal O. Biological subtypes of breast cancer: Prognostic and therapeutic implications. World J Clin Oncol. 2014;5(3):412.

19. Li Z, Hu P, Tu J, Yu N. Luminal B breast cancer: patterns of recurrence and clinical outcome. Oncotarget. 2016;7(40).

20. Tran B, Bedard PL. Luminal-B breast cancer and novel therapeutic targets. Breast Cancer Res. 2011;13(6):221.

21. Liu H, Fan Q, Zhang Z, Li X, Yu H, Meng F. Basal-HER2 phenotype shows poorer survival than basal-like phenotype in hormone receptor-negative invasive breast cancers. Hum Pathol. 2008;39(2):167–74.

22. De Munck L, Van Maaren MC, Strobbe LJA, Smidt ML, Poortmans PMP, Siesling S. 10-year recurrence risks for molecular subtypes in breast cancer: A large population-based study in the Netherlands. Eur J Cancer. 2018 Apr 1;92:S130.

23. Nakai K, Hung MC, Yamaguchi H. A perspective on anti-EGFR therapies targeting triple-negative breast cancer. Am J Cancer Res. 2016;6(8):1609–23.

24. Masoud V, Pagès G. Targeted therapies in breast cancer: New challenges to fight against resistance. World J Clin Oncol. 2017;8(2):120.

25. Moasser MM. The oncogene HER2; its signaling ans transforming functions ant its role in human cancer pathogenesis. Oncogene. 2007;26(45):6469–87.

26. Mendes D, Alves C, Afonso N, Cardoso F, Passos-Coelho JL, Costa L, et al. The benefit of HER2-targeted therapies on overall survival of patients with metastatic HER2-positive breast cancer - a systematic review. Breast Cancer Res. 2015;17(1):1–14.

27. Gomide LB, Matheus JP C dos RF. Morbidity after breast cancer treatment and physiotherapeutic performance. Int J Clin Pr. 2007;61(6):972–82.

28. Moor JS De, Mariotto AB, Parry C, Alfano CM, Padgett L, Kent EE, et al. Cancer Survivors in the United States: Prevalence across the Survivorship Trajectory and Implication for Care. Cancer Epidemiol Biomarkers Prev. 2014;22(4):561–70.

29. BCRF (Breast Cancer Research Foundation). Triple Negative Breast Cancer: New Treatments [Internet]. 2018 [cited 2018 Nov 17]. Available from: https://www.bcrf.org/blog/triple-negative-breast-cancer-new-treatments

30. Teplinski E, Jhaveri K. Antibody-Drug Conjugates and T-DM1 [Internet]. Contemporary Oncology. 2014. Available from: https://www.onclive.com/publications/contemporary-oncology/2014/february-2014/antibody-drug-conjugates-and-t-dm1

31. Medscape. The Role of MRD in Hematologic Malignancies [Internet]. 2015 [cited 2018 Aug 12]. Available from: https://www.medscape.org/viewarticle/849751_2

32. Ommen HB. Monitoring minimal residual disease in acute myeloid leukaemia: A review of the current evolving strategies. Ther Adv Hematol. 2016;7(1):3–16.

33. Gerber B, Freund M, Reimer T. Recurrent Breast Cancer. Medicine (Baltimore). 2010;107(6):85–91.

34. Colleoni M, Sun Z, Price KN, Karlsson P, Forbes JF, Thürlimann B, et al. Annual hazard rates of recurrence for breast cancer during 24 years of follow-up: Results from the international breast cancer study group trials I to V. J Clin Oncol. 2016;34(9):927–35.

35. Freedman, Gary M.; Fowble BL. Local Recurrence After Mastectomy or Breast-Conserving Surgery and Radiation. Oncology. 2000;14(11):1561–81.

36. Comen E, Norton L. Minimal Residual Disease and Circulating Tumor Cells in Breast Cancer. 2012;195:13–24.

37. Tachtsidis A, McInnes LM, Jacobsen N, Thompson EW, Saunders CM. Minimal residual disease in breast cancer: an overview of circulating and disseminated tumour cells. Clin Exp Metastasis. 2016;33(6):521–50.

38. Butler TP, Gullino PM. Quantitation of Cell Shedding Into Efferent Blood of Mammary Adenocarcinoma. Cancer Res. 1975;35(3):512–6.

39. Xu X ZJ. Circulating Tumor Cells and Melanoma Progression. J Invest Dermatol. 2010;130(10):2349–51.

40. Gerratana L, Fanotto V, Bonotto M, Bolzonello S, Minisini AM, Fasola G, et al. Pattern of metastasis and outcome in patients with breast cancer. Clin Exp Metastasis. 2015 Feb;32(2):125—133.

41. Blatter S, Rottenberg S. Minimal residual disease in cancer therapy--Small things make all the difference. Drug Resist Updat. 2015;21–22:1–10.

42. Sun YL, Patel A, Kumar P, Chen ZS. Role of ABC transporters in cancer chemotherapy. Chin J Cancer. 2012;31(2):51–7.

43. Zahreddine H, Borden KLB. Mechanisms and insights into drug resistance in cancer. Front Pharmacol. 2013;4 MAR(March):1–8.

44. Mansoori B, Mohammadi A, Davudian S, Shirjang S, Baradaran B. The different mechanisms of cancer drug resistance: A brief review. Adv Pharm Bull. 2017;7(3):339–48.

45. Brown R, Curry E, Magnani L, Wilhelm-Benartzi CS, Borley J. Poised epigenetic states and acquired drug resistance in cancer. Nat Rev Cancer. 2014;14(11):747–53.

46. Zhang S, Mercado-Uribe I, Xing Z, Sun B, Kuang J, Liu J. Generation of cancer stem-like cells through the formation of polyploid giant cancer cells. Oncogene. 2014 Jan 2;33:134.

47. Viale A, Petazzoni P, Lyssiotis CA, Ying H, Sanchez N, Marchesini M, et al. Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function. Nature. 2015;514(7524):628–32.

48. Zhang YX, Yuan YQ, Zhang XQ, Huang DL, Wei YY, Yang JG. HMGB1-mediated autophagy confers resistance to gemcitabine in hormone-independent prostate cancer cells. Oncol Lett. 2017;14(5):6285–90.

49. Ebinger S, Özdemir EZ, Ziegenhain C, Tiedt S, Castro Alves C, Grunert M, et al. Characterization of Rare, Dormant, and Therapy-Resistant Cells in Acute Lymphoblastic Leukemia. Cancer Cell. 2016;30(6):849–62.

50. Straussman R, Morikawa T, Shee K, Barzily-Rokni M, Qian ZR, Du J, et al. Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. Nature. 2012;487(7408):500–4.

51. Shen X, Zhi Q, Wang Y, Li Z, Zhou J, Huang J. Hypoxia Induces Multidrug Resistance via Enhancement of Epidermal Growth Factor-Like Domain 7 Expression in Non-Small Lung Cancer Cells. Chemotherapy. 2017;62(3):172–80.

52. Milane L, Duan Z, Amiji M. Role of hypoxia and glycolysis in the development of multi-drug resistance in human tumor cells and the establishment of an orthotopic multi-drug resistant tumor model in nude mice using hypoxic pre-conditioning. Cancer Cell Int. 2011;11(1):3.

53. Yu H, Kortylewski M, Pardoll D. Crosstalk between cancer and immune cells: Role of STAT3 in the tumour microenvironment. Nat Rev Immunol. 2007;7(1):41–51.

54. Kim M, Oskarsson T, Acharyya S, Nguyen DX, Xiang H, Norton L, et al. Tumor self-seeding by circulating cancer cells. Cell. 2010;139(7):1315–26.

55. Friberg S, Nystrom A. Cancer Metastases: Early Dissemination and Late Recurrences. Cancer Growth Metastasis. 2015;8:CGM.S31244.

56. Paget S. THE DISTRIBUTION OF SECONDARY GROWTHS IN CANCER OF THE BREAST. Lancet. 2018 Feb 19;133(3421):571–3.

57. Hosseini H, Obradovic MMS, Hoffmann M, Harper KL, Sosa MS, Werner-Klein M, et al. Early dissemination seeds metastasis in breast cancer. Nature. 2016;540(7634):552–8.

58. Rhim AD, Mirek ET, Aiello NM, Maitra A, Jennifer M, Mccallister F, et al. NIH Public Access. 2013;148:349–61.

59. Pavlidis N, Khaled H, Gaafar R. A mini review on cancer of unknown primary site: A clinical puzzle for the oncologists. J Adv Res. 2015;6(3):375–82.

60. Price TT, Burness ML, Sivan A, Warner MJ, Cheng R, Lee CH, et al. Dormant breast cancer micrometastases reside in specific bone marrow niches that regulate their transit to and from bone. Sci Transl Med. 2016;8(340):1–12.

61. Blatter S, Rottenberg S. Minimal residual disease in cancer therapy – Small things make all the difference. Drug Resist Updat. 2015 Jul 1;21–22:1–10.

62. Shackleton M, Quintana E, Fearon ER, Morrison SJ. Heterogeneity in Cancer: Cancer Stem Cells versus Clonal Evolution. Cell. 2009.

63. Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012.

64. Markowetz F. A saltationist theory of cancer evolution. Nature Genetics. 2016.

65. Gao X, Zhang M, Tang Y, Liang X. Cancer cell dormancy: mechanisms and implications of cancer recurrence and metastasis. Onco Targets Ther. 2017;Volume 10:5219–28.

66. Aguirre-Ghiso JA. Models, mechanisms and clinical evidence for cancer dormancy. Nat Rev Cancer. 2007;7(11):834–46.

67. Giancotti FG. Mechanisms Governing Metastatic Dormancy and Reactivation. 2015;155(4):750–64.

68. Osisami M, Keller E. Mechanisms of Metastatic Tumor Dormancy. J Clin Med. 2013;2(3):136–50.

69. Kennaway EL. The Dormant Cancer Cell. Br Med J. 1954;2(4898):1225–30.

70. Kuilman T, Michaloglou C, Mooi WJ, Peeper DS. The essence of senescence. Genes Dev. 2010 Nov 15;24(22):2463–79.

71. Aguirre-Ghiso JA. Models, mechanisms and clinical evidence for cancer dormancy. Nature Reviews Cancer. 2007.

72. Aktas H, Cai H, Cooper GM. Ras Links Growth Factor Signaling to the Cell Cycle Machinery via Regulation of Cyclin D1 and the Cdk Inhibitor p27 KIP1. 1997;17(7):3850–7.

73. Oki T, Nishimura K, Kitaura J, Togami K, Maehara A, Izawa K, et al. A novel cell-cycle-indicator, mVenus-p27K -, identifies quiescent cells and visualizes G0-G1 transition. Sci Rep. 2014;

74. Moore N, Lyle S. Quiescent, slow-cycling stem cell populations in cancer: A review of the evidence and discussion of significance. Journal of Oncology. 2011.

75. Polo SE, Theocharis SE, Klijanienko J, Savignoni A, Asselain B, Vielh P, et al. Chromatin Assembly Factor-1, a Marker of Clinical Value to Distinguish Quiescent from Proliferating Cells. CANCER Res. 2004;64:2371–81.

76. Coleman R, Gray R, Powles T, Paterson A, Gnant M, Bergh J, et al. Adjuvant bisphosphonate treatment in early breast cancer: Meta-analyses of individual patient data from randomised trials. Lancet. 2015;386(10001):1353–61.

77. Prommer EE. Toxicity of Bisphosphonates. J Palliat Med. 2009 Nov 1;12(11):1061–5.

78. Ignatiadis M, Reinholz M. Minimal residual disease and circulating tumor cells in breast cancer. Breast Cancer Res. 2011;13(5).

79. Beaver JA, Jelovac D, Balukrishna S, Cochran R, Zabransky DJ, Wong HY, et al. Detection of Cancer DNA in Plasma of Early Stage Breast Cancer Patients. 2015;20(June 2013):2643–50.

80. Olmedillas-López S, García-Arranz M, García-Olmo D. Current and Emerging Applications of Droplet Digital PCR in Oncology. Mol Diagn Ther. 2017 Oct;21(5):493–510.

81. Rosen PR, Groshen S, Saigo PE, Kinne DW, Hellman S. A long-term follow-up study of survival in stage I (T1N0M0) and stage II (T1N1M0) breast carcinoma. J Clin Oncol. 1989;7(3):355–66.

82. Ahmad A. Pathways to Breast Cancer Recurrence. ISRN Oncol. 2013;2013:1–16.

83. Coller HA. Is cancer a metabolic disease? Am J Pathol. 2014;184(1):4–17.

84. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57–70.

85. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell. 2011;144(5):646–74.

86. Pavlova NN, Thompson CB. The Emerging Hallmarks of Cancer Metabolism. Cell Metab. 2016;23(1):27–47.

87. Lewis NE, Abdel-Haleem AM. The evolution of genome-scale models of cancer metabolism. Front Physiol. 2013;4 SEP(September):1–7.

88. Kroemer G, Pouyssegur J. Tumor Cell Metabolism: Cancer's Achilles' Heel. Cancer Cell. 2008;13(6):472–82.

89. Vander Heiden M, Cantley L, Thompson C. Understanding the Warburg effect: The metabolic Requiremetns of cell proliferation. Science (80-). 2009;324(5930):1029–33.

90. Kishton RJ, Rathmell JC. Novel therapeutic targets of tumor metabolism. Cancer J. 2015;21(2):62–9.

91. Phan LM, Yeung SJ, Lee M. Cancer metabolic reprogramming : importance , main features , and potentials for precise targeted anti-cancer therapies. Cancer Biol Med. 2014;11:1–19.

92. Ward PS, Thompson CB. Metabolic Reprogramming: A Cancer Hallmark Even Warburg Did Not Anticipate. Cancer Cell. 2012;21(3):297–308.

93. Ahn CS, Metallo CM. Mitochondria as biosynthetic factories for cancer proliferation. Cancer Metab. 2015;3(1):1–10.

94. Spinelli JB, Yoon H, Ringel AE, Jeanfavre S, Clish CB, Haigis MC. Metabolic recycling of ammonia via glutamate dehydrogenase supports breast cancer biomass. Science (80-). 2017;358(6365):941–6.

95. Sullivan LB. Metabolic Frugality Marks Cancer Cells for Immune Targeting. Cell. 2018;174(6):1344–6.

96. Sotgia F, Martinez-Outschoorn UE, Pavlides S, Howell A, Pestell RG, Lisanti MP. Understanding the Warburg effect and the prognostic value of stromal caveolin-1 as a marker of a lethal tumor microenvironment. Breast Cancer Res. 2011;13(4):213.

97. Von Ahrens D, Bhagat TD, Nagrath D, Maitra A, Verma A. The role of stromal cancer-associated fibroblasts in pancreatic cancer. Journal of Hematology and Oncology. 2017.

98. Endo H, Okuyama H, Ohue M, Inoue M. Dormancy of cancer cells with suppression of AKT activity contributes to survival in chronic hypoxia. PLoS One. 2014;

99. Bartosh TJ. Cancer cell cannibalism and the SASP: Ripples in the murky waters of tumor dormancy. Mol Cell Oncol. 2017;4(1):e1263715.

100. Ameri K, Jahangiri A, Rajah AM, Tormos K V., Nagarajan R, Pekmezci M, et al. HIGD1A Regulates Oxygen Consumption, ROS Production, and AMPK Activity during Glucose Deprivation to Modulate Cell Survival and Tumor Growth. Cell Rep. 2015;

101. Kao S-H, Wu K-J, Lee W-H. Hypoxia, Epithelial-Mesenchymal Transition, and TET-Mediated Epigenetic Changes. Brenner DA, Kisseleva T, Fuxe J, editors. J Clin Med. 2016 Feb 4;5(2):24.

102. Stine ZE, Walton ZE, Altman BJ, Hsieh AL, Dang C V. MYC, metabolism, and cancer. Cancer Discov. 2015;5(10):1024–39.

103. Vermeersch K, Styczynski M. Applications of metabolomics in cancer research. J Carcinog. 2013 Jan 1;12(1):9.

104. Hirschey MD, Deberardinis RJ, Diehl AME, Janice E, Frezza C, Green MF, et al. Dysregulated metabolism contributes to oncogenesis. Semin Cancer Biol. 2015;35(2015):129–50.

105. Beger R. A Review of Applications of Metabolomics in Cancer. Metabolites. 2013;3(3):552–74.

106. Yates LR, Knappskog S, Wedge D, Farmery JHR, Gonzalez S, Martincorena I, et al. Genomic Evolution of Breast Cancer Metastasis and Relapse. Cancer Cell. 2017;32(2):169–184.e7.

107. Demeulemeester J, Kumar P, Møller EK, Nord S, Wedge DC, Peterson A, et al. Tracing the origin of disseminated tumor cells in breast cancer using single-cell sequencing. Genome Biol. 2016;17(1):1–15.

108. Chatterjee A, Erban JK. Neoadjuvant therapy for treatment of breast cancer: the way forward, or simply a convenient option for patients? Gland Surg. 2017;6(1):119–24.

109. Kasimir-Bauer S, Bittner AK, König L, Reiter K, Keller T, Kimmig R, et al. Does primary neoadjuvant systemic therapy eradicate minimal residual disease? Analysis of disseminated and circulating tumor cells before and after therapy. Breast Cancer Res. 2016;

110. Scheel C, Weinberg RA. Cancer stem cells and epithelial–mesenchymal transition: Concepts and molecular links. Semin Cancer Biol. 2012 Oct 1;22(5–6):396–403.

111. Visvader JE, Lindeman GJ. Cancer stem cells in solid tumours: Accumulating evidence and unresolved questions. Nature Reviews Cancer. 2008.

112. Karsten U, Goletz S. What makes cancer stem cell markers different? Springer Plus. 2013;2(301):1–8.

113. Deshmukh A, Deshpande K, Arfuso F, Newsholme P, Dharmarajan A. Cancer stem cell metabolism: A potential target for cancer therapy. Molecular Cancer. 2016.

114. Kleffel S, Schatton T. Tumor Dormancy and Cancer Stem Cells: Two Sides of the Same Coin? In: Enderling H, Almog N, Hlatky L, editors. Systems Biology of Tumor Dormancy. New York, NY: Springer New York; 2013. p. 145–79.

115. Fang JY, Tan SJ, Wu YC, Yang Z, Hoang BX, Han B. From competency to dormancy: A 3D model to study cancer cells and drug responsiveness. J Transl Med. 2016;14(1):1–13.

116. Li S, Kennedy M, Payne S, Kennedy K, Seewaldt VL, Pizzo S V., et al. Model of tumor dormancy/recurrence after short-term chemotherapy. PLoS One. 2014;9(5):1–8.

117. Preciado JA, Reátegui E, Azarin SM, Lou E, Aksan A. Immobilization platform to induce quiescence in dormancy-capable cancer cells. TECHNOLOGY. 2017 Sep 1;05(03):129–38.

118. Jechlinger M, Podsypanina K, Varmus H. Regulation of transgenes in three-dimensional cultures of primary mouse mammary cells demonstrates oncogene dependence and identifies cells that survive deinduction. Genes Dev. 2009;23(14):1677–88.

119. Havas KM, Milchevskaya V, Radic K, Alladin A, Kafkia E, Garcia M, et al. Metabolic shifts in residual breast cancer drive tumor recurrence. J Clin Invest. 2017;127(6):2091–105.

120. D'Cruz CM, Gunther EJ, Boxer RB, Hartman JL, Sintasath L, Moody SE, et al. c-MYC induces mammary tumorigenesis by means of a preferred pathway involving spontaneous Kras2 mutations. Nat Med. 2001;7(2):235–9.

121. Moody SE, Sarkisian CJ, Hahn KT, Gunther EJ, Pickup S, Dugan KD, et al. Conditional activation of Neu in the mammary epithelium of transgenic mice results in reversible pulmonary metastasis. Cancer Cell. 2002;2(6):451–61.

122. Miller DM, Thomas SD, Islam A, Muench D, Sedoris K. c-Myc and cancer metabolism. Clin Cancer Res. 2012;18(20):5546–53.

123. Iqbal N, Iqbal N, Iqbal N, Iqbal N. Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications. Mol Biol Int. 2014;2014:1–9.

124. Nair R, Roden DL, Teo WS, McFarland A, Junankar S, Ye S, et al. C-Myc and Her2 cooperate to drive a stem-like phenotype with poor prognosis in breast cancer. Oncogene. 2014;33(30):3992–4002.

125. Chatzispyrou IA, Held NM, Mouchiroud L, Auwerx J, Houtkooper RH, Diseases GM, et al. Tetracycline antibiotics impair mitochondrial function and its experimental use confounds research. Cancer Res. 2016;75(21):4446–9.

126. Xia Lab @McGill. MetaboAnalyst - statistical, functional and integrative analysis of metabolomics data [Internet]. Available from: http://www.metaboanalyst.ca/

127. Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, Nielsen J. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nat Commun. 2014;5(May 2013):1–11.

128. Väremo L, Nielsen J, Nookaew I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. Nucleic Acids Res. 2013;41(8):4378–91.

129. Oatley M, Bolukbasi OV, Svensson V, Shvartsman M, Ganter K, Zirngibl K, et al. Single-cell transcriptomics identifies CD44 as a new marker and regulator of haematopoietic stem cells development. bioRxiv. 2018;338178.

130. Mrabet Y. Glycolysis - Wikimedia commons [Internet]. 2009 [cited 2018 Oct 3]. Available from: https://commons.wikimedia.org/wiki/File:Glycolysis.svg

131. Ah Mew N, Simpson KL, Gropman AL, Lanpher BC, Chapman KA, Summar ML. Urea Cycle Disorders Overview [Internet]. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® [Internet] Seattle (WA): University of Washington, Seattle. [cited 2018 Oct 3]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1217/

132. Gonzalez-angulo AM, Iwamoto T, Liu S, Chen H, Do K, Hortobagyi GN, et al. Gene expression, molecular class changes and pathway analysis after neoadjuvant systemic therapy for breast cancer. Clin Cancer Res. 2013;18(4):1109–19.

133. Maire V, Némati F, Richardson M, Vincent-Salomon A, Tesson B, Rigaill G, et al. Polo-like kinase 1: A potential therapeutic option in combination with conventional chemotherapy for the management of patients with triple-negative breast cancer. Cancer Res. 2013;73(2):813–23.

134. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45(D1):D353–61.

135. Palmer A, Phapale P, Chernyavsky I, Lavigne R, Fay D, Tarasov A, et al. FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry. Nat Methods. 2016;14(1):57–60.

136. Ravi M, Paramesh V, Kaviya SR, Anuradha E, Solomon FDP. 3D Cell Culture Systems: Advantages and Applications. J Cell Physiol. 2014 Jun 9;230(1):16–26.

137. Vidi P-A, Bissell MJ, Lelievre SA. Three-Dimensional Culture of Human Breast Epithelial Cells: The How and the Why. Methods Mol Biol. 2013;945(1):193–219.

138. Chatterjee SJ, McCaffrey L. Emerging role of cell polarity proteins in breast cancer progression and metastasis. Breast Cancer Targets Ther. 2014;6(0):15–27.

139. Mayers JR, Vander Heiden MG. Famine versus feast: understanding the metabolism of tumors in vivo. Trends Biochem Sci. 2015;40(3):130–40.

140. Yeh ES, Vernon-Grey A, Martin H, Chodosh LA. Tetracycline-regulated mouse models of cancer. Cold Spring Harb Protoc. 2014;2014(10):1005–22.

141. Allred DC, Medina D. The relevance of mouse models to understanding the development and progression of human breast cancer. J Mammary Gland Biol Neoplasia. 2008 Sep;13(3):279—288.

142. Podsypanina K, Politi K, Beverly LJ, Varmus HE. Oncogene cooperation in tumor maintenance and tumor recurrence in mouse mammary tumors induced by Myc and mutant Kras. Proc Natl Acad Sci. 2008;105(13):5242–7.

143. Pfefferle AD, Herschkowitz JI, Usary J, Harrell J, Spike BT, Adams JR, et al. Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. Genome Biol. 2013;14(11):R125.

144. Hollern DP, Andrechek ER. A genomic analysis of mouse models of breast cancer reveals molecular features of mouse models and relationships to human breast cancer. Breast Cancer Res. 2014;16(3):1–16.

145. Carroll JS, Hickey TE, Tarulli GA, Williams M, Tilley WD. Deciphering the divergent roles of progestogens in breast cancer. Nat Rev Cancer. 2017;17(1):54–64.

146. Dontu G, Ince TA. Of Mice and Women: A Comparative Tissue Biology Perspective of Breast Stem Cells and Differentiation. J Mammary Gland Biol Neoplasia. 2015;20(1–2):51–62.

147. Lim E, Wu D, Pal B, Bouras T, Asselin-Labat M-L, Vaillant F, et al. Transcriptome analyses of mouse and human mammary cell subpopulations reveals multiple conserved genes and pathways. Breast Cancer Res BCR. 2010;12(2):R21.

148. Rotroff DM, Motsinger-Reif AA. Embracing Integrative Multiomics Approaches. Int J Genomics. 2016;2016.

149. Lu M, Zhan X. The crucial role of multiomic approach in cancer research and clinically relevant outcomes. EPMA J. 2018;9(1):77–102.

150. Cavill R, Jennen D, Kleinjans J, Briedé JJ. Transcriptomic and metabolomic data integration. Brief Bioinform. 2016;17(5):891–901.

151. Matesic DF, Ali A, Sidorova TS, Burns TJ. A Cell-Cell Communication Marker for Identifying Targeted Tumor Therapies. Curr Bioact Compd. 2013;9(3):255–62.

152. Civelek M, Lusis AJ, Genetics M, Angeles L. A New View into the Regulation of Purine Metabolism - The Purinosome. 2014;15(1):34–48.

153. Li H, Li B, Yang F, Duan C, Bai Y, Yang JJ, et al. Purine Biosynthesis in Drug Resistance and Tumor Relapse of Childhood ALL. Blood. 2015 Dec 3;126(23):2627 LP-2627.

154. Phang JM, Liu W, Hancock CN, Fischer JW. Proline metabolism and cancer: Emerging links to glutamine and collagen. Curr Opin Clin Nutr Metab Care. 2015;18(1):71–7.

155. Zhao T, Mu X, You Q. Succinate: An initiator in tumorigenesis and progression. Oncotarget. 2017;8(32):53819–28.

156. Pastò A, Bellio C, Pilotto G, Ciminale V, Silic-Benussi M, Guzzo G, et al. Cancer stem cells from epithelial ovarian cancer patients privilege oxidative phosphorylation, and resist glucose deprivation. Oncotarget. 2014;5(12).

157. Zhang C, Liu Z, Liu X, Wei L, Liu Y, Yu J, et al. Targeted metabolic analysis of nucleotides and identification of biomarkers associated with cancer in cultured cell models. Acta Pharm Sin B. 2013;3(4):254–62.

158. Peiris-Pagès M, Martinez-Outschoorn UE, Pestell RG, Sotgia F, Lisanti MP. Cancer stem cell metabolism. Breast Cancer Res. 2016;18(1):1–10.

159. Farnie G, Sotgia F, Lisanti MP. High mitochondrial mass identifies a sub-population of stem-like cancer cells that are chemo-resistant. Oncotarget. 2015;6(31):30472–86.

160. Sancho P, Barneda D, Heeschen C. Hallmarks of cancer stem cell metabolism. Br J Cancer. 2016;114(12):1305–12.

161. Song I-S, Jeong YJ, Han J. Mitochondrial metabolism in cancer stem cells: a therapeutic target for colon cancer. BMB Rep. 2015;48(10):539–40.

162. Viale A, Pettazzoni P, Lyssiotis CA, Ying H, Sánchez N, Marchesini M, et al. Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function. Nature. 2014 Oct 30;514(7524):628–32.

163. Celià-Terrassa T, Kang Y. Distinctive properties of metastasis- initiating cells.

164. Ciavardelli D, Rossi C, Barcaroli D, Volpe S, Consalvo A, Zucchelli M, et al. Breast cancer stem cells rely on fermentative glycolysis and are sensitive to 2-deoxyglucose treatment. Cell Death Dis. 2014;

165. Schieber MS, Chandel NS. ROS links glucose metabolism to breast cancer stem cell and emt phenotype. Cancer Cell. 2013.

166. Liao J, Qian F, Tchabo N, Mhawech-Fauceglia P, Beck A, Qian Z, et al. Ovarian cancer spheroid cells with stem cell-like properties contribute to tumor generation, metastasis and chemotherapy resistance through hypoxia-resistant metabolism. PLoS One. 2014;

167. Zhou Y, Zhou Y, Shingu T, Feng L, Chen Z, Ogasawara M, et al. Metabolic alterations in highly tumorigenic glioblastoma cells: Preference for hypoxia and high dependency on glycolysis. J Biol Chem. 2011;

168. Shiraishi T, Verdone JE, Huang J, Kahlert UD, Hernandez JR, Torga G, et al. Glycolysis is the primary bioenergetic pathway for cell motility and cytoskeletal remodeling in human prostate and breast cancer cells. Oncotarget. 2015;6(1):130–43.

169. Jones W, Bianchi K. Aerobic glycolysis: Beyond proliferation. Front Immunol. 2015;6(MAY):1–5.

170. Keshet R, Szlosarek P, Carracedo A, Erez A. Rewiring urea cycle metabolism in cancer to support anabolism. Nat Rev Cancer. 2018;18(10):634–45.

171. Heinecke JL, Ridnour LA, Cheng RYS, Switzer CH, Lizardo MM, Khanna C, et al. Tumor microenvironment-based feed-forward regulation of NOS2 in breast cancer progression. Proc Natl Acad Sci. 2014;111(17):6323–8.

172. Cappelletti V, Iorio E, Miodini P, Silvestri M, Dugo M, Daidone MG. Metabolic Footprints and Molecular Subtypes in Breast Cancer. Dis Markers. 2017;2017.

173. Zhang; L, Xu; L, Zhang; F, Vlashi; E. Doxycycline inhibits the cancer stem cell phenotype and epithelial-to-mesenchymal transition in breast cancer. Cell Cycle. 2017;16(8):737–45.

174. De Francesco EM, Bonuccelli G, Maggiolini M, Sotgia F, Lisanti MP. Vitamin C and Doxycycline: A synthetic lethal combination therapy targeting metabolic flexibility in cancer stem cells (CSCs). Oncotarget. 2017;

175. Ganapathy-Kanniappan S, Geschwind J-FH. Tumor glycolysis as a target for cancer therapy: progress and prospects. Mol Cancer. 2013 Dec 3;12:152.

176. Aschner PJ, Ruiz AJ. Metabolic Memory for Vascular Disease in Diabetes. Diabetes Technol Ther. 2012 May 31;14(S1):S-68-S-74.

177. Chen Z, Miao F, Paterson AD, Lachin JM, Zhang L, Schones DE, et al. Epigenomic profiling reveals an association between persistence of DNA methylation and metabolic memory in the DCCT/EDIC type 1 diabetes cohort. Proc Natl Acad Sci. 2016;113(21):E3002–11.

178. McLean AC, Valenzuela N, Fai S, Bennett SAL. Performing Vaginal Lavage, Crystal Violet Staining, and Vaginal Cytological Evaluation for Mouse Estrous Cycle Staging Identification. J Vis Exp. 2012;(67):4–9.

179. Andrews S. FastQC: a quality control tool for high throughput sequence data [Internet]. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

180. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17(1):10–2.

181. Lo C-C, Chain PSG. Rapid evaluation and quality control of next generation sequencing data with FaQCs. BMC Bioinformatics. 2014;15(336):1–8.

182. NCBI. Genome Reference Consortium Mouse Build 38 patch release 4 (GRCm38.p4) [Internet]. Available from: https://www.ncbi.nlm.nih.gov/assembly/GCF_000001635.24/

183. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2 : accurate alignment of transcriptomes in the presence of insertions , deletions and gene fusions. 2013;14(R36):1–13.

184. Anders S, Pyl PT, Huber W. Genome analysis HTSeq — a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166–9.

185. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(550):1–21.

186. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11(R106):1–12.

187. Strimmer K. fdrtool : a versatile R package for estimating local and tail area-based false discovery rates. Bioinformatics. 2008;24(12):1461–2.

188. Strimmer K. A unified approach to false discovery rate estimation. BMC Bioinformatics. 2008;9.

189. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. 4A-NCBI GEO: Archive for functional genomics data sets - Update. Nucleic Acids Res. 2013;41(D1):991–5.

190. Chrysanthopoulos PK, Goudar CT, Klapa MI. Metabolomics for high-resolution monitoring of the cellular physiological state in cell culture engineering. Metab Eng. 2010;12(3):212–22.

191. Vernardis SI, Goudar CT, Klapa MI. Metabolic profiling reveals that time related physiological changes in mammalian cell perfusion cultures are bioreactor scale independent. Metab Eng. 2013;19:1–9.

192. Kanani HH, Klapa MI. Data correction strategy for metabolomics analysis using gas chromatography–mass spectrometry. Metab Eng. 2007;9(1):39–51.

193. Fuhrer T, Heer D, Begemann B, Zamboni N. High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry. Anal Chem. 2011;83(18):7074–80.

194. Kanani H, Chrysanthopoulos PK, Klapa MI. Standardizing GC–MS metabolomics. J Chromatogr B. 2008;871(2):191–201.

195. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.

196. Bligh E., Dyer W. Canadian Journal of Biochemistry and Physiology. J Biochem Physiol. 1959;37(8):911–7.

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

αKG	alpha-ketoglutarate
2-DG	2-deoxyglucose
2HG	2-hydroxyglutarate
2PG	2-phosphoglyceric acid
3-BP	3-bromopyruvate
3PG	3-phosphoglyceric acid
6-P-gluconolactone	6-phosphate-gluconolactone
ABC transporters	ATP-binding cassette transporters
ABCC1	ATP-binding cassette subfamily C member 1 gene
ACC	acetyl-CoA carboxylase
ACLY	ATP-citrate lyase
ADH	alcohol dehydrogenase
AI	aromatase inhibitors
AIAP	(2S)-(+)-Amino-5-iodoacetamidopentanoic acid
Ak4	adenylate kinase 4
AKT	RAC-alpha serine/threonine-*protein* kinase
ALDOA	aldolase A
ALL	acute lymphoblastic leukemia
AMP	adenosine monophosphate
Ang2	angiopoietin
ARG1	arginase 1
ASCT2	solute carrier family 1 member 5
ASL	arginosuccinate lyase
ASS1	arginosuccinate synthase 1
ATP	adenosine triphosphate
Bak	Bcl-2 homologous antagonist/killer
Bax	BCL2 associated X, apoptosis regulator
Bcl-2	B-cell lymphoma 2
BGP	beta-galactosidase-like protein

LIST OF ABBREVIATIONS

Bim	Bcl-2-like protein 11
BRAF	v-raf murine sarcoma viral oncogene homolog B1
BRCA1	breast cancer 1, early onset
CAF-1	chromatin assembly factor 1
CAFs	cancer associated fibroblasts
CAV1	caveolin-1
CDH1	E-cadherin
CE	cholesteryl ester
Cer	ceramide
Chol	cholesterol
CPS1	carbamoyl phosphate synthetase
CPT1	carnitine palmitoyltransferase 1
CSC-like	cancer stem cell-like
CTCs	circulating tumor cells
CUP	cancer of unknown primary
DAG	diacylglycerol
DAPI	4’,6-diamidino-2-phenylindole
DCA	dichloroacetate
DCIS	ductal carcinoma in situ
ddPCR	droplet digital PCR
DHAP	dihydroacetone phosphate acyltransferase, isoform A
DHFR	dihydrofolate reductase gene
DM1	emtansine
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
Dox	doxycycline
DTCs	disseminated tumor cells
ECM	extracellular matrix
EGFL7	epidermal growth factor-like domain 7
EGFR	epidermal growth factor receptor

LIST OF ABBREVIATIONS

EMT	epithelial-mesenchymal transition
ENO	enolase
ER	estrogen receptor
ERK	extracellular signal–regulated kinase
ETC	electron-transport chain
F6P	fructose 6-phosphate
FADH	reduced flavin adenine dinucleotide
FADH2	reduced flavin adenine dinucleotide, hydroquinone form
FAO	fatty acid oxidation
FASN	fatty acid synthase
FBP	fructose-1,6-bisphosphate
FDG-PET	2-[^{18}F] fluoro-2-deoxy-D-glucose positron-emission tomography
FDR	false discovery rate
FFPE	formalin-fixed paraffin-embedded
FH	fumarate hydratase
G6P	glucose 6-phosphate
G6PD	glucose-6-phosphate dehydrogenase
G6PGLD	6-phosphogluconate dehydrogenase
GABA	gamma-aminobutyric acid
GAP	glyceraldehyde 3-phopshate
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GCK	glucokinase
GCMS	gas chromatography-mass spectrometry
GDH	glutamate dehydrogenase
TIC	tumor initiating cells
GER 11	aminoguanidine hydrochloride
GL	glycerolipids
GLB1	galactosidase beta 1
GLS	glutaminase

LIST OF ABBREVIATIONS

Glu	glutamine
Glucose-6-P	glucose-6-phosphate
GLUD	glutamate dehydrogenase
GLUL	glutamate-ammonia ligase
GLUT	glucose transporter
GM-130	Golgi matrix protein
GO	Gene Ontology
GOT	aspartate aminotransferase, cytoplasmic
GP	glycerophospholipids
HBV	hepatitis B virus
HCV	hepatitis C virus
HER2	human epidermal growth factor receptor 2
Hex2Cer	dihexosylceramide
HexCer	hexosylceramide
Hexose-P	hexose-phosphate
HGF	hepatocyte growth factor
HIF1	hypoxia-inducible factor
HIPCs	hormone-independent prostate cancer
HK2	hexokinase 2
HMR2	Human Metabolic Atlas 2
HPV	human papillomavirus
IDC	invasive ductal carcinoma
IDH1	isocitrate dehydrogenase 1
IDH2	isocitrate dehydrogenase 1
IF	immunofluorescence
IGF1R	the insulin-like growth factor 1 receptor
IHC	immunohistochemistry
IMS	imaging mass spectrometry
iNOS	inducible nitric oxide synthase
ITGA6	alpha-6 integrin

LIST OF ABBREVIATIONS

JHDMS	Jumonji-C domain histone demethylases
KEGG	Kyoto Encyclopedia of Genes and Genomes
Ki67	marker of proliferation Ki67
LAT1	large neutral amino acids transporter small subunit 1
LCIS	lobular carcinoma in situ
LCMS	liquid chromatography-mass spectrometry
LDH	lactate dehydrogenase
LDHA	lactate dehydrogenase A
LIN28A	lin28 homolog A
LPC	lysophosphatidylcholine
mAb	monoclonal antibody
MCT	monocarboxylate transporte
MCT4	monocarboxylate transporter 4
Me	methionine
ME1	malic enzyme 1
MEK	mitogen-activated protein kinase kinase
miRNAs	microRNAs
MMTV-LTR	mouse mammary tumor virus-long terminal repeat
MRD	minimal residual disease
MRI	magnetic resonance imaging
MRSI	magnetic resonance spectroscopy imaging
MS	mass spectrometry
MT-CO1	mitochondrial cytochrome c oxidase subunit 1
mTOR	*mammalian target of rapamycin*
MYC	V-myc avian myelocytomatosis viral oncogene homolog
NAD+	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
NADP+	nicotinamide adenine dinucleotide phosphate
NADPH	nicotinamide adenine dinucleotide phosphate
NAGS	N-acetylglutamate synthase

LIST OF ABBREVIATIONS

nano-ESI MS	nano-electrospray ionization mass spectrometry
NO	nitric oxide
NOS2	nitric oxide synthase 2
NSCLC	non-small-cell lung carcinoma
Nt5e	5'-nucleotidase
ORA	Over Representation Analysis
ORNT1	ornithine transporter, mitochondrial
OTC	ornithine transcarbamylase
OXPHOS	oxidative phosphorylation
p16	p16 protein; cyclin-dependent kinase inhibitor 2A
p21	p21 protein; cyclin-dependent kinase inhibitor 1
p27	p27 protein; cyclin-dependent kinase inhibitor 1B
p53	p53 protein
PA	phosphatidic acid
PanIN	pancreatic intraepithelial neoplasia
PARP	poly ADP ribose polymerase
PC	phosphatidylcholine
PC	principal component
PCA	Principal Component Analysis
PCR	polymerase chain reaction
PDAC	pancreatic ductal adenocarcinoma
PDH (PD)	pyruvate dehydrogenase
PDHK (PDK)	pyruvate dehydrogenase kinase
PEP	phopshoenolpyruvic acid
PFA	paraformaldehyde
PFK	phosphofructokinase
PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3
PG	phosphatidylglycerol
PGAM1	phosphoglycerate mutase 1
PGI	phosphoglucose isomerase

LIST OF ABBREVIATIONS

PGK	phosphoglycerate kinase 1
PGM	phosphoglycerate mutase
PI	phosphatidylinositol
PI3K	phosphoinositide 3-kinase
PI3KCA	phosphatidylinositol 3-kinase
piano	Platform for Integrative Analysis of Omics Data
PK	pyruvate kinase
PKM2	pyruvate kinase muscle isozyme M2
PPP	pentose phosphate pathway
PR	progesteron receptor
PS	phosphatidylserine
ptDNA	plasma tumor DNA
qPCR	quantitative polymerase chain reaction
RAF	rapidly accelerated fibrosarcoma
RAS	rat sarcoma
Ribose-5-P	ribose-5-phosphate
RNA	ribonucleic acid
ROS	reactive oxygen species
rtTA	reverse tetracycline transactivator
SAH	S-adenosyl homocysteine
SAM	S-adenosylmethionine
SCD	stearoyl-CoA desaturase-1
SCO2	supercritical carbon dioxide cytochrome c oxidase assembly protein
SDH	succinate dehydrogenase
SERM	selective estrogen receptor modulators
SHD	succinate dehydrogenase
SIRT1	sirtuin 1
SLC1A5	solute carrier family 1 member 5
Slc25A13	citrin
SM	sphingomyelin

LIST OF ABBREVIATIONS

SN2	system N transporter 2
SP	sphingolipids
ST	sterol lipids
TAG	triacylglycerol
TCA	tricarboxylic acid
TDLU	terminal ductal lobular unit
T-DM1	trastuzumab emtansine
TET	Ten-eleven translocation
TKI	tyrosine kinase inhibitor
TME	tumor microenvironment
TNBC	triple negative breast cancer
TPI	triosephosphate isomerase
UCP2	uncoupling protein 2
UCP3	uncoupling protein 3
VDAC	voltage-dependent anion channel protein
VEGF	v*ascular endothelial growth factor*
WHO	World Health Organization
ZO-1	Zonula Occludens 1