Dissertation

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

presented by
Eleni Kafkia, MSc
Born in Patras, Greece
Oral examination: June 25th, 2019

Uncovering functional metabolic pathways using metabolomics: case studies of mammalian nucleus and dormant cancer cells

Referees
Dr. Judith Zaugg
Prof. Dr. Rüdiger Hell

To my parents

“As you set out for Ithaka
hope your road is a long one,
full of adventure, full of discovery.
Laistrygonians, Cyclops,
angry Poseidon—don’t be afraid of them:
you’ll never find things like that on your way
as long as you keep your thoughts raised high,
as long as a rare excitement
stirs your spirit and your body.”
Constantine P. Cavafy

Acknowledgements

I would like to start by thanking and expressing my sincere gratitude to my supervisor, Dr. Kiran Patil, for giving me the opportunity to be part of his group and to work on very exciting projects. His almost endless ideas were always a source of excitement that shaped my scientific thinking for the better. I would also like to thank him for being supportive, giving me freedom and for the nice stories he always has to tell.

I would also like to thank my thesis advisory committee members: Dr. Judith Zaugg, Prof. Dr. Rüdiger Hell, Dr. Martin Beck and Dr. Martin Jechlinger for their time investment, valuable inputs and advises, and for always interesting TAC meetings. I would also like to thank Prof. Dr. Michael Knop for accepting to be a member in the defense committee.

During the last four years I had the opportunity to meet many interesting people that evolved into being something more than lab-mates.

I would like to thank Filipa for being my friend, for her scientific mind and inputs, and for always trying to help me any way she could. I promised her a life-long supply of loukoumades that I had to postpone for a while. May the time begin again!

Although a relatively new addition to the Patil lab, I have known Christoniki for many years now, and I couldn't feel more grateful for the fact that she has always being there for me, to help and support, and to teach me (almost) how to send appropriate emoticons in what's app. Words are not enough.

Of course I couldn't forget the other fellow Greek, Dimitris, who stoically suffers (or maybe not?) most of the things in the lab, and he only breaks yeast-pinatas when the right time comes. I thank him for his willingness to always listen, help, and transfer big boxes from the stores.

I would like to especially thank Paula, close to her I learned a lot, she is a great teacher and always provided valuable help.

Many thanks go to Sonja and Eva, especially because I tortured them with the German translation of this thesis' abstract. I thank Sonja for always being willing to help. I thank Eva for her help, open-mindedness, and for all the good times we had (and we will have).

I would like to thank Natalia for always trying to keep the spirit up (in an almost alien way to me), and Justyna for her willingness to help and more importantly for her cynical self that gives moments of joy in the lab. I would like to thank Kristina, aka "the kid", for helping me with the last nuclei experiments, her constant will to help, and for bringing an air of coolness in the lab.

I would like to thank all other members of Patil lab, current and former, certainly a unique mixture of personalities that never makes life boring. Special thanks go to Gisela for bringing my not-so-nice drawings into life.

I would like to especially thank Ksenija Radić Shechter for a great collaboration, and for always being willing to help me with the thesis preparation and experiments. I would also like to thank the former members of Dr. Martin Beck’s lab, Amparo and Marie, for helping me with experiments, protocols and providing me reagents whenever needed.

Lastly, I would like to thank my family, my mam and dad, for their love, constant support and respect in every decision I have taken so far. My thesis is dedicated to both of them. I would like to thank my brother for being my brother, and Dione for being Dione. Finally, I would like to thank Thanos, for his patience, support and understanding, and for feeding me through critical periods in life.

Summary

Beyond its fundamental role in fulfilling the nutritional and energetic needs of the cell, metabolism has emerged as an important component of cellular regulatory processes which are central to diverse biological phenomena, ranging from cell differentiation to cancer and longevity. These metabolic pleiotropic roles often converge on the crosstalk with gene expression regulation which is sensitive to the availability of specific metabolites utilized for chromatin and RNA chemical modifications. These required metabolites are often assumed to freely diffuse into the nuclear space, with their key biosynthetic pathways mainly assigned to function elsewhere. However, considering that the intracellular environment is a rather viscous space where the free diffusion between the different compartments could be restricted, a significant question arises of how the nucleus ensures a reliable supply of these essential metabolites, especially as it is often a reaction-diffusion scenario and not only diffusion. Along these lines, the aim of the current PhD thesis was to explore the hypothesis that the nucleus could harbor extended metabolic networks, and not only individual enzymatic steps, for local production of nuclear-relevant metabolites. To examine this, firstly, nuclear proteomics data and nuclear localization signal motif analysis were utilized to assess the potentiality of a nuclear presence of the corresponding metabolic enzymes. Next, by employing stable isotope [U-^{13}C]-based metabolomics analysis in isolated nuclei, we tracked an operational activity. Proximity ligation mass spectrometry for selected enzymatic players allowed us to examine their proximity interactome further corroborating a nuclear subcellular topology. Cumulatively, our data provided multi-level evidence for a functional metabolic pathway operating in a mammalian nucleus. The identified pathway is made of parts of the TCA cycle with intermediates having key roles in chromatin and RNA modifications, reflecting thus the presence of a metabolic nuclear niche ensuring a stable supply of essential metabolites with nucleus-relevant functionalities.

The aforementioned crosstalk between metabolism and gene expression regulation highlights the importance of considering metabolic deregulations in pathophysiological conditions. Cancer metabolic alterations are a well-studied phenomenon. Yet, little is known for the metabolic physiology of residual cancer cells that survive treatment and contribute to cancer relapse. The current PhD thesis contributed to the characterization of the metabolic particularities of residual cancer cells derived from a mouse model of breast cancer. The analysis indicated that the residual cells, although phenotypically similar to their normal counterparts and despite the absence of oncogenes expression, preserved a tumorous metabolic memory with main characteristics of an enhanced glycolysis, deregulated TCA and urea cycle. Considering glycolysis' central role, we next aimed at investigating the network-wide metabolic responses upon inhibition of two important facilitators of the pathway, namely lactate dehydrogenase A and the monocarboxylate transporters 1 and 2, involved in lactate generation and transportation, respectively, in cancer cell lines. The results revealed opposite changes in metabolite concentration pools in glycolysis and TCA cycle intermediates between the two inhibitors treatment, and an overall lower biosynthetic flux. Interesting metabolic nodes were identified that could potentially be therapeutically exploited.

Uncovering and understanding metabolic network activities in previously overlooked places, like the existence of a nuclear multistep metabolic network, or the perseverance in cancer regressed cells of a metabolic phenotype mnemonic to the tumorous state, can shed light on the hitherto unknown mechanisms of gene regulation and its interplay with the metabolic state of a cell.

Zusammenfassung

Neben seiner grundlegenden Rolle in der Versorgung der Zelle mit Nährstoffen und Energie, hat der Stoffwechsel sich als wichtiger Bestandteil regulatorischer Prozesse herausgestellt, welche eine zentrale Rolle in biologischen Vorgängen einnehmen, die von der Zelldifferenzierung bis hin zu Krebs und Langlebigkeit reichen. Diese metabolischen pleiotropen Rollen laufen oft beim Crosstalk mit der Genexpressionsregulierung zusammen, die auf die Verfügbarkeit metabolischer Intermediate reagiert, welche für die chemische Modifikation von Chromatin und RNA vonnöten sind. Obwohl bestimmte metabolische Enzyme im Zellkern gefunden wurden, die an der Biosynthese der benötigten Metaboliten beteiligt sind, werden die Hauptstoffwechselwege nach wie vor anderen Bereichen der Zelle zugeschrieben. Es wird daher angenommen, dass die erforderlichen Metaboliten frei in den Kern diffundieren. Diese Annahme führt jedoch zu erheblichen Einschränkungen bei der Effizienz der Modifikationen, da die erhöhte Viskosität der intrazellulären Umgebung die Diffusion sogar kleiner Moleküle einschränken könnte, benötigte Metaboliten von anderen Enzymen abgefangen werden könnten und besonders da chemische Modifikationen oft Reaktions-Diffusionsprozesse sind. In diesem Zusammenhang war es das Ziel der vorliegenden Doktorarbeit, die Hypothese zu untersuchen, dass der Zellkern ausgedehnte metabolische Netzwerke und nicht nur einzelne enzymatische Schritte für die lokale Produktion von kernrelevanten Metaboliten enthalten kann. Um dies zu untersuchen, wurden zunächst Zellkern-Proteomics-Daten und Kernlokalisierungssignalanalysen verwendet, um das Potenzial einer entsprechenden Präsenz von Stoffwechselenzymen im Kern zu bestimmen. Durch den Einsatz von stabilen Isotopen [U-^{13}C]-basierten Fluxanalysen in isolierten Zellkernen wurde die operative Aktivität verfolgt. Durch die Proximity-Ligations-Massenspektrometrie für ausgewählte enzymatische Akteure im Stoffwechsel konnten wir ihr Proximity-Interaktom untersuchen, was eine subzelluläre Zellkerntopologie weiter untermauert. Zusammengefasst liefern unsere Daten einen mehrstufigen Nachweis für einen funktionellen Stoffwechselweg in einem Säugetierzellkern. Der identifizierte Weg besteht aus Teilen des TCA-Zyklus mit Intermediaten, die eine Schlüsselrolle bei Chromatin- und RNA-Modifikationen spielen. Dies spiegelt das Vorhandensein einer metabolischen Zellkernnische wider, die eine stabile Versorgung mit essentiellen Metaboliten mit kernrelevanten Funktionalitäten gewährleistet.

Der zuvor erwähnte Crosstalk zwischen Metabolismus und Genexpressionsregulation unterstreicht die Bedeutung der Berücksichtigung von Stoffwechsel-Deregulierungen bei pathophysiologischen Zuständen. Krebstypische Stoffwechselveränderungen sind ein gut untersuchtes Phänomen. Über die Stoffwechselphysiologie von Krebszellen, die sich zurückentwickeln, die Behandlung überleben und zu Rückfällen beitragen ist jedoch wenig bekannt. Die vorliegende Doktorarbeit trug zur Charakterisierung der metabolischen Besonderheiten solcher residualen Krebszellen in einem Mausmodell für Brustkrebs bei. Die Analyse zeigte, dass die residualen Krebszellen trotz der phänotypischen Ähnlichkeit zu normalen Zellen und der Abwesenheit der Expression von Onkogenen ein tumoröses metabolisches Gedächtnis mit erhöhter Glykolyse und dereguliertem Harnstoff- und TCA-Zyklus bewahrten. In Anbetracht der zentralen Rolle der Glykolyse untersuchten wir in Krebszellen die netzwerkweiten metabolischen Auswirkungen der Inhibierung zweier zentraler Vermittler in der Glykolyse. Dabei handelte es sich um die Lactatdehydrogenase A und die Monocarboxylat-Transporter 1 und 2, die an der Bildung beziehungsweise dem Transport von Lactat beteiligt sind. Die Ergebnisse zeigten entgegengesetzte Änderungen in den Konzentrationen von Glykolyse und TCA-Zyklus-Intermediaten nach der Inhibitorbehandlung und einen allgemein verringerten Biosynthesefluss. In diesem Zusammenhang wurden interessante Metaboliten identifiziert, die therapeutisches Potential haben könnten.

Zusammengenommen haben die Entdeckung und das Verständnis von Stoffwechselaktivitäten in versteckten und bisher übersehenen Bereichen, wie die Existenz einer mehrstufigen Stoffwechselarchitektur im Zellkern oder das Fortbestehen des an Krebszellen erinnernde Stoffwechselzustand in eigentlich zurückentwickelnden Krebszellen, weitreichende Implikationen für die bisher unbekannten Mechanismen der Genregulation und ihr Zusammenspiel mit dem Stoffwechselzustand der Zelle.

Table of Contents

Introduction 19
  Metabolism and gene expression regulation 19
    Metabolites and chromatin modifications 19
    Nuclear topology of epigenetic-related metabolic enzymes 22
  Metabolism and cancer 25
Thesis Motivation and Objectives 31
Results and Discussion 33
  Chapter I: Uncovering metabolic pathways operating in the mammalian nucleus 35
    Nuclear Proteome and NLS motifs suggest a nuclear subnetwork made of TCA cycle enzymes 35
    [U-^{13}C]-Glutamine and [U-^{13}C]-Citrate label TCA cycle key intermediates delineating an active metabolic pathway operating in the HeLa nucleus 38
    Towards the identification of a nuclear niche using proximity biotinylation mass spectrometry 46
    Summary and Outlook 55
    Contributions 57
  Chapter II: Metabolic particularities of minimal residual disease in breast cancer 59
    Tumor and residual cells exhibit divergent response at the transcriptomic and metabolic level 59
    The metabolic alterations of the residual population are mnemonic of the tumorous state 63
    Ex vivo and *in vivo* validations support the altered glycolysis and urea cycle metabolic signatures in the residual population 70
    Discussion and Outlook 72
    Contributions 75
  Chapter III: Metabolic response of cancer cells to inhibition of lactate dehydrogenase and monocarboxylate transport activities 77
    Treatment with LDHA and MCT1 and 2 inhibitors results in distinctive metabolic profiles characterized by opposite alterations in glycolysis and TCA cycle 77
    Inhibition of MCT1 and 2 ablates pyruvate export but not lactate 80
    Inhibition of MCT1/2 or LDHA induces a decreased lipid biosynthesis flux 81
    Contribution of glutamine to TCA cycle is altered after treatment with LDHA or MCT1/2 inhibitors 83
    LDHA inhibition points to alterations in fructose metabolism 84
    Summary and Outlook 86
    Contributions 87
Concluding Remarks 91
Materials and Methods 93

Chapter I – Materials and Methods 95
  - Nuclei Purification 95
  - 13C-labeling experiments in isolated nuclei and cell lysates and metabolite extraction 95
  - Gas Chromatography - Mass Spectrometry (GC-MS) data acquisition and analysis 96
  - Generation of Constructs 97
  - Cell Culture and generation of stable inducible cell lines 98
  - BioID labeling and cell harvesting 99
  - Affinity purification of biotinylated proteins 100
  - Sample preparation and TMT labeling for mass spectrometry....101
  - Mass spectrometry data acquisition 101
  - Mass spectrometry data analysis 102
  - Statistical Analysis and Identification of proximal interactors 102
  - Gene ontology enrichment analysis 103
  - Dotplot generation 103
  - Cytoscape visualization 103
  - Immunofluorescence and Imaging 103
  - Immunoblotting 104

Chapter II and Chapter III – Materials and Methods 105
  - Three dimensional cell cultures 105
  - Three dimensional cell cultures experiments - Cell and Spent Growth Media Collection and Metabolite Extraction 106
  - *In vivo* mammary glands experiments 107
  - Gas Chromatography - Mass Spectrometry (GC-MS) data acquisition and analysis ..107
  - NOS enzymatic assay 108
  - Two dimensional cell cultures experiments - Cell and Spent Growth Media Collection and Metabolite Extraction 108

List of figures 111

List of Tables 113

Abbreviations 114

Publications 117

Appendix....119

Bibliography 125

Introduction

Introduction

Metabolism has often been viewed in a mono-perspective way, mainly reflecting the bioenergetic and nutritional requirements of the cell. However, over the past years a growing body of evidence points that metabolism is not simply a static bystander, rather an active component of many physiological processes. Being the most dynamic and environment sensitive level of cellular function, metabolism not only captures information that goes beyond the genome, it also mediates the environmental inputs into gene and protein regulatory responses. This is highlighted by the metabolic dependencies of epigenetic-related enzymes that utilize specific metabolites as substrates or co-factors for the chemical modifications on chromatin, as well as from the fact that cancer cells harbor particular metabolic characteristics that contribute to cancer growth, proliferation and evolution. The following introduction builds around these two aspects of metabolism presenting the current available knowledge and addresses open questions that initiated the research motivations of the current PhD thesis.

Metabolism and gene expression regulation

Metabolites and chromatin modifications

Amongst the different levels of organismal and intracellular interactions, the crosstalk between metabolism and gene regulation has been an increasingly established phenomenon. From the work of Jacob and Monod[1] in the 60s that led to the elucidation of lac operon, up to the characterization of more complex nutrient and hormonal sensing mechanisms and signaling cascades[2], these studies have highlighted how changes in the (micro)environmental nutritional status lead to changes in gene expression in order for an organism to adapt. More recent findings, however, have highlighted a supplemental and more intimate relationship between metabolism and gene regulation through the epigenome. It is nowadays well established that DNA and histones bare chemical modifications which modulate chromatin structure and accessibility, and in turn gene expression. These modifications, as well as their deposition and removal from the epigenetic modifying enzymes, require specific metabolites functioning as substrates or co-factors that are key intermediates of metabolism. Of the most well-known modifications, DNA and histone methylation depends on S-adenosyl methionine (SAM) as a methyl donor substrate, while the inverse reaction, the removal of methyl groups, is mediated by demethylases that make use of flavin adenine dinucleotide (FAD) or alpha-ketoglutarate as co-factors. Accordingly, acetyl-CoA is the universal acetyl donor for histone acetylation, while a certain class of deacetylases, the sirtuin family of proteins, are dependent on NAD^+ as a co-factor to exert their function. The list of the modification types and the relevant needed metabolites is not exhausted to the above mentioned. Additional marks have been described and are continuously discovered, including various modifications especially on histone core and tails, like succinylation, propionylation, malonylation, butyrylation, crotonylation, glutarylation, citrullination, *O*-GlcNAcylation, poly(ADP) ribosylation[3], or the most recently characterized histone serotonylation[4].

These metabolic dependencies of chromatin modifications connote that the metabolic physiology of a cell may, through the availability of particular metabolites, drive changes in chromatin, and in this way contribute to modulations in gene expression. Despite this being a rather logical connection, it remained largely underexplored, partially because it was not known whether the concentrations, or the range of variation, of these epigenetic-related metabolites could become limiting for the epigenetic-modifying enzymes. In other words, it was unclear whether the epigenetic modifying enzymes behave similar to the metabolic enzymes, whose activities are dependent on the metabolite abundances, or resemble mostly to kinases whose activities are independent of the levels of their substrate, the ATP, at least for most cases[5,6]. However, many rapid discoveries over the past decade started to provide evidence that variations on the metabolite concentration levels could indeed impact the respective epigenetic modifications.

It has been shown that acetyl-CoA abundance levels can influence both global histone acetylation levels, as well as histone acetylation at specific genes. Along these lines, alterations in glycolytic flux result in similar alterations in acetyl-CoA pools and subsequently in histone acetylation levels[7–9]. In multiple mammalian cell lines, the regulation of histone acetylation by glucose availability was found to be dependent on the activity of ATP-citrate lyase (ACLY)[7], the enzyme that produces acetyl-CoA from citrate. Under specific conditions, connected mostly to metabolic stresses like in hypoxic or nutrient-limiting environments, acetate can be utilized as an alternative carbon source for the production of acetyl-CoA and subsequently for histone acetylation through the activity of acetyl-coenzyme A synthase 2 (ACSS2)[10,11]. The same enzyme was identified as the predominant contributor of acetyl-CoA generation for histone acetylation in mouse hippocampal neuronal cells[12], highlighting that the relative pathways contributing to histone acetylation can be differentiated by cell type and physiology state. In addition, ACSS2 was found to be important for the recycling of acetate produced by histone deacetylases (HDACs)[11]. Apart from global acetylation, acetyl-CoA abundance levels have also been implicated to influences of histone acetylation of specific genes, for instance genes related to glycolysis, growth and proliferation[7,13], as well as neuronal genes[12].

Similarly to acetyl-CoA, fluctuations in S-adenosyl methionine (SAM) concentration can impact total methylation levels. Studies have shown that decrease in the concentration of

SAM as a result of depletion in metabolites directly contributing to its generation, like methionine[14], or threonine for the case of mouse embryonic stem cells[15], has been connected to reduced methylation levels and induction of stem cell differentiation. Stem cells pluripotency and differentiation has also been associated to demethylation dynamics through the action of another metabolite, that of alpha-ketoglutarate, which is a necessary co-factor for a specific family of demethylases. In mouse embryonic stem cells, increase in concentration levels of alpha-ketoglutarate promote demethylation contributing to pluripotency maintenance[16]. On the contrary, in human primed pluripotent stem cells, alpha-ketoglutarate supplementation leading to global histone and DNA demethylation has the opposite cell fate effect, supporting stem cell differentiation[17], italicizing thus how the same mechanism exerts a differential outcome depending on context-specificity. Around the same epigenetic node, the metabolites downstream of alpha-ketoglutarate, namely succinate and fumarate, has been shown to inhibit the activity of the alpha-ketoglutarate demethylases, and in this way contribute to global hypermethylation, a recurring phenotype in tumors harboring mutations in succinate dehydrogenase (SDH) and fumarate hydratase (FH) that lead to supra-physiological levels of those metabolites[18]. Similarly, D- and L-2-hydroxyglutarate, two metabolites produced by mutant versions of isocitrate dehydrogenases 1 and 2 (IDH1 and IDH2), and wild type lactate dehydrogenase (LDH) and malate dehydrogenase (MDH) promiscuous activities under acidic or hypoxic conditions, respectively, are also inhibitors of the alpha-ketoglutarate demethylases inducing similar hypermethylating effects[18,19].

Modifications on RNAs, both on coding and non-coding, require also specific metabolic intermediates and play key role in their stability and function representing thus multiple additional layers in gene expression regulation[20]. The most known RNA modifications involve methylations which require SAM as a methyl donor, however many others stem from more complex metabolic pools and combinations[21]. As such, potential threads between metabolism and RNA could also apply, in an analogy to those of histones and DNA. A recent work have highlighted in leukemia and glioma the involvement of D-2-hydroxyglutarate in the inhibition of the mRNA demethylase FTO, resulting in a global increase in the N^6-methyladenosine (m^6A) mRNA modifications, which in turn renders mRNA transcripts unstable and prone to degradation, especially those associated with MYC/CEBPA pathways[22]. Another example involves the N^4-acetylcytidine (ac^4C) modification found in rRNAs and tRNAs, which requires at least acetyl-CoA as a donor of the acetyl moiety. Studies in *Saccharomyces cerevisiae* and *Caenorhabditis elegans* found that perturbations contributing in decreased concentration levels of acetyl-CoA result also in ac^4C reduced levels[23,24]. The human RNA acetyltransferase responsible for the catalysis of this modification, namely NAT10[25], has been found to bind CoA suggesting that the enzyme could be regulated by feedback inhibition[26]. While studies in this field have been limited so far, the aforementioned examples begin to illuminate another level of complexity in gene regulation and yet another role for metabolism's contribution into this.

Nuclear topology of epigenetic-related metabolic enzymes

It is evidently from the above that the cell nucleus relies on specific metabolites to confer chemical changes on DNA, RNA and histones. A common perception for the needed metabolites is to freely diffuse into the nucleus, given especially the large size of the nuclear pores. However, often neglected is the fact that the intracellular environment is a rather viscous space where the free diffusion of metabolites between the different compartments, like cytoplasm and nucleus, could be restricted. A significant emerging aspect is thus how the efficiency of the epigenetic enzymatic reactions is ensured regarding the required metabolite accessibility and availability. While compartment-specific metabolic measurements are difficult, recent studies have highlighted the existence of some enzymes involved in the generation of epigenetic-related metabolites inside the nucleus.

Of the most prominent examples is the methionine adenosyltransferase II (MATII) which catalyzes the biosynthesis of SAM from methionine. In mammalian cells, MATII localizes in the nucleus where it associates with chromatin remodeling factors and locally supplies methyltransferases with SAM[27]. In yeast nucleus, the respective enzyme is part of a multi-enzymatic complex that additionally contains a pyruvate kinase, serine metabolic genes and an acetyl-CoA synthetase (SESAME complex, SErine-responsive SAM-containing Metabolic Enzyme complex). The complex interacts with methyltransferases providing them with SAM, and in this way combines the methylation and phosphorylation (through the activity of the pyruvate kinase) of histone H3 in specific genes in response to glucose availability through the glucose-derived serine[28].

Analogous to SAM, acetyl-CoA is locally synthesized in the nucleus through three distinct enzymatic players the contribution of which is mainly shaped from condition and/or tissue particularities. In mammalian cells, under nutrient and oxygen sufficient conditions, ATP-citrate lyase (ACLY), an enzyme that localizes both in the cytoplasm and in the nucleus, produces acetyl-CoA from citrate utilized for global, as well as gene-specific, histone acetylation as mentioned earlier[7]. Recently, ACLY has also been implicated to histone acetylation as a response to DNA damage induced by ionization. ACLY was found to synthesize acetyl-CoA used for histone H4 and H2A acetylation close to double-strand DNA breaks. Absence of ACLY, or prevention of ACLY entry into the nucleus, resulted in reduced histone acetylation and impairment of repair by homologous recombination[29]. Apart from ACLY, another enzyme involved in acetyl-CoA production in the nucleus is the pyruvate dehydrogenase complex (PDC), a complex traditionally encountered as of mitochondrial topology only. In cancer cells, PDC was found to translocate from mitochondria into the nucleus for local acetyl-CoA production from pyruvate, especially during S to G2 cell cycle phase transition and as a response to growth stimulating factors, mild mitochondrial stresses[30] or during a specific developmental time window in mice embryos[31]. Additional evidence, implicate PDC as being part of a larger multi-enzymatic complex in the nucleus, involving the pyruvate kinase M2, the acetyltransferase p300 and the arylhydrocarbon receptor (AhR). Pyruvate synthesis by PKM2 is further utilized by PDC for the generation of acetyl-CoA which in turn is used by p300 for the acetylation of specific genes that are targets of AhR[32]. Finally, the third enzyme connected to a nuclear acetyl-CoA production is ACSS2 that makes use of acetate as a substrate. As it was mentioned in the above section, ACSS2 nuclear activity has been connected to conditions of metabolic stress[10,11] or to specific tissues (i.e. neurons)[12], as well as to nuclear reprocessing of acetate produced by histone deacetylases towards acetyl-CoA local regeneration[11].

While the aforementioned examples refer to the activities of mostly singular enzymes in the nucleus, a recent study in mouse embryos characterized the existence of a functional, partial TCA cycle in the nucleus during zygotic genome activation[31]. The pathway covered all the enzymatic reactions from pyruvate up to alpha-ketoglutarate, combining thus the biosynthesis of two important epigenetic-related metabolites, that of acetyl-CoA and alpha-ketoglutarate, highlighting in this way the actuality, within the nuclear compartment, of metabolic channeling facilitating chromatin modifications. The pathway was found to be absent from the nucleus in later stages and to localize only in the mitochondria, in alignment with the traditional perspective regarding the subcellular topology of TCA cycle enzymes. Interestingly, aconitase 2 (ACO2), one of the enzymes of the aforementioned pathway that catalyzes the citrate to isocitrate formation, is present in yeast nuclei with uncharacterized function so far[33]. The subsequent to ACO2 enzymatic step involving the isocitrate conversion to alpha-ketoglutarate is facilitated by isocitrate dehydrogenase (IDH) which has three isoforms in mammalian cells. While their primary localization is assigned to mitochondria for IDH2 and IDH3, and cytoplasm for IDH1, both IDH1 and IDH3 appear to have a nuclear localization albeit in seemingly smaller pools and with yet unexplored function(s). Recently, another TCA cycle enzyme has been located in the nucleus of mammalian cell lines. The alpha-ketoglutarate dehydrogenase complex (OGDC), that generates succinyl-CoA from alpha-ketoglutarate, has been found to interact and supply the acetyltransferase KAT2A with succinyl-CoA. KAT2A in turn functions as a succinyltransferase mediating the succinylation of histones in gene promoter regions[34]. Finally, two additional TCA cycle-related enzymes that appear to have a nuclear localization are the succinate dehydrogenase (SDH) and fumarate hydratase (FH). SDH is a heterodimer consisting of SDHA and SDHB subunits mediating the

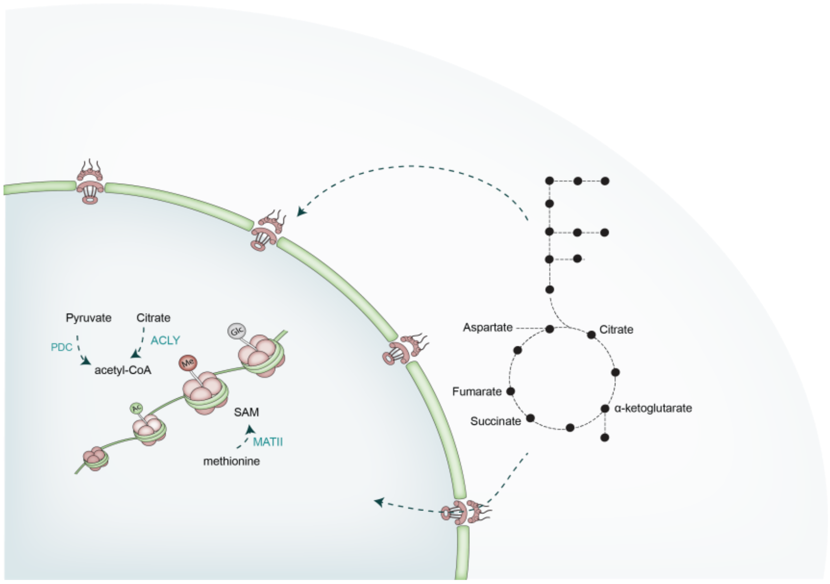

Figure 1: Paradigms of subcellular localization of metabolic enzymes related to the production of epigenetic relevant metabolites.

conversion of succinate to fumarate in the mitochondria where it is also coupled to oxidative phosphorylation. Studies from the 60s have spotted both subunits in the nucleoli of cancer cells exhibiting an enzymatic activity, although that was not directly assigned to fumarate production[35,36]. Since then, the nuclear presence and potential functions of the heterodimer have not been further investigated. Finally, fumarate hydratase (FH), which catalyzes the reversible fumarate to malate reaction, was of the first mitochondrial TCA cycle enzymes to be described in the nucleus where its enzymatic activity was connected to chromatin modifications. More specifically, it has been shown that FH translocate to the nucleus in response to DNA damage induced by ionization[37]. Later it was found that the enzyme interacts with histones close to double strand DNA breaks and locally produces fumarate that inhibits histone H3 demethylation, setting in turn a sequence of events for recruitment of specific factors leading to DNA repair by non-homologous end-joining[38].

Cumulatively, the aforementioned paradigms begin to highlight that although metabolic intermediates diffuse into the nucleus, their local biosynthesis has biologically meaningful implications. As a consequence of this, many interesting aspects begin to emerge, for instance, the inquiry of the existence of more extended metabolic networks around epigenetic-related metabolites within the nuclear compartment, apart from only distinctive enzymatic steps. As an extension to the nuclear partial TCA cycle identified in mouse embryos, it is enticing to hypothesize the existence of multi-step metabolic organization coopted in additional cells and/or conditions that could allow for more efficient metabolic channeling, and/or allow for multiple pathway couplings regarding metabolite biosynthesis with nuclear relevant extensions.

Metabolism and cancer

Although cancer metabolism was actively explored before the 1970s, the advent of genomic era with the molecular and genetic approaches has dampened the interest towards it. However, over the past decade cancer metabolism has been greatly reconsidered as it is now well established that cancer is not only the final outcome of genetic mutations, instead for its effective understanding non-genetic determinants must also be taken into account in a relationship that is not simply one of cause and effect. Cancer cells are characterized by metabolic particularities that offer them great robustness, enabling them to adapt and support their growth and proliferation even under limiting conditions by (re)shaping the pathway activities and their relative contributions[39]. These idiosyncrasies are mentored by both cell-internal factors, representing metabolic network constraints mainly imposed by cancer genetics and tissue of origin, as well as external factors, referring to the microenvironment as it is shaped by nutrient and oxygen availabilities and the interactions with matrix and neighbouring cells[40]. Although there are no universal metabolic features common to all cancer types, there are still some general hallmarks[39], whose interpretation in the context of the aforementioned internal and external to each cancer type factors could offer therapeutic opportunities.

Pleiotropic roles of aerobic glycolysis in cancer cells

Of these hallmarks, the most commonly recurring across many cancer types is an alteration in glucose metabolism[41]. In non-cancerous and non-proliferating mammalian cells, glucose is converted to pyruvate which upon entry in mitochondria is metabolized in TCA cycle and oxidized to CO_2 yielding high amounts of ATP through oxidative phosphorylation. Cancer cells, as well as normal proliferating and developing cells, display an increase in glucose uptake which in turn is converted to lactate, even under oxygen replete conditions. Subsequently, lactate is secreted in the extracellular space. This phenomenon, known as aerobic glycolysis or the Warburg effect, was first described by Otto Warburg almost a century ago[41,42]. Recent studies have highlighted that aerobic glycolysis could be induced both by oncoproteins, like MYC, RAS or AKT, as well as by environmental inputs, such as growth factors[40,43–46]. However, despite the magnitude of studies, how exactly aerobic glycolysis promotes or contributes to cancer growth still remains unclear. Studies argue that glycolysis functions mainly by supporting the high energetic and biosynthetic requirements of cancer cells[41]. Although it is now well established that ATP is not a limiting factor for rapidly proliferating cells, which instead present only a slight increase in ATP consumption[39], a major requirement is the reducing equivalents, more specifically NADPH[47]. The various branching pathways stemming from glycolysis can indeed couple the glycolytic function to the biosynthesis of NADPH, as well as to various other macromolecules like nucleotides, lipids and amino acids such as serine and glycine. The latter two represent an example of glycolytic carbon divergence in certain cancers, where up to 50% of glucose contributes to their production through the rate-limiting enzyme phosphoglycerate dehydrogenase (PHGDH)[48]. Serine and glycine are in turn connected to various pathways such as generation of glutathione, NADPH, as well as SAM for epigenetic modifications. However, despite the above paradigm, a limitation for the generalization of the biosynthesis as the driving force of the aerobic glycolysis in cancer cells comes from the fact that biomass accumulation would require the glucose carbons to be retained and not secreted as lactate[41], which in many cases accounts even for 93% of the glucose uptake[49]. These findings imply that biosynthesis could occur only in times of a very high glycolytic flux, and in turn lactate secretion helps to maintain the glycolysis active[43].

Additionally, the above numbers suggest that glycolysis may serve more functions, beyond bioenergetics and biosynthesis. Possible connections are now implicating lactate which has been shown to exert multifactorial effects. Lactate, through the acidification of the microenvironment, can promote disruption of the tissue architecture enabling thus tumour invasion[50]. In addition, lactate has been found to act as a signal for tumour-associated macrophages modulating their gene expression[51]. These altered macrophages can in turn induce tube-like morphogenesis in adjacent cells that allow blood perfusion in nutrient limited areas[52]. Recent studies have expanded even more the functional relevance of lactate, attributing to it a more centralized role, as it has been shown that is the primary carbon source for TCA cycle metabolites, both in normal tissues (in mice)[53], as well as in lung cancer cells[54]. The studies showed that lactate's turnover fluxes exceeded those of glucose both in fed and fasting mice, and that glucose contributed in TCA cycle through circulating lactate[53]. Lactate dehydrogenase that catalyses the reversible conversion of pyruvate to lactate exist as homo or heterotetramers of lactate dehydrogenase A and B (LDHA, LDHB) subunits. In HeLa cells, LDHB has a mitochondrial topology, in addition to the cytoplasmic, catalysing lactate oxidation to pyruvate and subsequent utilization in TCA cycle as well as in lipid biosynthesis[55]. These studies, further complement earlier models suggesting a symbiosis between glycolytic and oxidative tumour cells where lactate produced from one type is utilized by the other[56]. Cumulatively, lactate's pleiotropic roles render him a potential therapeutic target with studies examining the inhibition of the enzymes catalysing its biosynthesis (LDH) or its transportation (monocarboxylate transporters, MCTs). So far, none of the developed LDH inhibitors have progressed into clinical trials, owing to inefficacies, to developed drug resistance or unacceptable toxicities[57], while MCT inhibitors are currently being tested as an alternative[40].

Mitochondria and TCA cycle in cancer

Apart from aerobic glycolysis, the importance of mitochondrial metabolism to tumour proliferation is now beginning to emerge. It is well established that most cancer cells still maintain functional mitochondria whose relevance to cancer growth is highlighted by studies where depletion of mitochondrial DNA (mtDNA) reduces their tumorigenic potential[58]. Tumorigenesis is restored once cells acquire host mtDNA through a process that involve whole mitochondrial transferring from stroma cells[59]. Cells lacking mtDNA display a defective electron transport chain (ETC)[60]. Surprisingly, the feature that mainly contributes to the attenuation of their proliferation is not an ATP deficiency, as it was initially thought, rather the lack of electron acceptors from ETC whose presence is necessary for the continuation of enzymatic reactions involved mainly in nucleotide biosynthesis. More specifically, such reactions include the dihydroorotate dehydrogenase (DHODH)[61], a rate-limiting enzyme whose activity is coupled to respiration and is involved in *de novo* pyrimidine synthesis, as well as the oxidation reactions in TCA cycle leading to aspartate generation which is an essential carbon and nitrogen donor for purine and pyrimidine biosynthesis[62,63]. As a result, targeting of electron transport chain[64–66] and especially its combinatorial targeting with glycolysis[57] are examined as alternative potent strategies for cancer growth inhibition.

In addition to nucleotide biosynthesis, TCA cycle further contributes to other metabolic precursors which complement the ones generated from glycolysis. These precursors include non-essential amino acids as well as fatty acids. Fatty acids, in contrast to the majority of non-cancerous tissues, present a high increase in many cancer cells and apart from being integral components in membrane synthesis they can also function as signalling molecules[67]. In mammalian cells, the most predominant pathway for their biosynthesis involves the citrate which is produced inside the mitochondria through TCA cycle and subsequently is exported to cytosol for generation of oxaloacetate and acetyl-CoA from ATP citrate lyase (ACLY). Part of acetyl-CoA is converted by acetyl-CoA carboxylase (ACC) to malonyl-CoA, the latter being the first committed step to fatty acid generation[68]. Both ACLY and ACC, along with the fatty acid synthase (FASN), have been found upregulated in human cancers, while their inhibition reduces cancer cell growth *in vitro* and *in vivo*[67,69,70]. This carbon divergence away from TCA cycle for lipogenesis necessitates its replenishment through additional nutrients in order for the cycle to continue the supply of other anabolic and energetic intermediates. As such, apart from glucose, a prominent anaplerotic carbon source is glutamine which feeds the TCA cycle mainly after conversion to glutamate and subsequently to alpha-ketoglutarate[49]. Oxidation of alpha-ketoglutarate in the cycle contributes to the generation of aspartate and citrate amongst other metabolites. Under specific contexts, associated with hypoxic conditions or defective mitochondria, glutamine contributes to the generation of citrate mainly through a reversal of a part of the TCA cycle, where alpha-ketoglutarate is reductively carboxylated to isocitrate and then citrate, through the activities of isocitrate dehydrogenases (IDH1 or IDH2) and aconitase (ACO1 or ACO2), respectively[71–73]. Apart from a carbon donor, glutamine is also an important nitrogen source for nucleotide biosynthesis[49]. Most cancer cells in culture exhibit an increased dependence on glutamine anaplerosis[49], and recent studies have highlighted that this feature can be regulated by oncoproteins, more specifically MYC[74]. Along these lines, inhibitors targeting glutaminase, the enzyme which converts glutamine to glutamate, are explored as therapeutic cancer interventions[75].

The aforementioned paradigms signify how classical central carbon metabolic pathways such as glycolysis and TCA cycle are deregulated to contribute to cancer growth and proliferation through novel and formerly underappreciated biological mechanisms. As already mentioned, these metabolic particularities are often additionally shaped by the tissue metabolic architecture and the tumour microenvironment[76]. As a result, the same oncogene can induce differential metabolic phenotypes along the same pathway in different tissues, and different cancerous tissues can employ differential nutrients or strategies for their acquisition and biosynthesis. This is illustrated in myc induced lung and liver cancers, where the first display an upregulation in glutamine uptake and catabolism, while the latter are characterized by increased glutamine synthesis and accumulation[77]. Regarding nutrient acquisitions, kras transformed lung cells can directly uptake branched chain amino acids (BCAAs) from the blood stream[78], instead kras pancreatic tumours obtain them indirectly from extracellular proteins through macropinocytosis[79]. Lastly, in a manner dependent both on tumour microenvironment and tissue genetics, aspartate biosynthesis can derive either from glutamine through TCA cycle[62,63], or from glucose through a sequence of reactions with fewer oxidative steps[80]. All the above emphasize the perplexity in identifying metabolic targets for cancer therapy since identical cells in distinctive microenvironments may exhibit differential metabolic phenotypes and variations in responses, for instance in drugs. Furthermore, given that the tumour microenvironment substantially differs from the more tractable one in *in vitro* cell culture systems challenges exist regarding the *in vivo* relativity of metabolic reprograming schemes and extensive verifications are required. This was recently highlighted in kras lung tumours - cancer cells derived from these tumours extensively utilized glutamine when cultivated *in vitro*, however for the respective *in vivo* tumours glutamine utilization was minimal[80]. Cell and tissue morphology may also influence nutrient accessibilities, so cell culture systems better resembling *in vivo* tissue morphology, for instance 3D cell cultures[81], offer significant advantages when system controllability is required. Finally, although the substantial amount of studies over the last years have shaped many important aspects around cancer metabolism, much remains unknown for the metabolic idiosyncrasies of tumour dormant cells which contribute to cancer retrogression. Their characterization regarding persisting and uniquely metabolic attributes is imperative for further understanding cancer pathophysiology and therapeutics.

Thesis Motivation and Objectives

Every stage in the life of a human cell, from embryonic development to ageing, critically depends on precise regulation of gene expression. To achieve this highly complex task, the cell nucleus relies on specific metabolites utilized as substrates or cofactors to confer modifications on DNA, RNA and chromatin. Although certain metabolic enzymes related to the biosynthesis of these metabolites have been spotted in the nucleus, their key metabolic pathways, e.g. the tricarboxylic acid cycle, are mainly assigned to function in the cytoplasm or mitochondria. The metabolites required for these modifications are thus assumed to freely diffuse across the nuclear pores and within the nucleoplasmic space. However, considering that the intracellular environment is a rather viscous space where the free diffusion between the different compartments could be restricted, a significant question arises of how the nucleus ensures a reliable supply of these essential metabolites, especially as it is often a reaction-diffusion scenario and not only diffusion. Towards these lines, the principle purpose of the current PhD thesis was to explore the hypothesis that the nucleus could harbor extended metabolic networks, and not only distinctive enzymatic steps, around epigenetic-related metabolites (Chapter I). The existence of nuclear multistep metabolic architecture, in an analogy to known metabolons described so far for other subcellular organelles, can reveal novel extensions in gene expression regulation and the interplay with the cellular metabolic state.

The aforementioned metabolic dependencies of the epigenetic regulating enzymes suggest that the metabolic physiology of a cell may, through the availability of metabolites, drive changes in chromatin and in this way contribute to modulations in gene expression. As such, the metabolic alterations underlying many diseases, being either the outcome or the effect, may in turn have enabling, promoting, or reshaping activities on these disordered states. Along these lines, cancer metabolic particularities are now a well-established phenomenon presenting potential therapeutic interventions. Yet, far less is known for the metabolic physiology of residual, dormant cancer cells that contribute to cancer relapse. Understanding their metabolic pathophysiology is important as it may reveal mechanisms sustaining this population and signatures enabling their *in vivo* identification which in turn can lead in devised therapeutic strategies for their successful eradication. By employing a three-dimensional cell culture system derived from a mouse model that carry inducible oncogenes (i.e. MYC/NEU), and thus able of simulating tumor progression, regression and relapse, the current thesis contributes to the characterization of the metabolic reformations of the residual population (Chapter II). Lastly, building around a well-established phenomenon present across many cancer types, that of aerobic glycolysis, the effects of two potent drug inhibitors involved in lactate generation and transportation are comprehensively addressed at the metabolic level (Chapter III).

Results and Discussion

Chapter I: Uncovering metabolic pathways operating in the mammalian nucleus

The cell nucleus critically depends on metabolic factors for mitigating a precise regulation of gene expression[82]. Considering the large size of the nuclear pores, a common perception for these required metabolites is that they freely diffuse into the nuclear space, with their key metabolic pathways functioning elsewhere. However, often neglected is the fact that the intracellular environment is a rather viscous space where the free diffusion of metabolites between the different compartments could be restricted[83], or metabolites could be siphoned to other metabolic processes until they reach the nuclear sites of need. This diffusion scenario in turn poses questions for the efficiency of the epigenetic enzymatic reactions and how it is ensured regarding the required metabolite accessibilities and availabilities. Recent studies started to highlight the existence of specific metabolic enzymes involved in the biosynthesis of epigenetic-related metabolites inside the nucleus. The most prominent paradigms include the nuclear production of SAM[27], and acetyl-CoA which can be locally synthesized through three distinct enzymatic players the contribution of which is dependent on specific conditions and/or tissue particularities.[7,30]. These studies thus reveal that a local nuclear generation can have biologically meaningful implications. Along these lines, here, we opted to investigate the hypothesis that the nucleus harbors not only distinctive enzymatic steps, but extended metabolic networks which could increase the efficiency of substrate channeling conferring a more reliable supply of essential metabolites rate-limiting in the control of gene expression. We explored this by employing a combination of proteomics, [U-^{13}C]-based stable isotope metabolomics and proximity labeling mass spectrometry to draft the potentiality and activity of the respective metabolic pathways.

Nuclear Proteome and NLS motifs suggest a nuclear subnetwork made of TCA cycle enzymes

As a first step of this study, we opted to explore literature-available proteomics datasets of isolated nuclei for the presence of metabolic enzymes. The proteomics data corresponding to nuclei isolated from five commonly used mammalian cell lines were obtained from Beck and colleagues[84]. Data were also available for crude cytoplasmic fractions referring to cytoplasm and all other subcellular compartments except the nucleus. 2146 metabolic enzymes covering 78 pathways of central carbon metabolism were extracted from KEGG database[85–87] and used as a seeding dataset to search against the aforementioned nuclear data. In total, 181 enzymes (out of the 2146, ~8.5% coverage) were consistently detected in the nucleus of all five cell lines, while the respective number for the crude cytoplasmic fraction was 567 enzymes (out of 2146, 27% coverage) (Figure 2). When the nuclear detected metabolic enzymes were mapped onto the KEGG pathways, we observed the presence of TCA cycle related enzymes (Figure 2a). The relative abundance distribution of the metabolic enzymes within the crude cytoplasmic fraction did not readily connect their nuclear presence to a cytoplasmic contamination as it seemed to be mostly concentration-independent (Figure 2b). For instance, of the most abundant cytoplasmic enzymes, malate dehydrogenase 1 (MDH1), aconitase 1 (ACO1) and isocitrate dehydrogenase 1 (IDH1) were not detected in the nuclear fractions. For the mitochondrial compartment and TCA cycle per se, enzymes with similar abundances in the crude cytoplasmic fraction were not equally present in the nucleus (i.e. PCK2 vs IDH3G, SUCLG2 vs OGDH).

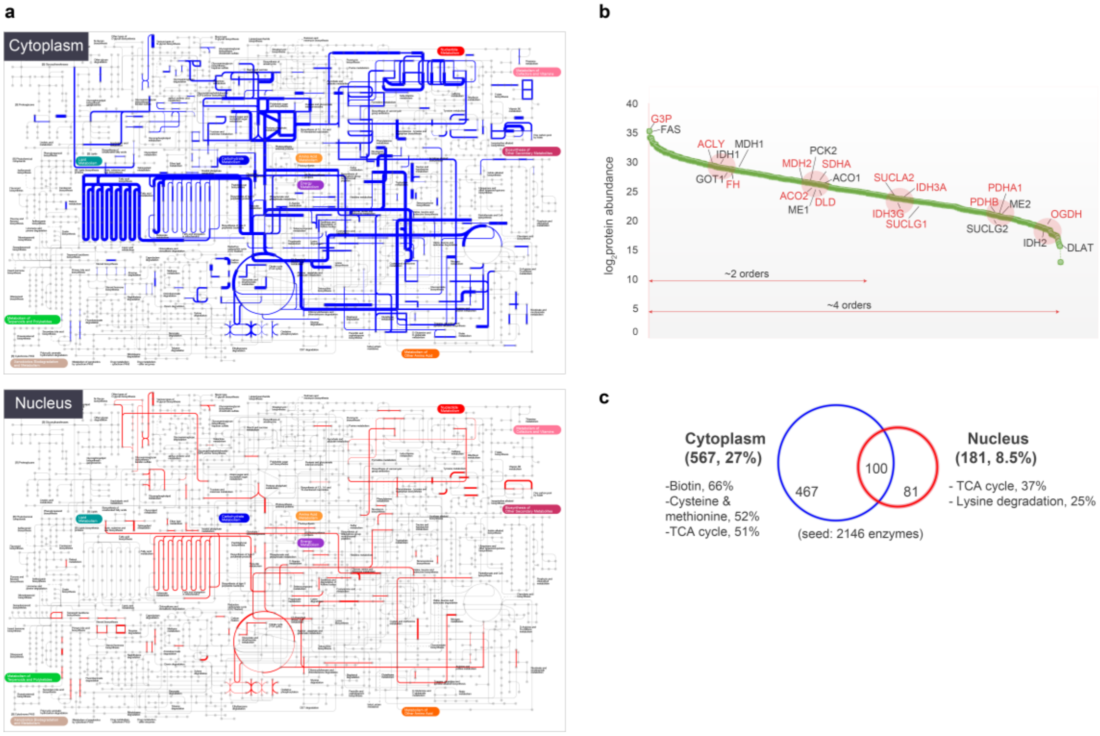

Figure 2: The nuclear metabolic proteome. (a) Metabolic enzymes detected in the crude cytoplasm (upper panel) and in the nucleus (lower panel). The line width corresponds to the relative abundance of the enzymes within the respective compartment. (b) Relative abundance distribution of the metabolic enzymes within the crude cytoplasm. The metabolic enzymes that are detected in the nucleus are depicted with red letters. (c) Top enriched metabolic pathways in the crude cytoplasmic and nuclear fractions.

The reconstruction of parts of the TCA cycle in the nucleus was thus intriguing. As already discussed in the introduction, individual enzymes of the TCA cycle have a nuclear localization connected to the biosynthesis of metabolites utilized for histone and DNA modifications[30,88]. As such, the existence of larger networks that could facilitate more effective substrate channeling for a localized production of important metabolites might not come as a surprise. This idea of metabolons has already been studied for individual pathways, including TCA cycle, albeit within the more traditional mitochondrial perspective[89–93]. This prompted us to further investigate the TCA enzymes for the presence of nuclear localization motifs (NLS) in their sequence which, although not a necessity, would provide additional evidence for the possibility of a nuclear topology. Except citrate synthase (CS), we identified NLS for every enzyme, or individual partners when it came to complexes, up to the generation of fumarate (Figure 3). In total, we identified NLS for 9 enzymes directly associated to TCA cycle, and for 12 enzymes when a more extended network was included (Figure 3). Apart from OGDC[34], none of the other enzymes had a previously characterized NLS. For PDC, which has been shown to localize in the nucleus with unknown mechanism so far[30], we identified NLS for two subunits, namely PDHB and PDHA2, while for OGDC, we identified an additional NLS for the OGDH subunit, apart from the already described one on DLST[34].

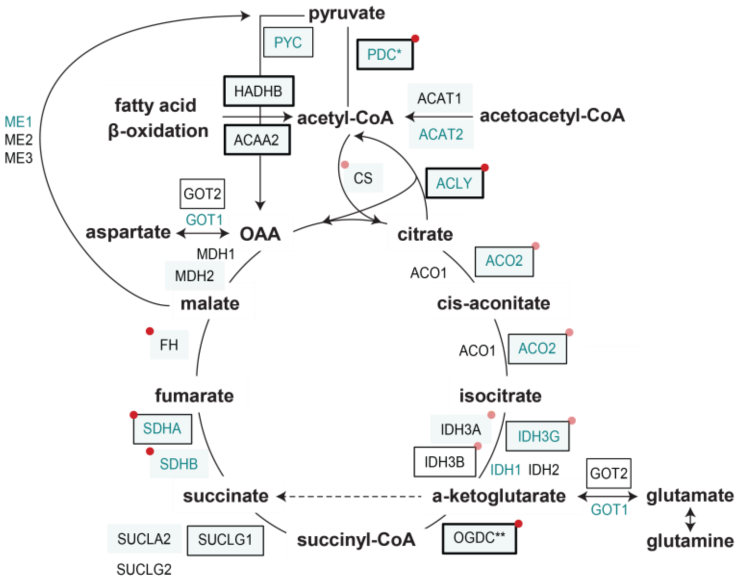

Figure 3: Nuclear Proteome and Nuclear Localization Signal motifs suggest the potentiality of a nuclear subnetwork of TCA cycle enzymes. With light background are the enzymes detected in the nuclear proteomics datasets. A black box around the enzymes denotes the presence of Nuclear Localization Signal (NLS) in their sequence. Green letters represent the enzymes with a nuclear topology in Human Protein Atlas[94,95] database. Enzymes proven by other studies to localize in the nucleus are depicted with a red dot. Enzymes with a condition-specific localization in the nucleus are marked with a light red dot.

Other NLS motifs for enzymes connected to TCA cycle anaplerosis included those for the pyruvate carboxylase (PC) and for aspartate aminotransferase (GOT2). Interestingly, we also identified NLS for two enzymes involved in fatty acid β-oxidation which is connected to acetyl-CoA production, namely the 3-ketoacyl-CoA thiolase (ACAA2) and the beta subunit of trifunctional enzyme (HADHB). To have an initial spatial resolution, we mapped the TCA cycle enzymes to the Human Protein Atlas[94,95] and we found nuclear immunofluorescence evidence for a number of them, while two enzymes presented a cell line-specific nuclear topology (i.e. ACO2, PC). Figure 3 encompasses all the aforementioned information which, collectively, pointed towards the potentiality of a TCA cycle subnetwork present in the mammalian nucleus.

[U-^{13}C]-Glutamine and [U-^{13}C]-Citrate label TCA cycle key intermediates delineating an active metabolic pathway operating in the HeLa nucleus

The cumulating nuclear proteomics, NLS and imaging data described above indirectly inferred the possibility of a local nuclear production of specific TCA cycle-related metabolites by the presence of the corresponding enzymes. The quantification of the metabolites of interest would provide additional evidence, however, compartment-specific metabolic measurements are inherently difficult since metabolism readily changes upon experimental stresses during the laborious organelle isolation procedures. The fast turn-over rates of the metabolites along with their excessive diffusion through the membranes result in a non-representative capture of the *in vivo* subcellular metabolic physiology and composition. Additionally, only the presence of an enzyme is not necessarily coupled to an enzymatic/metabolic activity. To directly assay the metabolic functionality of such a network, we therefore decided to introduce in isolated nuclei heavy isotope (^{13}C) labeled compounds representing key entries and nodes of TCA cycle, and trace labeled carbon incorporation to the following metabolites.

Our experimental scheme (Figure 4) involved the isolation of nuclei from HeLa cells, incubation with uniformly ^{13}C-labeled substrates (U-^{13}C) for 1 or up to 5 hours, harvesting and washing with PBS to remove any extra-nuclear contaminants, and finally collection for metabolomics analysis. Polar metabolites were extracted from the collected nuclei and the mass isotopomers of TCA cycle intermediates were quantified by Gas Chromatography - Mass Spectrometry (GC-MS). In parallel, total cell lysates (lysed cells corresponding to a mixture of cytoplasm with all subcellular compartments, some partially lysed and some intact) were incubated under the same buffer and conditions, serving as a positive control for the labeling tracing. Finally, to assess the purity of the isolated nuclei, samples were collected for electron microscopy analysis, immunoblots and immunofluorescence against known mitochondrial proteins. Of note, we decided to add all necessary cofactors (NAD$^+$, FAD$^+$, ADP, ATP, etc, see methods) in equimolar concentrations to ensure that the reactions that did not occur were not due to cofactor(s) limitation.

Assessment of nuclei purity

To assess the status of the isolated nuclei regarding possible contaminations, we performed electron microscopy analysis, as well as immunoblotting and immunofluorescence against known cytoplasmic and mitochondrial proteins. Electron microscopy showed that the nuclei were free of intact mitochondria, however they were surrounded by membranes the origin of which is not possible to be verified by electron microscopy (Figure 5). Immunoblots with antibodies against beta-tubulin revealed no cytoplasmic contamination, while antibodies against mitochondria gave no signal for known matrix or shuttling mitochondrial proteins (isocitrate dehydrogenase 2, IDH2, and cytochrome c, respectively), and a positive signal for the mitochondrial membrane protein COXIV (Figure 5). When immunofluorescence for COXIV was performed, no staining of the nuclei was observed (Figure 5). These findings denoted that the isolated nuclei, although free of whole intact mitochondria, bear fragments of mitochondrial membranes. As a result, we decided to include into our 13C-labeling scheme, 13C-substrates that could serve as guides of a possible mitochondrial contamination.

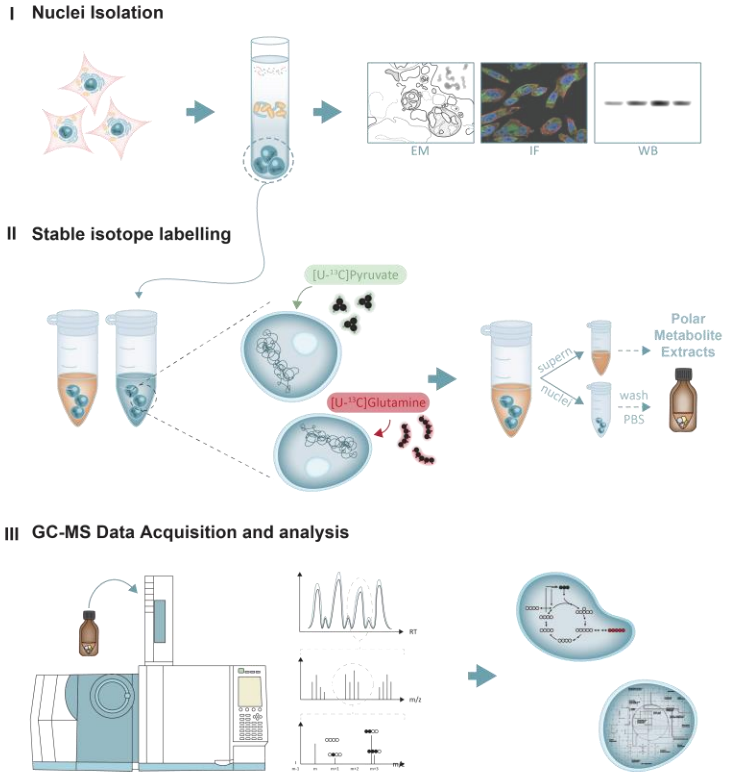

Figure 4: Experimental design of the nuclei labeling with ^{13}C metabolic substrates.

Along these lines, as a first step, we incubated nuclei with [U-^{13}C]-pyruvate. On a cellular-wide context, pyruvate is converted to acetyl-CoA (pyruvate dehydrogenase complex, PDC activity) which subsequently condensates with oxaloacetate for the generation of citrate (citrate synthase, CS activity). Part of the citrate further continues in the TCA cycle for the production of intermediates (Figure 6). Utilizing [U-^{13}C]-pyruvate as a substrate in the isolated nuclei did not result in the detection of any labelled TCA cycle intermediate. In terms of metabolite presence, apart from the [U-^{13}C]-pyruvate, we only detected small amounts of unlabeled succinate and aspartate. The latter two metabolites were consistently detected in the intra-nuclear space of all isolated nuclei along with various amino acids and sugar compounds (i.e. glucose, fructose). In the cell lysates, the utilization of [U-^{13}C]-pyruvate resulted in labeling in expected compounds, including aspartate and succinate, denoting thus that the cultivation conditions were not a limiting factor for the absence of labeling in the isolated nuclei (appendix, Figure 30). Moreover, the absence of labeled metabolites were in line with expectations, since although nuclear PDC has been shown to produce acetyl-CoA for histone acetylation in certain mammalian cell lines[30], nuclear citrate synthase (CS) presence and activity has been defined only for a very specific developmental stage in mice embryos[31]. As such, pyruvate carbon entry in metabolites other than acetyl-CoA is not expected, and the [U-^{13}C]-pyruvate tracer served as an internal negative control regarding the purity of our system. Acetyl-CoA cannot be measured with GC-MS, thus we could not conclude whether PDC was active and the subsequent enzymes were not, or whether in our system this reaction did not occur too.

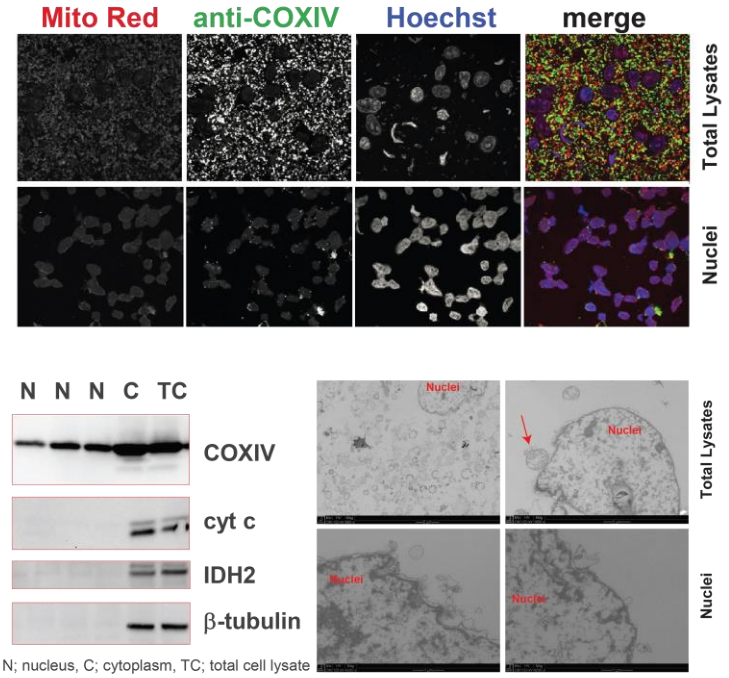

Figure 5: Nuclei purity assessment. (a) Immunofluorescence of total cell lysates (upper panel) and nuclei (lower panel). Mito Red (red) stains cardiolipin, a lipid detected in the inner mitochondrial membrane. COXIV (green) is an inner mitochondrial membrane protein. DNA was labeled with Hoechst staining. (b) Immunoblotting for COXIV (inner mitochondrial membrane), cytochrome c (mitochondrial shuttling protein), isocitrate dehydrogenase 2 (IDH2, mitochondrial matrix protein) and beta-tubulin (cytoplasmic protein) in nucleus (N), cytoplasm (C) and total cell lysates (TC) samples. (c) Electron microscopy of total cell lysates (upper panel) and nuclei (lower panel). Red arrows point to mitochondria.

[U-^{13}C]-glutamine is a precursor of key TCA intermediates in the HeLa nucleus

A major anaplerotic substrate of TCA is the amino acid glutamine which feeds the cycle predominantly at the level of alpha-ketoglutarate. Alpha-ketoglutarate is an important epigenetic-related metabolite, as it is a necessary co-factor for the function of specific DNA and histone demethylases. As such, we sought to determine whether [U-^{13}C]-glutamine supplementation could contribute in the production of alpha-ketoglutarate or any other metabolite in the nucleus. Incubating HeLa nuclei with [U-^{13}C]-glutamine resulted in the production of labeled glutamate, succinate, fumarate and aspartate (Figure 6). Regarding alpha-ketoglutarate, we were not able to detect even the unlabeled form. A possible explanation could attributed to a very low concentration, or to its highly labile nature, as well as to its instability, as mentioned in studies performed in cell lysates[96,97]. Regarding the detected labeled metabolites, glutamate and aspartate labeling reached 90% and 84%, respectively, already after 1h of incubation, while succinate and fumarate approached 45% at 5h. The slight reduction in the labelling percent of succinate at 5h compared to 1h, could be coupled to its conversion to fumarate whose labelling percent increased from 21% (1h) to 45% (5h) (Figure 6). It should be noted that although the labeling percent in fumarate approximated the one in succinate at 5h of incubation, the peak area for the labeled fumarate ion was at the detection limit, implying that very low amounts were generated.

The presence of labeled glutamate is not unexpected as glutamine can non-enzymatically convert to glutamate. Glutamate subsequently enters the TCA cycle upon conversion to alpha-ketoglutarate, a step that in mammalian cells is catalyzed by glutamate dehydrogenase (GLUD1) or by aspartate transaminase (GOT1 or GOT2) (Figure 6). GLUD1 and GOT2 are currently considered mitochondrial enzymes both of which possess an NLS, albeit with no apparent nuclear staining in literature data or in Human Protein Atlas. GOT1 resides mainly in the cytoplasm, however there are supported evidence in Human Protein Atlas for a nuclear presence too with unexplored function so far. Upon alpha-ketoglutarate formation, succinate production can derive from two possible pathways (Figure 6). The first, and more prevailing in cells, pathway encompasses the sequential activities of alpha-ketoglutarate dehydrogenase complex (OGDC) and succinate-CoA ligase (SUCL) for the generation of succinyl-CoA and succinate, respectively. The topology of OGDC in the nuclear compartment has been recently discovered and was connected to a local generation of succinyl-CoA, the succinyl moiety of which is used for the succinylation of histones[34]. The alternate option involves the activities of alpha-ketoglutarate dependent dioxygenases (demethylases) which directly produce succinate from the decarboxylation of alpha-ketoglutarate in the process of hydroxylating various substrates, including histones and 5-methylcytosine (5mC) residues in DNA. For these demethylases, succinate can antagonize alpha-ketoglutarate and act as a competitive inhibitor[18]. Succinate will then convert to fumarate through a reaction that requires the catalytic activity of SDHA-SDHB heterodimer (Figure 6). Although in the mitochondria the heterodimer is anchored in the inner mitochondrial membrane through the subunits SDHC and SDHD, upon several conditions it dissociates still retaining its functional activity[98,99]. Interestingly, SDHA-SDHB presence and activity in nucleoli of normal and cancerous cells was first demonstrated in the early 60s[35,36]. Current topological evidence from Human Protein Atlas support a clear localization of the catalytic subunits in the nucleus and nucleolus, however their role in this compartment has not been further investigated.

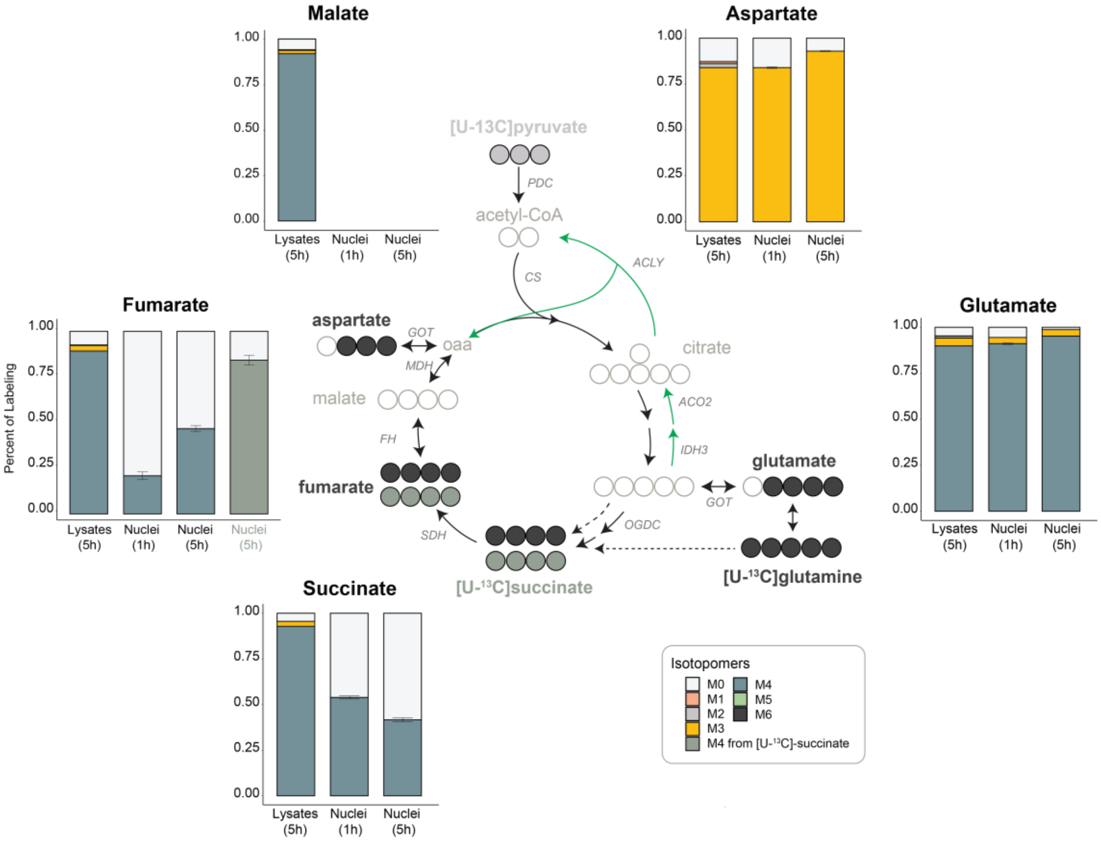

Figure 6: Glutamine is a precursor of key TCA intermediates in the HeLa nucleus. Mass isotopomers analysis of specific TCA metabolites in the nucleus and total cell lysates after incubation with [U-^{13}C]-glutamine, or [U-^{13}C]-succinate, or [U-^{13}C]-pyruvate, for 1h or 5h. The figure also includes a schematic representation of the TCA cycle with the expected carbon atom (circles) transitions. Filled black, dark green and grey circles denote the ^{13}C carbons derived from [U-^{13}C]-glutamine, [U-^{13}C]-succinate and [U-^{13}C]-pyruvate, respectively. ^{12}C carbons are depicted as empty circles. For aspartate and glutamate, instead of the full carbon backbone, the detected ions corresponded to a three and four fragment ion, respectively (the missing carbon is depicted as a grey circle). Dark and grey colored metabolites correspond to quantified and non-quantified/non-detected features in the nuclei samples, respectively. Error bars denote the mean ± standard error of the mean (sem) of 3 independent experiments. Abbreviations: oaa, oxaloacetate.

Aspartate is a proteinogenic metabolite, as well as a nitrogen and carbon donor for the synthesis of purines and pyrimidines[62,63,100,101]. The presence of labeled aspartate in the nucleus necessitates the presence of labeled oxaloacetate, since the carbons of the latter are required for aspartate biosynthesis through the activity of aspartate aminotransferase (GOT 1 or GOT2)[49]. To generate oxaloacetate, the most predominant pathway is through the oxidative route (clockwise) of TCA cycle - alpha-ketoglutarate produces succinate followed by conversion to fumarate, malate and eventually to oxaloacetate (Figure 6). An alternative pathway, operating under specific conditions, involves the reductive carboxylation of alpha-ketoglutarate (anti-clockwise) into isocitrate and then to citrate[102,103]. Citrate is eventually cleaved by ACLY to acetyl-CoA and oxaloacetate (Figure 6). The oxidative and reductive TCA pathways can be dissolved by the presence of specific isotopomers in the labelling patterns of individual metabolites[96,97]. This however requires the detection and quantification of ions that correspond to the full carbon backbone of the metabolites, which, for the case of aspartate was not possible owing to the detection of only a three carbon fragment and not of the whole four carbon backbone.

To indirectly resolve which pathway contributes to aspartate biosynthesis, we decided to supplement the nuclei with [U-^{13}C]-succinate. The presence of labeled aspartate in this case would imply the activity of the oxidative route, considering that alpha-ketoglutarate to succinate reactions can be regarded to operate only in the forward and not backward direction[96]. The [U-^{13}C]-succinate tracer did not result in the detection and labeling of any metabolite apart from fumarate (Figure 6), suggesting that the fumarate to oxaloacetate axis is not taking place in the nucleus. The absence of labeling in aspartate indirectly connoted that the reductive route should then contribute to its generation. In this scenario, a metabolic tracer towards the end of the reductive pathway, for example citrate, would provide for the labeled aspartate carbons.

[U-^{13}C]-citrate reconstructs a citrate to succinate axis in HeLa nuclei

As a consequence of the aforementioned results, we next aimed to assess whether citrate, a key metabolic intermediate laying in the intersection for the generation of two important epigenetic-related metabolites in cells, that of acetyl-CoA and alpha-ketoglutarate, could in turn contribute in the production of aspartate or other TCA-related intermediates in the nucleus. Feeding HeLa nuclei with [U-^{13}C]-citrate as a labeling substrate resulted in the detection and labeling of a single metabolite, that of succinate, 80% of which was labelled after 5h of incubation (Figure 7). In total cell lysates, all expected metabolites were detected, including the ones downstream of succinate (i.e. fumarate, malate, aspartate etc.), along with the respective isotopomers (Figure 7).

In mammalian cells, the predominant pathway for succinate generation from citrate involves as a first step the citrate conversion to isocitrate by aconitase (ACO1 or ACO2), and then to alpha-ketoglutarate by isocitrate dehydrogenase (IDH1, IDH2 or IDH3) (Figure 7). Both ACO2 and IDH3 feature a metabolic activity in the nucleus connected to alpha-ketoglutarate production, albeit such an activity was linked only to mouse embryonic development and found absent in later stages[31]. Interestingly, ACO2 was identified in yeast nuclei with uncharacterized function so far[33], while topological evidence from mammalian cells reveal cell line-specific nuclear localization as it is observed in nuclei of U-251 MG (glioblastoma) cells. IDH3G (one of the subunits of IDH3), and IDH1 to lesser extent, show a nuclear localization in multiple mammalian cell lines. Finally, as discussed earlier, our sequence analysis identified NLS for both ACO2 and IDH3G enzymes (Figure 3). To verify whether succinate production in the nucleus involved indeed the activity of aconitase, we treated isolated nuclei with the aconitase inhibitor fluorocitrate in the presence of [U-^{13}C]-citrate. This resulted in a two-fold reduction in the labeling of succinate (40%) supporting such an activity to be present in the nucleus of HeLa cells (Figure 7).

Regarding our hypothesis for aspartate generation through the reductive route, the [U-^{13}C]-citrate tracer did not confirm it as we did not observe any labeled aspartate (Figure 7). Retrospectively, we hypothesized that this could partially be attributed to the concentration of citrate used in this study - citrate is a known metal chelator and as such inhibits the activities of many enzymes. In addition, citrate is a known allosteric inhibitor for several metabolic enzymes, including SDH[104]. The latter could explain why the [U-^{13}C]-citrate tracer did not reach a labeled fumarate in the nucleus (Figure 7) contrary to the [U-^{13}C]-glutamine (Figure 6). Indeed, when total cell lysates were incubated with a concentration of [U-^{13}C]-citrate tracer that contributed in 80% of succinate being labeled (approximating thus the percentage of labeled succinate found in the nucleus), fumarate was unlabeled indicating an inhibition of SDH activity (Figure 7). In the same experiment, we observed labeled malate pointing to its generation from oxaloacetate which derived from citrate through the activity of ACLY (Figure 7). Although oxaloacetate contributed to malate, there was no labeling in aspartate denoting that GOT enzyme did not operate under these conditions and was inhibited by citrate (Figure 7c). Overall, the results in total lysates could partially explain our hypothesis regarding possible reasons why reductive aspartate synthesis was absent in the nucleus after [U-^{13}C]-citrate supplementation, however, further experimental repetition with lower citrate concentration levels in nuclei is required to validate its actual presence and contribution in the production of aspartate.

Collectively, the labeling data revealed that a functional metabolic pathway operate in the nucleus, a pathway that is made of parts of the TCA cycle. Several lines of evidence offered by internal negative control labeling results supported the purity of our system, for instance the absence of contribution in metabolic nuclear pools after [U-^{13}C]-pyruvate supplementation, or the absence of labeling of specific TCA metabolites in the nucleus (i.e. malate) contrary to total cell lysates. In addition, the design of our experimental scheme involved the washing of the nuclei from the extranuclear medium prior to collection and metabolite extraction, denoting that the labeling results reflected intranuclear labeled pools. Our results thus indicated that [U-^{13}C]-glutamine, [U-^{13}C]-citrate and [U-^{13}C]-succinate are actively converted to metabolites with previously recognized roles in epigenetic modifications, and as such, they complement existing studies by further reconstructing an entire axis up to the formation of fumarate being active in the HeLa nucleus.

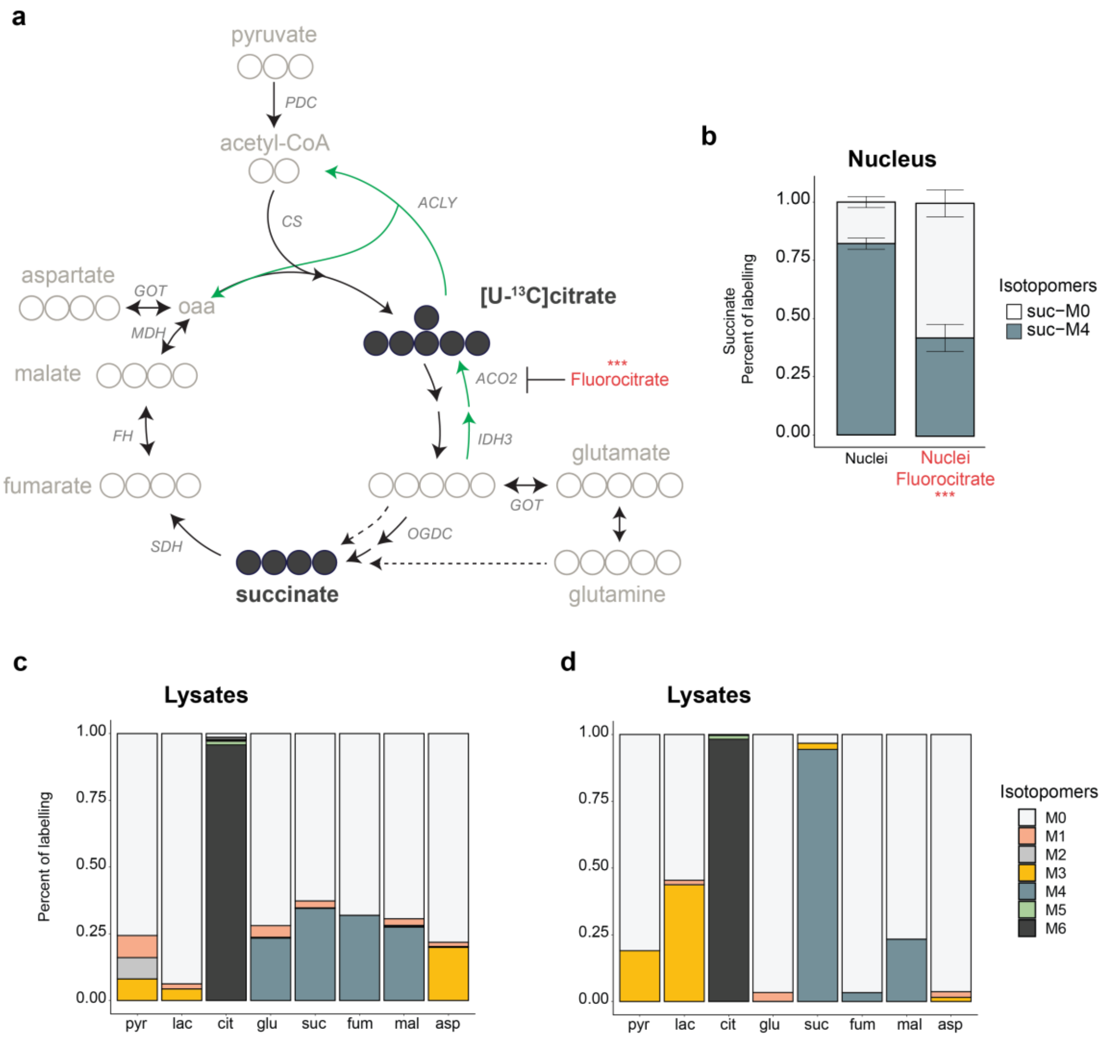

Figure 7: [U-^{13}C]-citrate supplementation reconstructs a citrate to succinate axis actively present in HeLa nuclei. (a) Schematic representation of the TCA cycle along with the carbon atom (circles) transitions of specific metabolic intermediates obtained after supplementation and 5h incubation of isolated nuclei with [U-^{13}C]-citrate. Filled black circles denote the ^{13}C carbons derived from [U-^{13}C]-citrate. ^{12}C carbons are depicted as empty circles. Dark and grey colored metabolites correspond to quantified and non-quantified/non-detected features in the nuclei samples, respectively. Abbreviations: oaa, oxaloacetate. (b) Succinate biosynthesis in the nucleus involves the activity of aconitase as a first enzymatic step. After [U-^{13}C]-citrate supplementation, ~80% of succinate was labeled (M4 isotopomer). Incubation with an aconitase inhibitor (fluorocitrate) in the presence of [U-^{13}C]-citrate resulted in a two-fold reduction of the labeled succinate (M4 isotopomer). The graph corresponds to the mass isotopomers analysis of succinate (mean ± standard error of the mean (sem) of 3 independent experiments). (c) Mass isotopomers analysis of TCA intermediates after supplementing total cell lysates with 2µM [U-^{13}C]-citrate – all expected metabolites and isotopomers are detected. (d) Mass isotopomers analysis of TCA intermediates after supplementing total cell lysates with 10µM [U-^{13}C]-citrate results in inhibition of several enzymatic steps as depicted by the absence of specific labeled metabolites. Abbreviations: pyr, pyruvate; lac, lactate; cit, citrate; suc, succinate; glu, glutamate; fum, fumarate; mal, malate; asp, aspartate. In (b), (c) and (d), M0 (unlabelled) to Mn (labelled in n carbons) indicate the different mass isotopomers.

Towards the identification of a nuclear niche using proximity biotinylation mass spectrometry

To further investigate the aforementioned subnetwork of metabolic enzymes, we decided to explore their protein-protein interaction (PPI) network. While these enzymes have been studied with a wide range of biochemical methods, very few have been used as baits in interaction networks and often are confined to a mitochondrial metabolic perspective. A systematic characterization of their PPI could point to additional localizations and functionalities. To this end, we used proximity labeling mass spectrometry based on the BioID system[105]. In this method, the protein of interest is fused to a promiscuous prokaryotic biotin ligase which, in the presence of excess biotin, catalyzes the biotinylation of lysine residues in the interacting and neighboring proteins. The biotinylated proteins are subsequently affinity-purified with streptavidin beads and identified by mass spectrometry. The advantage of this method that prompted us to incorporate it in our experimental scheme is that it allows for mapping of the protein-protein interactome in living cells. Because of this, and owing to the nature of the modification which remains present even after harsh lysis conditions, weak and transient interactions are also able to be captured. Since the biotin ligase allows the biotinylation of any protein within a 10-50 nm radius[106,107], the identified interactors are not constrained to only direct but also to proximal ones, allowing thus to probe the molecular environment and identify additional subcellular neighborhoods. In this way, the BioID can offer higher sensitivity in capturing more dynamic or smaller niches especially when the relative distribution of the examined protein is enriched for a specific compartment which could mask the detection of the protein elsewhere when conventional methods are used (i.e fluorescence microscopy).

Our targets included aconitase 2 (ACO2), isocitrate dehydrogenase 3 (IDH3) and the 2-oxolglutarate dehydrogenase complex (OGDC) covering the citrate to succinly-CoA pathway (Figure 8). IDH3 has three distinct subunits, IDH3A, IDH3B and IDH3G, and is organized as a heterotetramer composed of the heterodimers IDH3A-IDH3B and IDH3A-IDH3G[108]. Out of the three subunits, the IDH3B and IDH3G contain a nuclear localization signal and we chose to tag IDH3G owing to a more visible nuclear staining in the Human Protein Atlas database[94]. OGDC is comprised of three subunits, the 2-oxoglutarate dehydrogenase (OGDH), the dihydrolipoamide succinyl-transferase (DLST) and the dihydrolipoamide dehydrogenase (DLD). Since the OGDH subunit contains a nuclear localization signal we decided to tag this one as a representative of the whole complex. As a negative control sample, we used the isocitrate dehydrogenase 2 (IDH2) which is a purely mitochondrial enzyme. In the study we included as additional control samples, the pyruvate dehydrogenase B (PDHB), a subunit of the PDC complex which has been shown that although it mainly localizes in the mitochondria it can shuttle to the nucleus as well[30], and isocitrate dehydrogenase 1 (IDH1) which has a predominantly cytoplasmic localization but a small fraction is detected in the nucleus too[94].

The six aforementioned enzymes were fused to a smaller version of biotin ligase (BioID2) which has been shown to minimize improper subcellular localization of the engineered proteins[106]. In addition to BioID2, FLAG tag was also included in the sequence. For all the enzymes, the BioID2-FLAG was introduced in the C-terminus, while for PDHB, IDH2 and OGDH, N-terminally fused versions were created as well. Each engineered enzyme was stably integrated in the tetracycline inducible Flp-In T-Rex HeLa cell system. The induction of the protein expression was facilitated by addition of tetracycline for 24hr, followed by biotin supplementation for protein biotinylation for another 24hr. The labeling and affinity purification of the biotinylated proteins was performed for four biological replicates per case, including wild type parental cells, and the mass spectrometry was performed with 10-plex tandem mass tag (TMT) labeling (Figure 8).

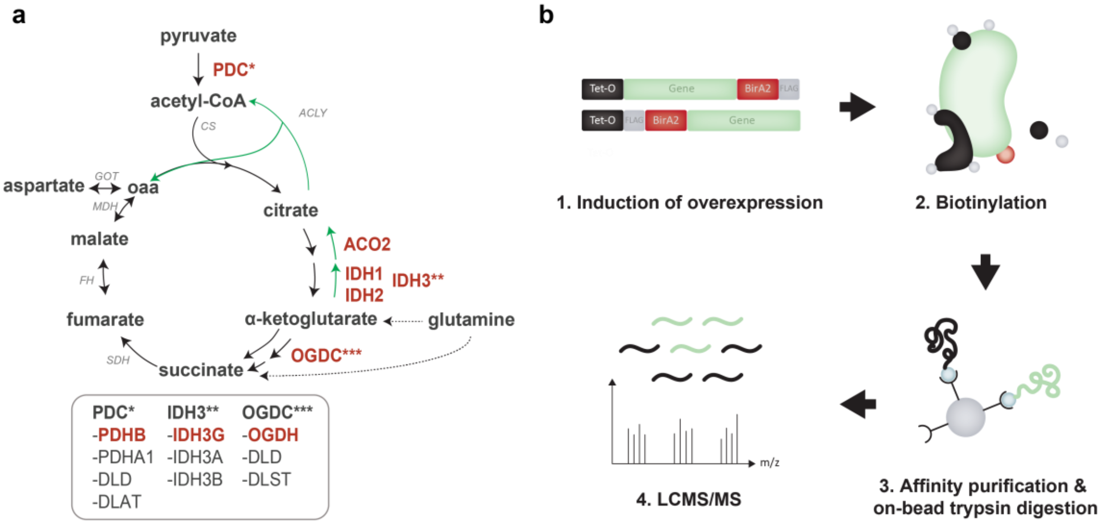

Figure 8: Proximity biotinylation mass spectrometry based on the BioID method. (a) Selected targets are depicted with red color. For the complexes PDC, IDH3 and OGDC, the selected subunits was based on the presence of NLS in their protein sequences. (b) Graphical abstract of the experimental scheme.

To create a first rough network of the proximal interactome of each enzyme, each stable cell line was compared to wild-type parental HeLa cells. The significant results were further refined by excluding proteins that have been identified in negative control HeLa-BioID samples deposited in the CRAPome database[109]. Pairwise pearson correlation analysis between the rough proximal networks revealed two main clusters (Figure 9a). The C-tagged versions of the mitochondrial enzymes were positively correlated with each other, forming one of the two main clusters, and negatively correlated with the cytoplasmic IDH1. Consistent with the expectations, Gene Ontology (GO) analysis revealed the mitochondrion as the top enriched cellular component term (appendix, Table 5). However, for the three enzymes that were analyzed as both C- and N-terminally fused versions, the correlation between the two termini was negative, with the N-fused ones grouping with IDH1, forming all together the second main cluster (Figure 9a). Here, gene ontology (GO) analysis revealed as top cellular component enrichments the cytoplasm and cytoskeleton (Supplementary section, Table 5). Further subcellular annotation of the hits showed that no purely mitochondrial proteins were detected in the N-fused versions. In addition, the known interaction partners of the OGDH and PDHB were only captured in the C-tagged fusions. While the absence of the known partners could be associated with the sampling of different interactors by the two termini, altogether the observations pointed to inability of the N-fused enzymes to access the mitochondria. This is possibly attributed to inaccessibility of the mitochondrial import signal that most mitochondrial proteins carry in their N-terminus[110]. As a result, we decided to exclude the N-fused versions from further analysis.

The aforementioned results depict a first rough representation of the proximal interactome. To further refine them, we decided to use one of the N-fusions, namely N-fused IDH2, as an additional negative control sample which would account for further eliminating unspecific proteins as a result of the exogenous expression of the biotin ligase in the cells. With the addition of this control, the final number of significant proximal interactors ranged from 108 (C-tagged OGDH) to 271 (C-tagged IDH2) (Figure 9b). An overview of pearson correlation coefficients can be viewed in Supplementary Figure 2. ACO2 and IDH2 had the highest correlation (R^2 = 0.95), consistent with the fact that they share an immediate neighborhood participating in the catalysis of sequential reactions in the mitochondrion. Together with PDHB, ACO2 and IDH2 formed a highly correlated cluster (core), sharing 239 proximal interactors, followed by IDH3G which shared 139 proximal interactors with the above core (139 out of the 157 total interactors, 88% overlap) (Figure 9b). OGDH shared the least number of proximal interactors with the core (83 out of 108, 76% overlap), and more with IDH3G (89 out of 108, 82% overlap) (Figure 9b).

While the vast majority of the proximal interactors represented the primary molecular neighborhood of the enzymes (Figure 9d), we were able to resolve smaller clusters with distinct biological processes and probable subcellular niches associated with specific enzymes. This is instantiated in Figure 9d with clusters 2, 4 and 5. Of particular interest, cluster 4 represents a highly enriched group in OGDH and IDH3G enzymes, followed by mediocre enrichment in ACO2 (out of the total number of proteins detected in the cluster, 11% were statistically significant in ACO2) but largely excluded from IDH2 and PDHB. Gene ontology analysis revealed as top biological process the rRNA metabolic modifications (p value, $3.3\text{x}10^{-6}$), and as subcellular components the pre-ribosome (p value, $4.3\text{x}10^{-7}$) and the nucleolus (p

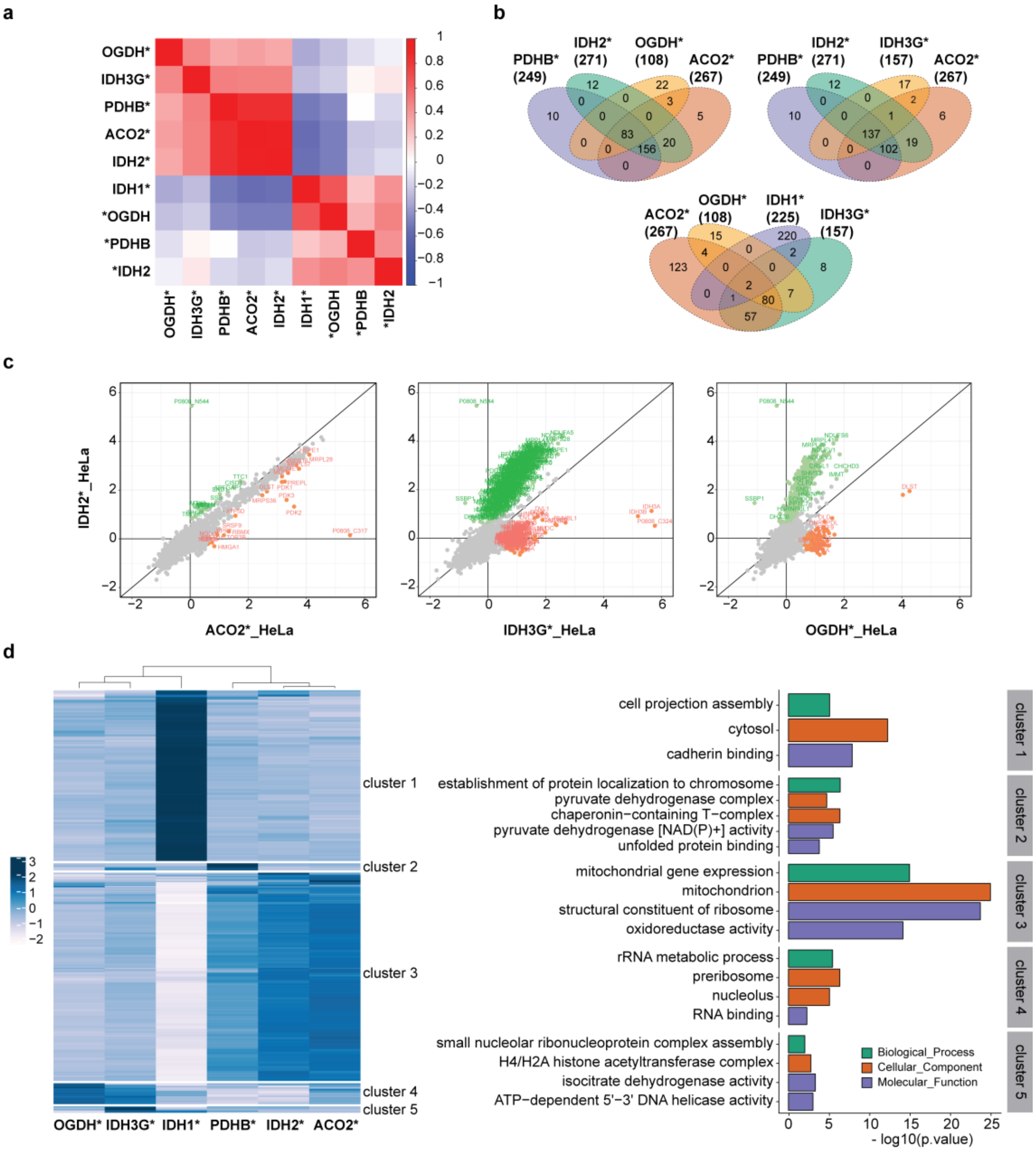

Figure 9: Exploration of the proximal interactomes of ACO2, OGDH and IDH3G. (a) and (b), Pairwise Pearson correlation analysis and Venn diagram of the significant biotinylated proteins detected in each engineered cell line, respectively. An asterisk at the beginning and the end of the protein name is representative of the C- and N-fusion of the biotin ligase-FLAG, respectively. (c) Pairwise comparisons of the significant proximal interactors (log2 fold changes when compared to wild type HeLa) between the C-fused i) ACO2 and IDH2, ii) IDH3G and IDH2, iii) OGDH and IDH2. Red and green color indicates the proximal interactors that are significantly enriched in ACO2, IDH3G or OGDH when compared to IDH2 cell line, and in reverse, respectively. (d) k-means clustering of the significant proximal interactors revealed five major clusters (left panel) that were further characterized by Gene Ontology (GO) enrichment analysis for the terms “Biological Process”, “Molecular Function” and “Cellular Component” (right panel) using the g:Profiler[112,113]. Color gradient for the k-means clustering corresponds to $\log_2$ fold changes of each engineered cell line compared to wild type HeLa.

value, 8.1×10^{-6}) which is connected to the transcription of ribosomal genes and rRNA processing and assembly with ribosomal proteins[111]. Consistent with the aforementioned, the cluster was also enriched for RNA binding molecular function (p value, 5.1×10^{-3}). Cluster 5, presented in IDH3G and moderately in IDH1, was enriched for isocitrate dehydrogenase activity, as expected. Interestingly, we also detected proteins of a complex with histone acetyltransferase, as well as DNA helicase activity (Figure 9d).

To have a better overlook of the topological distribution of the identified proximal proteins, we used subcellular localization information derived from Uniprot, Human Protein Atlas[94] and Compartment[114] databases (Figure 10). OGDH had the highest fraction of proteins characterized as of “purely nuclear” topology (14.9%, 16 proteins), followed by IDH1 (9.8%, 22 proteins) and IDH3G (5.2%, 8 proteins). The same topological category in ACO2 involved five proteins (1.9%). However, two of them (SPRYD4 and TRMT11 - the first of unknown function and the latter connected to tRNA methyltransferase activity) were also detected in the negative control IDH2, as well as in PDHB, being the only “purely nuclear” proteins in these samples (corresponding to 0.8%). In a recently published BioID study[107], these two proteins were present amongst the proximal interactors of four mitochondrial matrix proteins (CH60, TR61B, PDK1 and MRM1) and absent from the interactome of any other non-mitochondrial protein, corroborating thus that although they are annotated as purely nuclear in the databases, they are associated with the mitochondria. We further analyzed the results of this study to get an estimate of their “purely nuclear” topological distribution and compare it to ours. The proximal interactomes of CH60 and PDK1 had a 0.8% characterized as of purely nuclear topology involving the same two proteins described above (TRMT11, SPRYD4). For TR61B we found one additional nuclear protein (3 proteins in total, 1.3%) and for MRM1 three more (5 proteins in total, 2%). The fact that in at least one of these mitochondrial protein there is still a 2% of nuclear proteins confounded the significance of our ACO2 “purely nuclear” results. For AIFM1, which is a mostly mitochondrial matrix protein but also detected in the nucleus and cytoplasm, there was a 5.2% of purely nuclear features (9 proteins in total), similar to the percentage of our IDH3G.

Taken together, the aforementioned indicated the correct localization of the C-tagged enzymes to their primary mitochondrial compartment (or cytoplasmic for the case of IDH1), but also revealed a fraction of proteins connected exclusively to the nucleus for the cases of OGDH, IDH3G and IDH1. For ACO2, we identified three “purely nuclear” proteins, however, owing to the small percentage, the nuclear topology for ACO2 itself remained ambiguous. Finally, although PDHB has been described in the literature as having a nuclear topology[30] in addition to the mitochondrial one, we were not able to identify any purely nuclear protein in our dataset.

The high number of proximal interactors detected with the above analysis is consistent with the fact that the BioID method captures the immediate environment of the tagged protein. In order to identify more preferential and more specific interactors even within the same

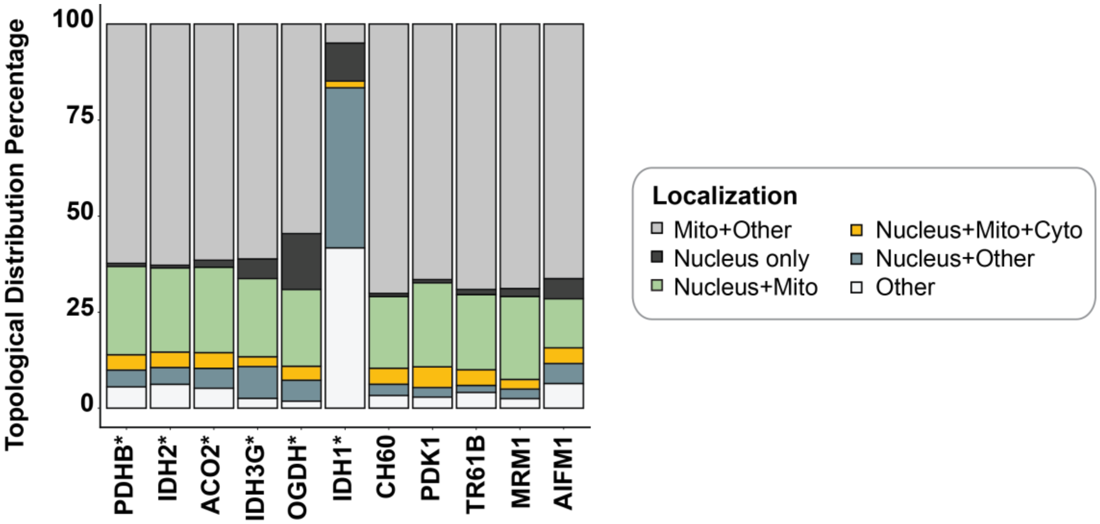

Figure 10: Subcellular distribution of the proximal interactors. The topological annotation was based on Uniprot, Human Protein Atlas[94,95] and Compartments[114] databases. The proximal interactors of CH60, PDK1, TR61B, MRM1 and AIFM1 were obtained from the study of Liu et al[107] and refer to mitochondrial proteins fused to a larger version of biotin ligase stably integrated in Flp-In T-REx 293 cells.

environment, we decided to compare each enzyme to another fellow, in addition to the wild type parental HeLa cells and the N-fused IDH2. This way for example, ACO2 was compared to the C-fused IDH2 or C-fused PDHB, to parental HeLa and N-fused IDH2, and the union of the significant interactors was taken into consideration. The number of significant proximal interactors per enzyme was now significantly reduced – for instance, for the case of ACO2 from the initial 267 proximal interactors, 17 were now included in the final list (Figure 11a). A list of all higher specificity interactors is provided in appendix, in Table 6. This analysis allowed for more specific features to appear. For instance, while IDH3A and IDH3B, known interactors of IDH2 and partners of IDH3G, were previously included in the proximal interactomes of all mitochondrial enzymes as part of the shared immediate neighborhood, now were highlighted for the cases of IDH2 and IDH3G only. With this analysis, the subcellular topological distribution highlighted further the purely nuclear fraction present in OGDH, IDH3G, IDH1 and now more clearly for the case of ACO2 (Figure 11b). Regarding ACO2, out of the 17 identified proximal interactors 3 were nuclear proteins, in contrast to the zero of the negative control IDH2. The respective proteins were the RBMX, HMGA1 and TOP2B - these were also detected amongst the significant interactors of OGDH and IDH3G but not of IDH1 (Figure 11a and 11d). All three proteins have been connected to chromatin binding and remodeling[115–117], while RBMX contains an RNA-binding domain too and has been involved in pre-mRNA processing[118]. Interestingly, both TOP2B and HMGA1 have been associated to development, cellular differentiation and malignancy[119–121]. As already shown before, OGDH and IDH3G had the highest percentage of nuclear proteins, sharing many overlapping proximal interactors ranging from 53% to 78%, respectively, which was hence reflected into many shared enriched GO terms too. For both enzymes, for the nuclear fraction, the most enriched biological feature

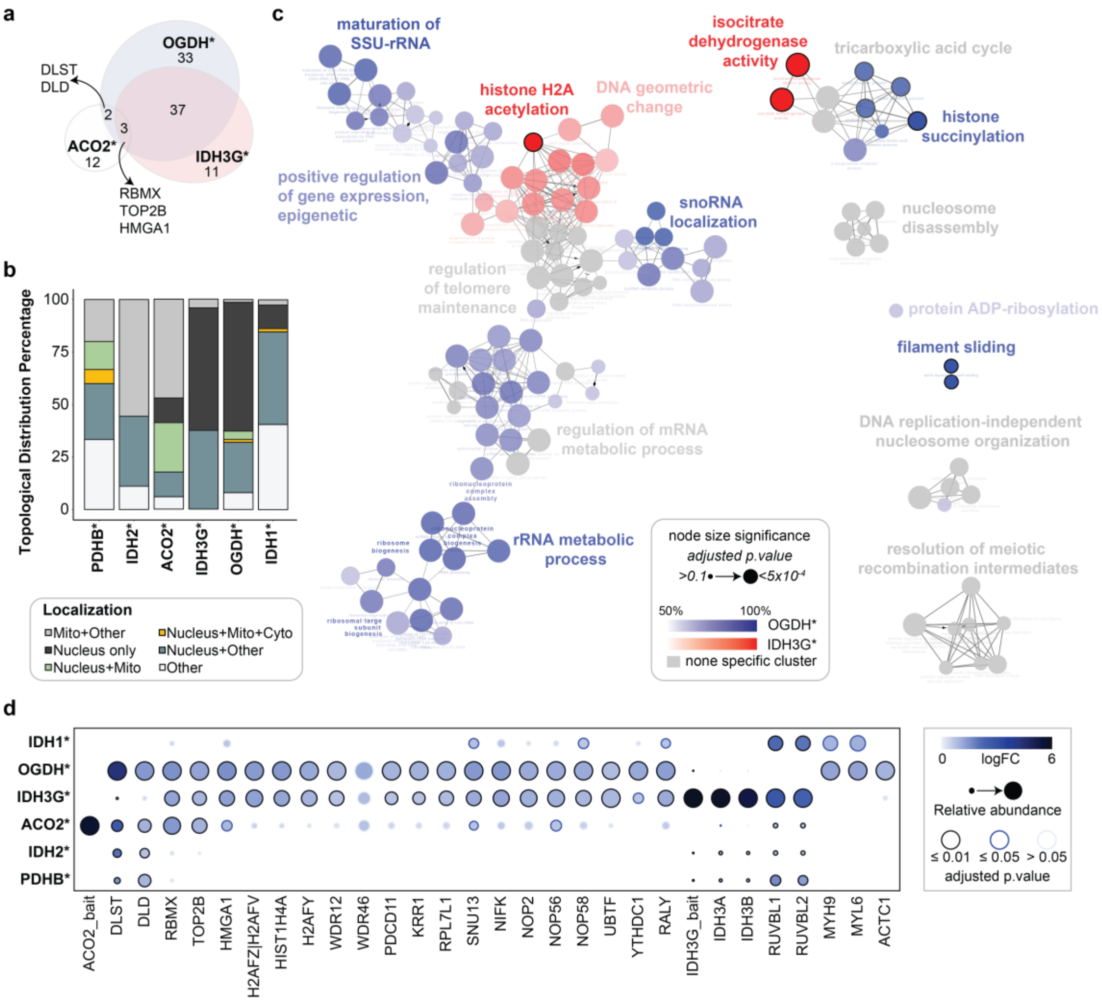

Figure 11: Proximity biotinylation mass spectrometry reveals a potential nuclear niche for ACO2, OGDH and IDH3G coupled to histones, chromatin binding and RNA modifications. (a) and (b), Venn diagram and topological distribution of the higher specificity proximal interactors, respectively. For the topological annotation see details of Figure 10. For a detailed list of the higher specificity proximal interactors refer to supplementary material, Table 5. (c) GO Biological Process Network based on the higher specificity proximal interactors of the C-fused IDH3G and OGDH. Each node represents a GO enrichment “Biological Process” term and each group of nodes is described by a representative term (bold letters). The size of the node corresponds to the significance level (Benjamin-Hochberg adjusted p values). The color and color gradient indicates the specificity of the GO term for each engineered cell. GO terms equally enriched in both cell lines are colored grey. GO terms uniquely present in a cell line are marked with a black-lined circle. The GO enrichment analysis was performed with ClueGO122 and visualized with Cytoscape[123]. (d) Dotplot representation of selected proximity interactors. Histones and rRNA metabolic process related proteins are highlighted. Node color gradient represents the $\log_2$ fold change of the protein in each engineered cell line compared to wild type HeLa. The size of the node indicates the relative abundance of the above log2 fold changes amongst the different engineered cells. The color of the line around the node is indicative of the significance (Benjamin-Hochberg adjusted p values). Dotplot was generated using the ProHits Visualization web-tool[124].

was the RNA processing (Figure 11c), albeit the enrichment was 2-fold higher for the case of OGDH. Out of the shared enriched GO terms, while the mRNA metabolic process presented equal enrichment for both enzymes, the terms that drove the higher “RNA processing” enrichment in OGDH were attributed to ncRNA metabolic processes and more specifically to those of rRNA and tRNA. Finally, we additionally identified two processes that were rather specific to each enzyme and absent from the other. Both processes were connected to histone modifications, which, for the case of OGDH was histone succinylation, and for IDH3G histone acetylation (Figure 11c). While the first one was connected to the activity of the enzyme which produces succinyl-CoA that can be used for the succinylation of histones and generally of other proteins too, the latter was attributed to the activities of proteins present in the proximal interactome of IDH3G, namely the RUVBL1 and RUVBL2. Both are components of the NuA4 histone acetyltransferase complex involved in the acetylation of histones H4 and H2A[125].

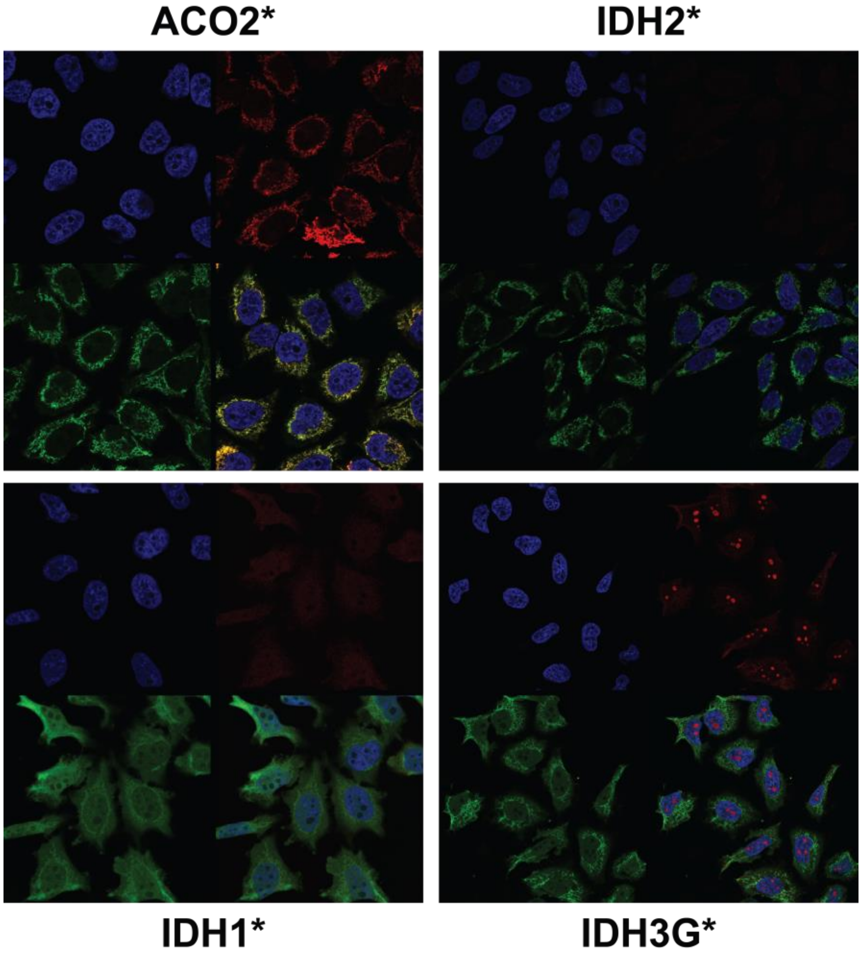

Figure 12: Immunofluorescence microscopy of the engineered cell lines expressing the protein of interest fused to biotin ligase 2. The following stable cell lines are depicted: ACO2*, IDH2*, IDH3G* and IDH1*. Each image has four panels corresponding to DNA labeled with Hoechst dye (blue), anti-flag (red) (for IDH1* and IDH2*) or anti-native protein (red) (anti-ACO2 for ACO2*, anti-IDH3G for IDH3G*), biotinylated proteins labelled with Streptavidin-Alexa 488 (green) and merged. An asterisk at the beginning or the end of the protein name denotes the fusion of biotin ligase at the N-terminus or C-terminus, respectively.

As a final step, we aimed to monitor the localization of the engineered enzymes and the biotinylated proteins using immunofluorescence microscopy based on anti-FLAG and streptavidin-staining, respectively (Figure 12 and 13). Anti-FLAG did not give an immunofluorescence signal, contrary to immunoblotting results, and hence we utilized antibodies targeting the endogenous proteins. Possible explanations for the lack of signal are the general low expression of the engineered enzymes, and/or the inaccessibility of the site when the whole complexes are formed for the cases of OGDH and IDH3G. For ACO2 (Figure 12) and OGDH (Figure 13), both the engineered/endogenous enzymes and the biotinylated proteins were observed predominantly in the mitochondria.

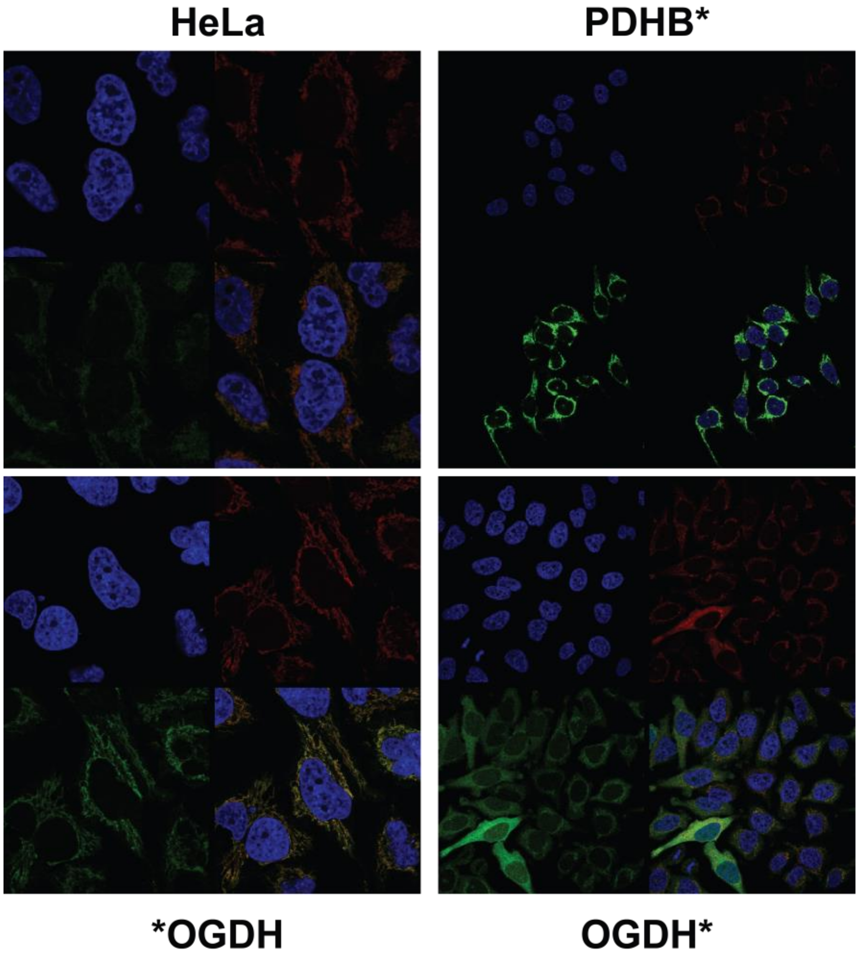

Figure 13: Immunofluorescence microscopy of the engineered cell lines expressing the protein of interest fused to biotin ligase 2. The following cell lines are depicted: wild type HeLa, PDHB*, OGDH* and *OGDH. Each image has four panels corresponding to DNA labeled with Hoechst dye (blue), anti-flag (red) (for PDHB* and HeLa) or anti-native (red) (anti-OGDH for OGDH* and *OGDH), biotinylated proteins labelled with Streptavidin-Alexa 488 (green) and merged. An asterisk at the beginning or the end of the protein name denotes the fusion of biotin ligase at the N-terminus or C-terminus, respectively.

Although the subcellular annotations of the proximal proteins included a fraction characterized as of purely nuclear topology, we were not able to quantify a detectable signal in the nuclear compartment. We speculate that this is likely due to the uneven subcellular distribution of the enzymes with a primarily localization in the mitochondria and a relative small proportion interacting with nuclear proteins as shown by the mass spectrometry results. Regarding IDH3G, for the enzyme itself we detected a very clear presence in the nucleolus, while the biotinylated proteins displayed a more diffused pattern throughout the nucleus and mitochondria (Figure 12). Finally, IDH1 proximal biotinylated partners showed a mainly cytoplasmic distribution (Figure 13), while the negative control IDH2 and PDHB appeared in the mitochondria (Figure 12 and 13).

Summary and Outlook

Collectively, our [U-^{13}C]-labeling data provided strong and direct evidence that a functional metabolic pathway representing part of the TCA cycle operates in the HeLa nucleus. While existent studies have provided so far the nuclear activity of specific enzymes under specific contexts, we identified citrate and glutamine as two important feeding points in the nucleus reconstructing an entire active pathway supplementing intermediates having key roles in epigenetic modifications. For instance, alpha-ketoglutarate is a necessary cofactor for a specific class of demethylases, succinyl-CoA is the succinyl-donor for histones succinylation, while other downstream metabolites like succinate and fumarate act as demethylase inhibitors. A nuclear generation of aspartate could potential be attributed to its requirement for purine and pyrimidine synthesis, being close to known nucleotide-related enzymes, like CAD, whose nuclear presence increases during S phase when the demands for nucleotides are maximal[126,127]. The metabolic pathway was further supported by the presence of the respective enzymes in literature available nuclear proteomics datasets, as well as by the identified NLS in their protein sequences. Although with the current study we cannot exclude the nuclear topology of these enzymes owing to distinct, moonlighting activities, and hence the generation of these metabolites as a lateral effect, the apparent connection of these metabolites to histone, DNA or RNA modification processes couples their nuclear biosynthesis to a more efficient architecture by being in close proximity to sites of need.

Proximity labelling mass spectrometry identified a nuclear niche for OGDH and IDH3G, and potentially also for ACO2. For ACO2, although the number of nuclear interactors was limited to three, these were included amongst the higher specificity and preferential proximal interactome. In addition, several nuclear proteins detected in OGDH and IDH3G had a minor presence in ACO2, albeit not significant, which could point to a condition-specific enrichment of ACO2 in the nuclear compartment (Figure 11d). Further exploration of their proximal interactome revealed projections to chromatin, histones and RNA metabolic processes. A nuclear presence could be connected to a local generation of succinyl-CoA by OGDH further used for the succinylation of histones. While our work was in progress, one additional study was published showing that KAT2A acetyltransferase interacts with OGDH and utilizes the generated succinyl-CoA for the succinylation of histone H3[34]. Although in our dataset we did not recapitulate the OGDH-KAT2A-H3 axis, we identified connections to histone H4 and the histone H2A variants H2AFZ or H2AFV and H2AFY, denoting a similar role (Figure 11d). The same study quantified that only 1–1.6% of total OGDH, DLST, and DLD is localized in the nucleus, further supporting the absence of a clear nuclear staining in our immunofluorescence results too. Finally, the presented enrichment in RNA processes is a more complex and intriguing one. This relationship could be attributed to the fact that several metabolic enzymes, including TCA cycle enzymes, have been characterized as RNA binding proteins, with this characteristic being currently of unknown biological role[128–130]. Apart from or conjointly to this, a connection might be drawn to RNA modifications per se, where, similarly to DNA and histones, specific metabolites are utilized as substrates or co-factors for their formation. Although our investigated pathway did not include enzymes that generate metabolites directly employed for these RNA modifications (i.e. SAM), it was coupled to alpha-ketoglutarate production, a necessary co-factor for specific classes of DNA and histone demethylases, few of which have been lately discovered for mRNA and tRNA, though not rRNA yet[20]. All the aforementioned provide just the basis of some initial functionally relevant connotations which however render of advanced systematic characterization and further validation. Overall, the proximal interaction dataset provides a resource for furthering the exploration and organize follow up experiments laying at the conjunction of metabolism and its gene expression regulation pleiotropic roles.

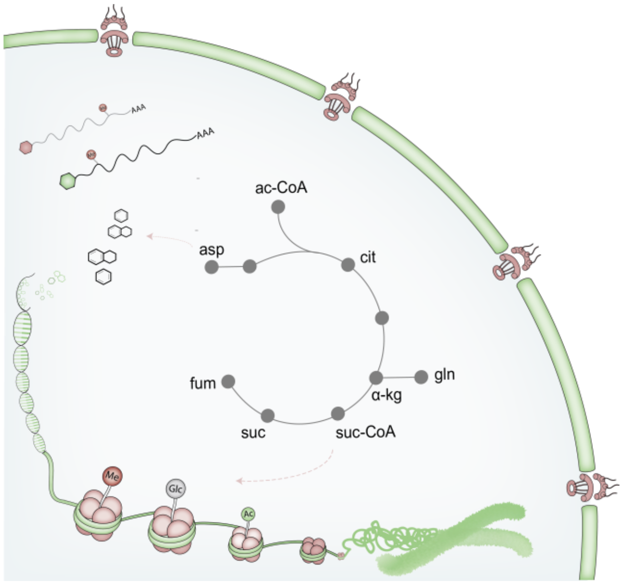

Figure 14: Proposed nuclear metabolic network and relevant functions.

Contributions

- Eleni Kafkia (Structural and Computational Biology Unit, EMBL Heidelberg, Germany) was responsible for all experiments, design, data analysis and interpretation
- Paola Jouhten (VTT, Technical Research Center of Finland) and Christoniki Manga-Nteve (Structural and Computational Biology Unit, EMBL Heidelberg, Germany) helped with the natural abundance correction of the labeling samples
- Markus Seiler (Buchman Institute for Molecular Sciences, Goethe University Frankfurt) performed NLS motifs analysis
- Amparo Andres Pons (Friedrich Miescher Institute for Biomedical Research, Basel), conducted the immunofluorescence and immunoblots of the isolated nuclei, and helped at initial protocol set-up of the nuclei isolation experiments
- Electron microscopy of the isolated nuclei was performed at the Electron Microscopy Core Facility, EMBL, Heidelberg, by Paolo Ronchi
- Proximity biotinylation samples were analyzed at the Proteomics Core Facility, EMBL, Heidelberg by Mandy Rettel, and initial data analysis was performed by Frank Stein
- Kiran Raosaheb Patil (Structural and Computational Biology Unit, EMBL Heidelberg, Germany) supervised the project and designed experiment

Chapter II: Metabolic particularities of minimal residual disease in breast cancer

Metabolic reprogramming is fundamental to the robustness of cancer cells. These enable them to adapt and support their growth and proliferation in the face of hindering environments and genetic mutations. Although these altered metabolic features have been extensively studied in tumors, much less is known about the metabolic states that characterize residual, dormant cancer cells which escape treatment and contribute to cancer recurrence and therapy resistance (known as minimal residual disease). Understanding the metabolic pathophysiology of this population is important both for their effective diagnosis and for identifying pathways that could be limiting for their growth. A recent study from Havas and colleagues[131], showed that regressed mouse breast cancer cells, although phenotypically similar to their normal counterparts, possessed a distinctive transcriptional profile which pointed to metabolic alterations. Following this lead, the current thesis contributed to the comprehensive characterization of the metabolic physiology of these residual cells. To achieve this, metabolomics analyses were employed to complement transcriptomics and phenotypic data in three-dimensional cell cultures of primary epithelial cells isolated from mammary glands of the breast cancer mouse model. The identified key metabolic idiosyncrasies were further verified in the *in vivo* mouse model.

Tumor and residual cells exhibit divergent response at the transcriptomic and metabolic level

Considering the inherent challenges involved in the recognition and isolation of residual, dormant cancer cells *in vivo*, animal models correlative to human cancers offer an advantageous alternative system. The system employed in the current study was based on a mouse model of breast cancer, wherein the cancer formation can be initiated with the doxycycline-inducible and mammary-specific overexpression of two oncogenes, Myc and Neu. Myc oncogene is mutated human Myc, and Neu is coming from rat and is equivalent of the human epidermal growth factor receptor 2 (HER2). Increased MYC and HER2 expression is frequently observed in breast cancers, with HER2 overexpression occurring in approximately 25-30% of the known cases[132]. Once oncogene overexpression is seized (i.e. upon doxycycline withdrawal), tumor regresses and mammary glands are phenotypically similar to their normal counterparts. However, spontaneous tumors may reoccur in these mice, representing thus a system that simulates the kinetics of tumor progression, regression and relapse. This *in vivo* cancer kinetics can be recapitulated *in vitro*, in three-dimensional cell cultures of primary epithelial cells isolated from the mammary glands of the breast cancer mouse model[131]. When the cells are seeded in a basement membrane extract (Matrigel) mixed with collagen are able to form three-dimensional structures that resemble the respective ones found *in vivo* and which are named acini. The breast acini are units built from single layer of polarized epithelial cells localized around the lumen of the duct and surrounded by an outer layer of myoepithelial cells and the basement membrane. When doxycycline is added to the growth medium, oncogene overexpression is induced resulting in the formation of tumors represented as filled acini. Upon doxycycline withdrawal, the majority of the cells undergo apoptosis, which results in the lumen clearing out and structures regressing and organizing back to a single-layered rim of polarized epithelial cells. These surviving cells represent the cancer residual population (Figure 15).

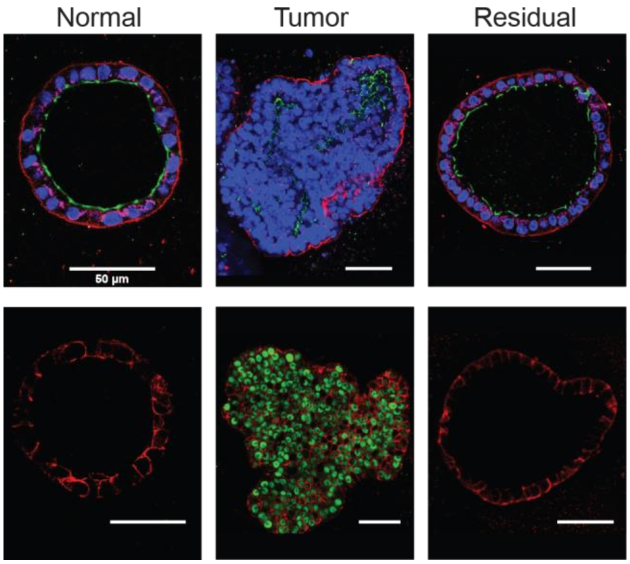

Figure 15: Residual cells are phenotypically similar to their normal counterparts. Upper panel: ITGA6 (red), ZO-1 (green), GM-130 (magenta). Lower panel: human MYC oncogene (green); E-cadherin (CDH1, in red) as a counterstain. From left to right: normal, tumor and residual structures. Scale bar: 50 µm. (figure adapted from adapted from Ksenija Radić Shechter, PhD thesis).

To examine the cancer residual population, the three-dimensional cell cultures were collected for immunofluorescence, RNA-sequencing, lipidomics and metabolomics in the following states: a) cells where oncogenes overexpression was absent (hereafter denoted as normal), b) cells treated with doxycycline for five days and thus overexpressing the oncogenes (hereafter denoted as tumors), and c) cells treated with doxycycline for five days and overexpressing the oncogenes, followed by doxycycline withdrawal for seven days and thus absence of oncogenes overexpression (hereafter denoted as regressed cells). For the analysis at the metabolic level, apart from the intracellular samples, extracellular spent growth medium was also collected prior to cell harvesting, corresponding to the extracellular metabolome of the last 24h of the cell cultures. For the metabolomics analyses, two differential methods were utilized. Untargeted analysis was performed using direct injection in a hybrid quadrupole-orbitrap mass spectrometer (Q Exactive MS)[133]. For targeted analysis, Gas

Chromatography - Mass Spectrometry (triple quadrupole) (GC-MS) was utilized covering central carbon metabolic pathways. For the metabolomics analyses, one additional control was included where non-inducible wild type cells were treated with doxycycline (hereafter denoted as wild type) to account for any potential metabolic effects attributed solely to the doxycycline treatment.

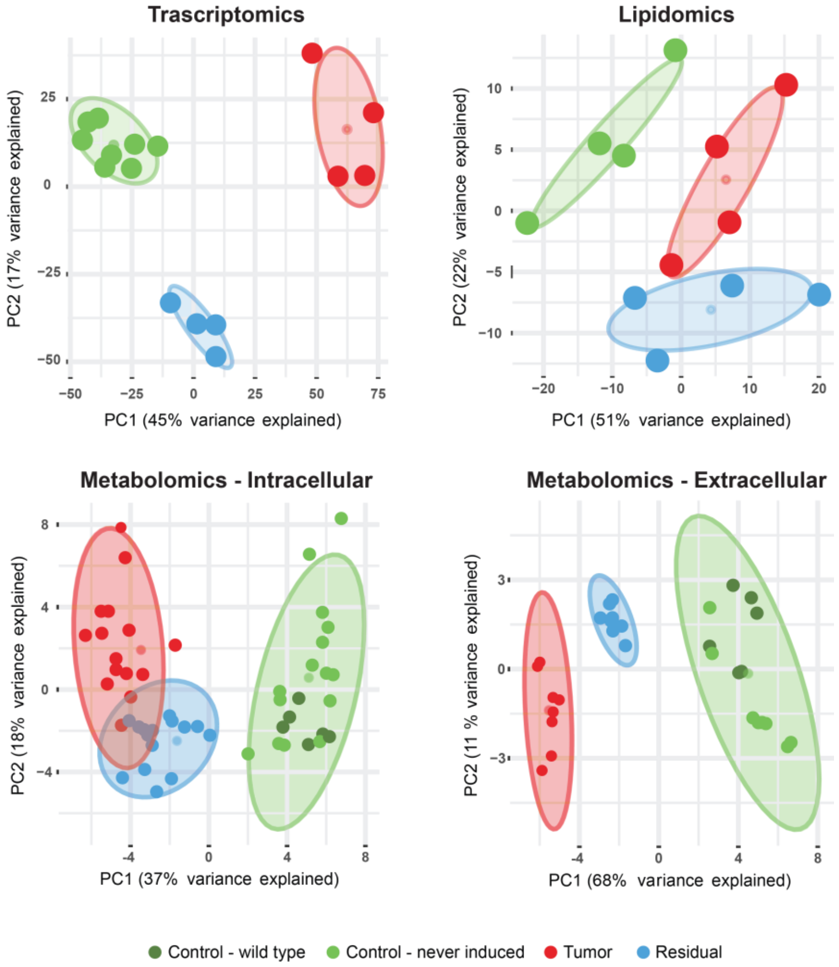

Figure 16: Tumor and residual cells exhibit divergent response at the transcriptional and metabolic level. Principle component analysis graphs of the transcriptomics, lipidomics and metabolomics (intracellular and extracellular) profiles for the control (never induced and wild type), tumor and residual populations. Eclipses represent the confidence interval of a multivariate normal distribution (confidence intervals 0.95).

The immunofluorescence imaging highlighted the phenotypic similarity among the normal and regressed cells, and verified the absence of the MYC oncoprotein in these two states compared to the highly increased levels in tumor cells (Figure 15) (Ksenija Radić Shechter Shechter, PhD thesis). The expected expression levels of both oncogenes, MYC and HER2, were additionally confirmed at the transcript level for all three states (Ksenija Radić Shechter Shechter and Katharina Zirngibl, PhD theses). The RNA-sequencing derived transcriptome profiles exhibited a clear separation between the three states, with the regressed cells being closer to the normal than to the tumor, as indicated by the principle component analysis (PCA) graph (Figure 16). This result further confirmed the initial study from Havas and colleagues, where transcriptome differences based on microarray analysis were also observed for all three states[131]. Notably, despite the transcriptome profile of the regressed cells being more close to the normal population, the distances at the metabolic level amongst the three states exhibited a differential pattern. As defined by both intracellular untargeted and targeted metabolomics analyses, the metabolic profile of the regressed cells was more similar to the tumorous population than to their normal counterparts. This is reflected in the principle component analysis of the intracellular targeted metabolomics results (Figure 16), based on the quantification of 53 polar metabolites covering central carbon metabolism, as well as in the hierarchical clustering of the untargeted metabolomics results based on 2832 detected ions (Figure 17). A similar pattern was also observed at the extracellular targeted metabolic profiles (based on 36 quantified metabolites), albeit with a more clear separation of the residual from the tumor state (Figure 16), as well as at the level of lipids (Figure 16). To assess whether the distances between the three states at the transcriptomic level could resemble the ones present at the metabolic level if solely the gene expression of metabolic enzymes is included, principle component analysis of the RNA-sequencing data was performed excluding all non-metabolic genes. Despite this, the principle component analysis did not result in changes in the transcriptional distances (Ksenija Radić Shechter and Katharina Zirngibl, PhD theses), suggesting that the higher metabolic similarity between the residual and tumor cells points towards the existence of a tumorous metabolic memory preserved in the residual population.

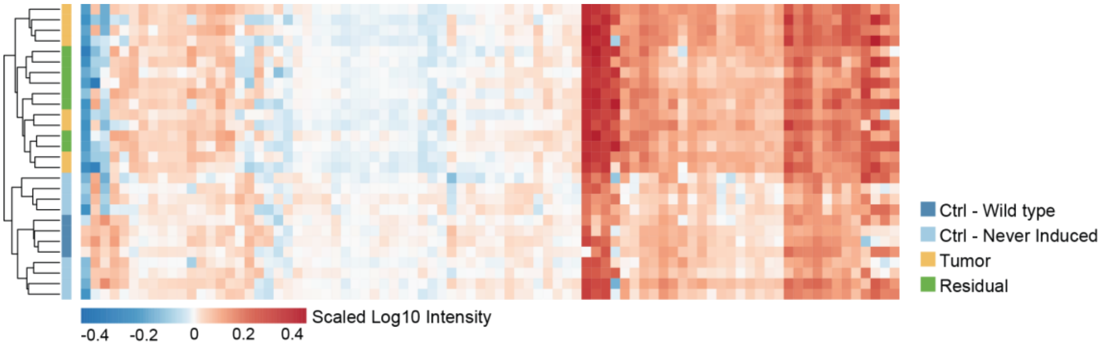

Figure 17: The metabolic profile of residual cells is more similar to the tumor than the control state. Hierarchical clustering of control (never induced and wild type), tumor and residual populations based on the untargeted metabolomics method for intracellular samples (unsupervised hierarchical using the Manhattan distance) (figure adapted from adapted from Katharina Zirngibl and Ksenija Radić Shechter, PhD theses).

The metabolic alterations of the residual population are mnemonic of the tumorous state

To gain a deeper insight of the metabolic physiology that characterizes the residual as well as the tumor cells, we next examined the intracellular and extracellular metabolic changes identified in these two populations when compared to the normal state. The untargeted metabolomics analysis provided preliminary information regarding altered metabolic pathways, however the multiple annotations per ion renders the identification of metabolites challenging. For this reason, we utilized the targeted metabolomics analyses owing to the precise identification of the detected metabolites which covered multiple central carbon metabolic pathways, including glycolysis along with branching pathways, TCA cycle, glutamine metabolism, urea cycle, amino acids and several free fatty acids. An overview of the quantified metabolites in intracellular samples is provided in Figure 18. The hierarchical clustering based on the metabolite concentration pools revealed once again the higher similarity between the residual and tumor cells, as well as that doxycycline treatment did not induce overall changes in the metabolic physiology of the control cells. When compared to the control state, the tumor population exhibited changes in the concentration of several metabolites connected to metabolic pathways frequently deregulated in cancers characterized by overexpression of the oncoproteins MYC and HER2. Notably, the residual population, despite the absence of the overexpression of the two oncogenes, retained similar changes in their metabolic profiles, with differential magnitude in different pathways. An overview of all significant alterations in the intracellular and extracellular samples is provided in the appendix section, Figure 32. The significantly detected changes are discussed below per pathway, in parallel for the tumor and regressed populations.

Of the most prominent metabolic alterations was the glucose intracellular and extracellular pool sizes which were largely decreased in both tumor and residual cells when compared to the normal state (Figures 19-21). Pyruvate, the end glycolytic metabolite which can be converted to lactate, alanine or acetyl-CoA, was also reduced intracellularly and extracellularly both in tumor and regressed cells. On the other hand, lactate and alanine extracellular pools were significantly increased in both states. The reduced glucose concentration levels combined with the significantly increased levels of lactate in the extracellular spent growth medium suggested a higher glycolytic flux present in both populations. Interestingly, lactate concentration levels in the intracellular environment were significantly higher only in the case of the residual population. Additionally, the reduced pyruvate levels compared to the concentration by which pyruvate was detected in pure growth media (i.e. growth medium prior to its addition to the cell cultures), were also suggestive of an increased pyruvate consumption in both states. Overall, the aforementioned results pointed towards an enhanced glycolytic phenotype which was present in tumor and residual cells.

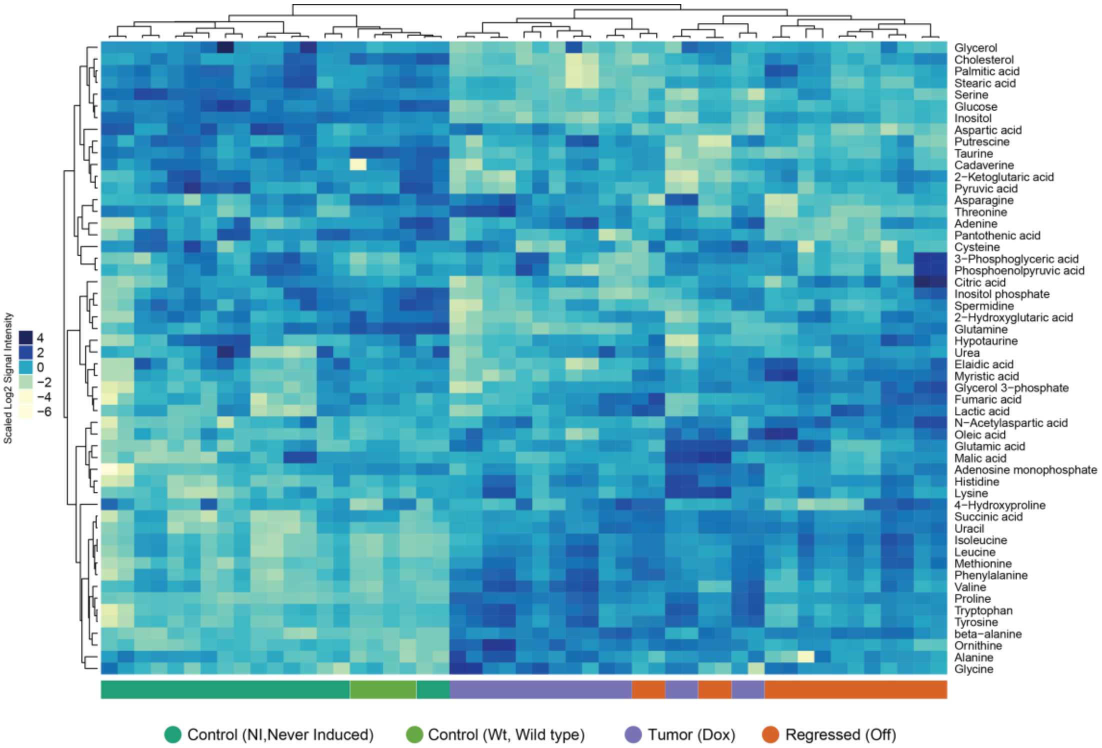

Figure 18: The metabolic profile of residual cells is more similar to the tumor than the control state. Heatmap of all intracellular metabolites quantified with the targeted metabolomics method (GC-MS) for the control (never induced and wild type), tumor and residual populations (hierarchical bi-clustering, pearson correlation).

Glutamine is an important amino acid involved in several metabolic pathways, including TCA cycle anaplerosis, nucleotide metabolism, glutathione generation and redox homeostasis[49]. Glutamine concentration pools were significantly decreased in tumor cells, both intracellularly and extracellularly, indicating thus an increase in glutamine catabolism (Figure 19-21). Residual cells displayed a decrease in extracellular glutamine too, with the concentration pools exhibiting a tendency towards restoration by approximating the levels found in normal cells. Glutamine can generate glutamate whose intracellular abundance was only slightly increased and to a similar extent in both tumor and residual cells. Subsequently, glutamate is coupled to multiple pathways[49], including entering the TCA cycle after conversion to alpha-ketoglutarate. In addition, glutamate is involved in the biosynthesis of proline, which was significantly increased in both populations, intracellularly and extracellularly. Collectively, these results highlighted common alterations in glutamine and proline metabolism in tumor and residual cells, albeit with varying degrees of magnitude in the residual population for the different sub-pathways.

Since glucose and glutamine are major carbon sources contributing to TCA cycle, and a deregulation in their pathways was observed for tumor and residual cells, we next examined the concentration levels of TCA cycle intermediates. Alpha-ketoglutarate and 2-hydroxyglutarate were significantly decreased in the tumor samples, while no significant changes were detected for the residual cells. However, the residual population exhibited a significant increase for the metabolites further downstream of alpha-ketoglutarate, namely succinate and fumarate (Figure 19-21).

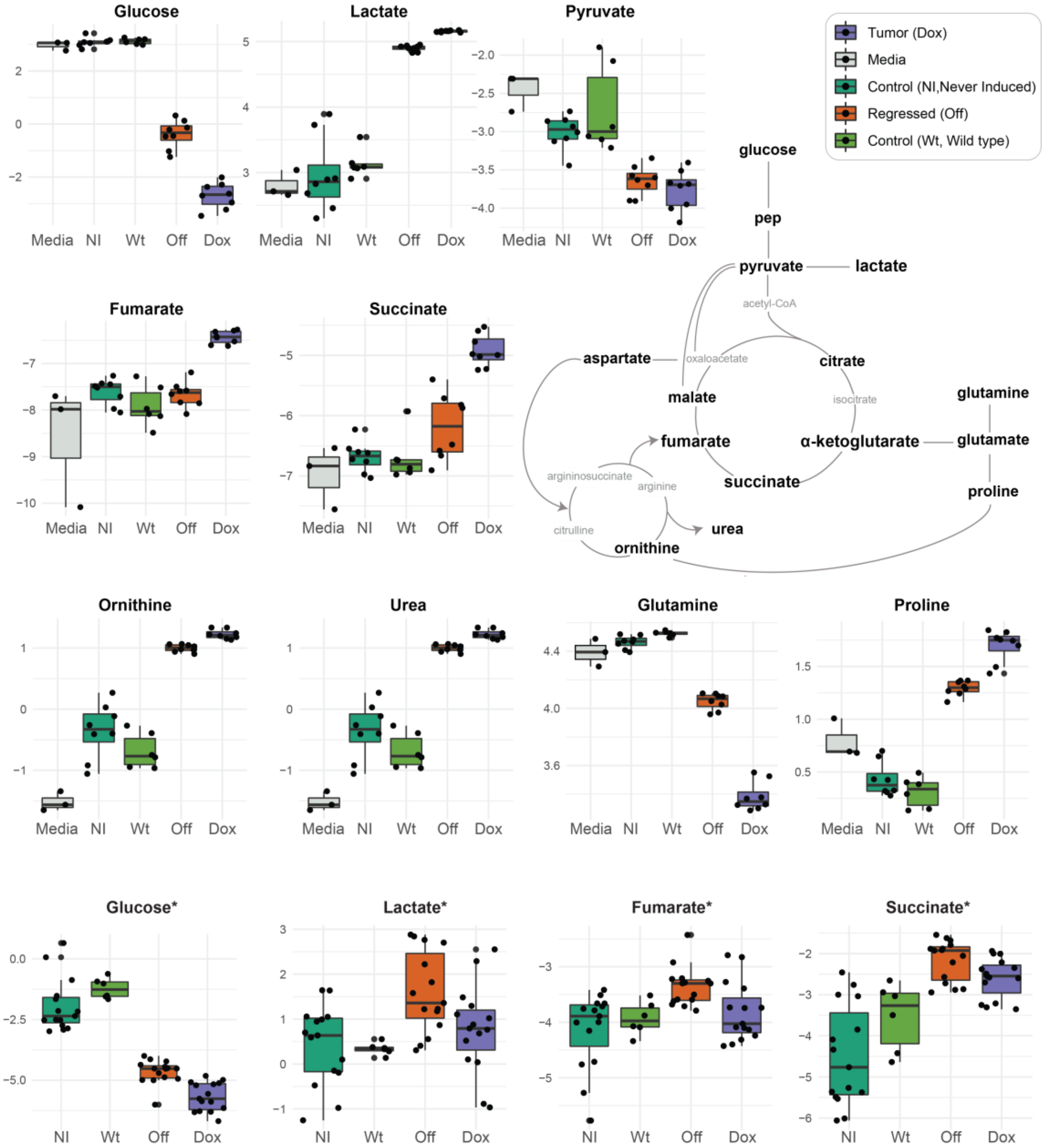

Figure 19: Metabolic alterations in the residual and tumor population. Boxplots showing selected cases of significantly altered metabolites (limma analysis, Benjamini-Hochberg adjusted p value ≤ 0.01, see methods) in the spent growth medium of control (never induced and wild type), tumor and residual cells. The concentrations of the respective metabolites in the growth media prior addition to the cell cultures (denoted as media) are also depicted. Marked with an asterisk are selected significant metabolites from the intracellular samples. The concentrations are measured as signal intensities (area under the curve) and are depicted in the $\log_2$ scale.

Malate was also increased, albeit the change was statistically less powerful. Tumor cells displayed increase only for succinate, with the relative change being lower than the observed one in the residual population (Figure 21 and 22). However, both succinate and fumarate were detected in the spent growth medium of tumor cells with significantly increased levels (Figure 20). Succinate can be released from the cells, although the process is poorly understood so far[134]. Concomitantly, succinate's intracellular accumulation can derive from multiple pathways, which apart from an altered TCA cycle activity, include the catabolism of the branched chain amino acids isoleucine and valine, the glutamine-derived and TCA cycle by-passing GABA shunt pathway, or from fumarate produced from aspartate catabolism, urea cycle and purine nucleotide cycle through a reversal of the succinate dehydrogenase activity (SDH)[135]. Overall, we detected changes in TCA cycle with differential behaviors for specific metabolic intermediates between the two populations. Interestingly, we observed a preferential accumulation of succinate and fumarate in the intracellular space of residual cells, while for the case of tumor cells both metabolites were released in the extracellular environment.

Apart from the aforementioned alterations, we additionally identified changes in metabolites and metabolic pathways connected to TCA cycle, namely aspartate and urea cycle (Figures 20 and 21). Mammalian cells can uptake aspartate from the extracellular environment, however the process is often inefficient[136]. As a result, aspartate is generated from oxaloacetate which is connected either to TCA cycle[62,63], or in specific contexts to glycolysis[80]. Aspartate was significantly decreased in both residual and tumor populations, intracellularly and extracellularly. This decrease, compared to aspartate's levels in the pure media samples and in the spent growth medium of normal cells, was suggestive of an increased catabolism. Aspartate is coupled to nucleotides, as it is an essential nitrogen and carbon source for their biosynthesis[100]. For both populations, we observed a significant increase in the intracellular concentration pools of metabolites linked to nucleotides, namely uracil and beta-alanine (Figure 20). In addition, aspartate is also linked to urea cycle, where argininosuccinate synthase (ASS1) utilizes aspartate and citrulline to synthesize argininosuccinate[137]. We found urea and ornithine to be the second most increased metabolites after lactate in the extracellular medium of both tumor and regressed cells, indicative of an enhanced urea cycle (Figure 19). From the intracellular samples, ornithine was significantly increased in the tumor population, and a minor, albeit significant, increase in the regressed cells was also observed. Ornithine in turn is involved in the generation of polyamines[138] and the biosynthesis of proline[139], which, as described above, was also significantly increased in tumor and residual population (Figures 19-21). Finally, urea cycle generates fumarate[139] that can further fuel the TCA cycle. Overall, the aforementioned results described significant alterations along the urea pathway interconnected to alterations in aspartate, TCA cycle and proline metabolism (Figure 19-21).

To conclude, additional affected pathways involved free fatty acids (i.e. cholesterol, stearate, palmitate) which were significantly decreased in the case of tumor cells (Figure 23). For the residual population, the concentration pools of the aforementioned fatty acids approximated the levels detected in normal cells. An exception was the myristic acid, which, although unaltered in tumor cells, exhibited a two-fold increase in the case of residual cells. Glycerol-3-phosphate was another metabolite specifically increased only in residual cells. Alterations in the concentration pool of this metabolite could be connected to glycerolipid metabolism, as well as to glycerol-3-phosphate shuttle utilized for electron transport between the cytoplasm and the mitochondria, coupling thus glycolysis and oxidative phosphorylation. Finally, other significant changes observed in both states included increased histidine, taurine and N-acetyl-aspartate levels, as well as in amino acids, with the respective changes following varying degrees of magnitude between the two populations (Figures 20 and 21).

To understand whether the above metabolic alterations were reflected in the expression of the respective genes, the transcript levels, as defined by the RNA-sequencing analysis, were examined. Utilizing the same transcriptional data, reporter metabolite analysis was performed to predict which metabolites were likely to change as a further verification to the metabolomics data, as well as to capture potential changes in metabolites that were not included in the targeted method (Katharina Zirngibl and Ksenija Radić Shechter, PhD theses). Moreover, flux predictions based on a genome-scale reconstructed mouse model were also estimated (Katharina Zirngibl and Ksenija Radić Shechter, PhD theses). An overview of all the acquired levels of information and predictions is presented in Figures 21 and 22 for the tumor and residual population, respectively. Focusing on the residual state, the transcript levels of glucose transporters as well as the majority of the glycolytic enzymes were increased compared to the normal state. In addition, specific enzymes in the urea cycle were also elevated, as well as the mitochondrial transporter SLC25A13 that facilitates the exchange of aspartate and glutamate connected to the urea cycle. Subsequently, the predictions suggested alterations in several glycolytic intermediates, the urea and nitric oxide metabolites, pyrimidine nucleotides and in the pentose phosphate pathway. Finally, flux changes were predicted for glycolysis (increase) and two reactions in the TCA cycle (decrease). For the tumor state, similar to the residual cells alterations were observed and predicted. Collectively, the transcriptional changes and predictions were in agreement with the metabolic phenotype of the residual and tumor populations.

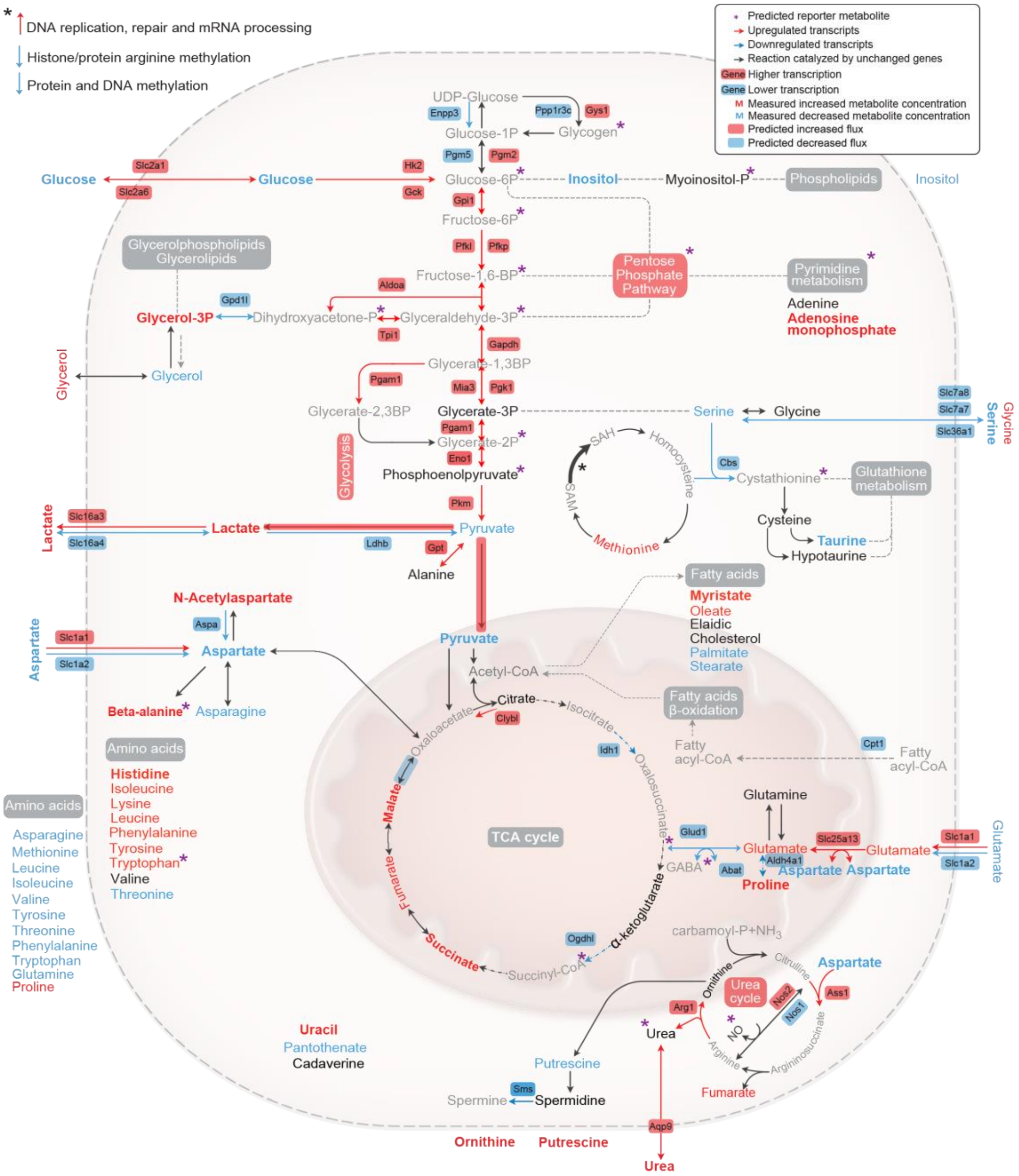

Figure 20: Global overview of central carbon metabolic alterations of the residual population compared to the control state. The figure encompasses all available information acquired from targeted metabolomics measurements, transcriptional changes of the metabolic enzymes, reporter metabolite analysis based on the transcriptional data, and flux predictions based on a genome-scale reconstructed mouse model (Ksenija Radić Shechter, PhD thesis).

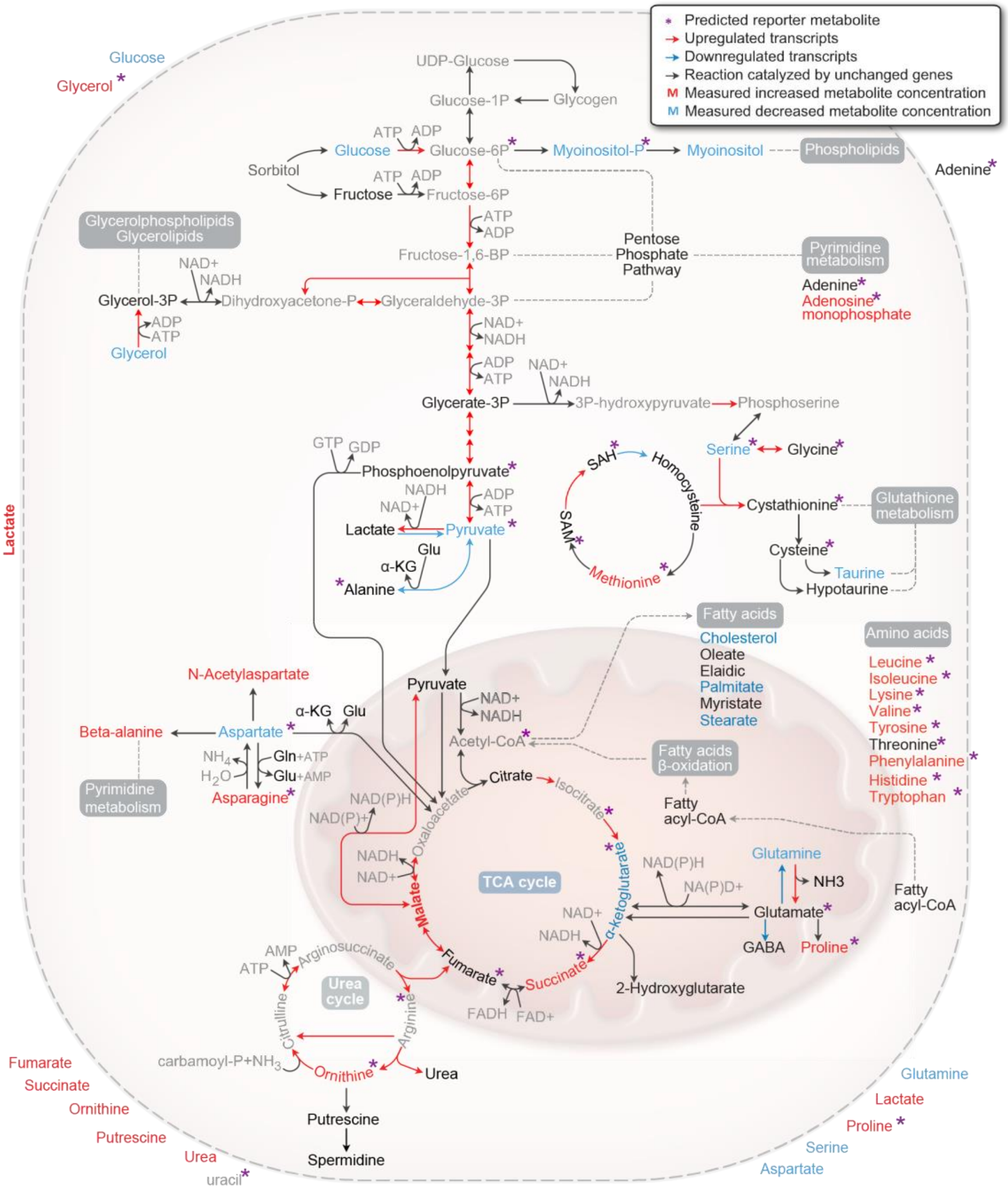

Figure 21: Global overview of central carbon metabolic alterations of the tumorous population compared to the control state. The figure encompasses all available information acquired from targeted metabolomics measurements, transcriptional changes of the metabolic enzymes, reporter metabolite analysis based on the transcriptional data, and flux predictions based on a genome-scale reconstructed mouse model (Ksenija Radić Shechter, PhD thesis).

Taken together, the targeted metabolomics data highlighted that the residual cells, despite the phenotypic similarity to their normal counterparts, preserved specific metabolic alterations reminiscent of the tumorous population. This metabolic memory was also reflected at the transcriptional level of the corresponding enzymes for the majority of the reactions. The metabolic particularities that characterized the residual population pointed towards an enhanced glycolysis and urea cycle, alterations in proline metabolism and accumulation of specific TCA cycle intermediates.

Ex vivo **and** ***in vivo*** **validations support the altered glycolysis and urea cycle metabolic signatures in the residual population**

We next aimed to validate in the mouse model system the key metabolic alterations observed in the *in vitro* three dimensional cell cultures. For this purpose, animals corresponding to the regressed state (i.e. mice treated with doxycycline for four to six weeks to induce tumor formation followed by doxycycline withdrawal for nine weeks for tumor regression), as well as age-matched wild type animals (i.e. wild type mice treated with doxycycline followed by doxycycline withdrawal for the same period) were utilized to isolate the mammary glands. The isolated tissues were cultivated in the presence of uniformly labeled glucose, [U-^{13}C]-glucose, for 8 hours, and intracellular and extracellular samples were collected to measure ^{13}C-lactate levels. In parallel, the same experiment was performed with unlabeled glucose for extracellular targeted metabolomics analysis.

From the *ex vivo* labeling experiment, we detected statistically higher intracellular ^{13}C-lactate pools in the regressed mammary glands compared to the control ones (Figure 22d). In the *ex vivo* extracellular metabolomics experiment, the regressed mammary glands when compared to the control ones, showed a trend for decrease in glucose levels and an increase in lactate levels, albeit not statistically significant (Figure 22b). Overall, the labeling and extracellular metabolomics results combined suggested that the relative glucose contribution to lactate was higher in the regressed cells compared to the control state.

The extracellular metabolic profile of the regressed mammary glands highlighted once again their distinctive metabolic nature from the control state (Figure 22a), and verified the previously defined alterations in the urea cycle as indicated by the higher extracellular abundances in urea, ornithine and putrescine (Figure 22b). Pyruvate extracellular levels exhibited the same alteration as in the three dimensional cell cultures, being significantly lower in the regressed state compared to the control (Figure 22b). Aspartate extracellular pools displayed an opposite alteration than the one observed in the cell cultures, being significantly higher in the regressed state compared to the control. Hypotheses regarding this disagreement could be attributed to the fact that aspartate uptake from the cells is usually not an efficient process which combined with the differential nutrient accessibilities in a *ex vivo* cultivated tissue sample compared to the cell culture system results in a decreased uptake and hence higher extracellular levels. It should also be noted that the *ex vivo* experiments were performed after a cultivation period of 8 hours, compared to the *in vitro* cell cultures where the extracellular samples were corresponding to the spent medium of the last 24h of cultivation. The latter could also explain why the extracellular unlabeled glucose and lactate concentration changes, albeit exhibiting the expected trend, did not reach a statistical significance.

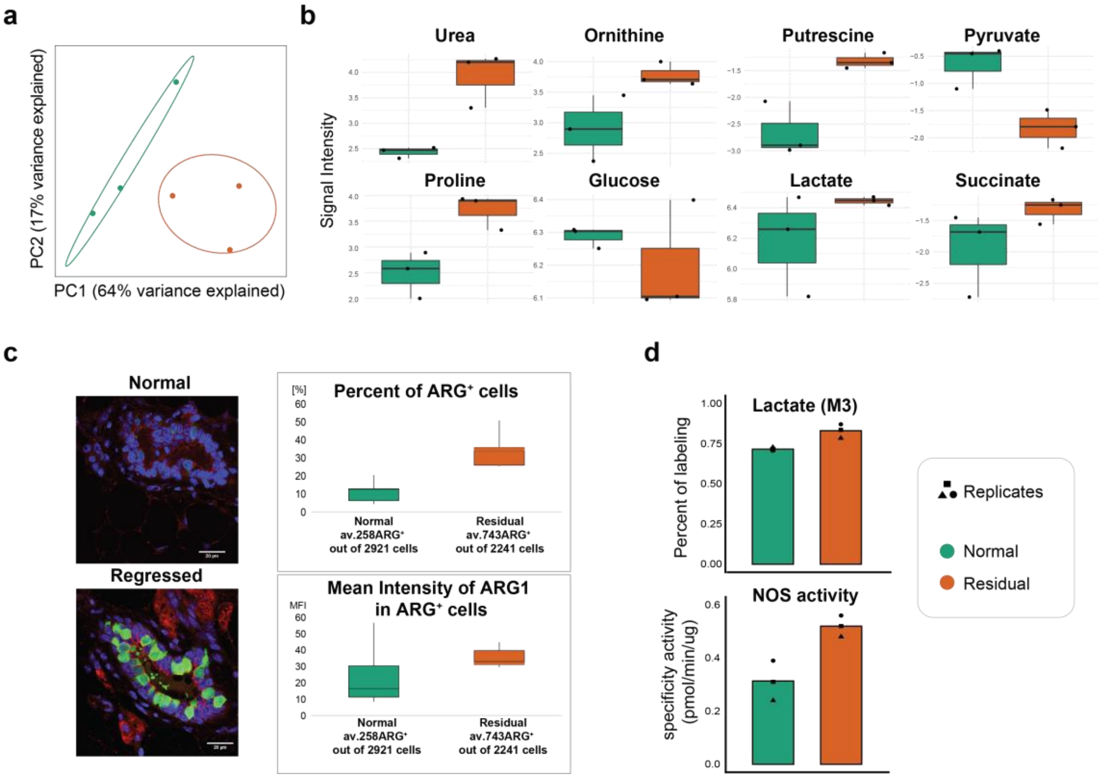

Figure 22: *Ex vivo* and *in vivo* validation supports the altered glycolysis and urea cycle metabolic signatures in the residual population. (a) Principle component analysis of extracellular metabolic profiles of regressed and control mammary glands. Color legend same as in panel (d). (b) Boxplots of selected extracellular metabolites of the regressed and control mammary glands. Apart from glucose, lactate and succinate, all other metabolites are statistically significant (limma analysis, Benjamini-Hochberg adjusted p value ≤ 0.01). The concentrations are measured as signal intensities (area under the curve), and are depicted on the log2 scale. Color legend same as in panel (d). (c) Arginase1 verification on the tissues of animals 9 weeks off doxycycline (regressed mammary glands) compared to healthy age-matched controls. Left panel: IF stains – E-cadherin in red, ARG1 in green; nuclei stained with DAPI (blue). Scale bar: 50 µm. Right panels: Percentage (upper) and mean intensity (lower) of ARG1 positive cells (Ksenija Radić Shechter, PhD thesis). Color legend same as in panel (d). (d) In vivo intracellular metabolic verifications on the regressed and control mammary glands. Upper panel: Intracellular lactate labeling from [U-13C]-glucose (M3 indicates the mass isotopomer labeled in three carbons). Lower panel: enzymatic activity for nitric oxide synthase (NOS).

To further monitor the urea cycle, regressed and control mammary glands isolated from the same animals utilized for the aforementioned metabolomics experiments, were assessed for the enzymatic activity of the nitric oxide synthase (NOS) which catalyzes the production of nitric oxide from arginine. The particular enzyme was chosen owing to the high transcriptional levels of NOS2 in the regressed cells compared to the control and tumor state. The lysates of the regressed mammary glands displayed a significantly higher activity than the control ones (Figure 22d). In addition, immunofluorescence staining for another enzyme of the urea cycle, that of arginase 1 (ARG1) which catalyzes the conversion of arginine to ornithine and urea, was significantly increased in the regressed mammary glands compared to the control (Figure 22c). The quantification of the number of $ARG1^+$ cells, as well as the comparison of the mean intensity in $ARG1^+$ cells on the analyzed tissue sections were also increased in the regressed state compared to the control (Figure 22c) (Ksenija Radić Shechter Shechter, PhD thesis).

Collectively, the *ex vivo* [U-^{13}C]-labeling, extracellular metabolomics, NOS enzymatic assay and immunofluorescence imaging verified the metabolic changes observed in the three dimensional cell cultures for two key metabolic characteristics present in the residual population, namely an enhanced glycolysis and urea cycle activity.

Discussion and Outlook

Cancer metabolic alterations are nowadays a well-established phenomenon. Yet, far less is known for the metabolic physiology underlying residual cells that survive therapy and can contribute to cancer recurrence. This limitation is largely attributed to the difficulties in characterizing and obtaining these populations *in vivo*. In the presented study, to explore the metabolic particularities of residual cells, we employed three dimensional primary epithelial cell cultures derived from a mouse model of breast cancer based on the doxycycline-inducible overexpression of the oncoproteins MYC and HER2. The addition of doxycycline followed by withdrawal could thus simulate cancer induction, formation and regression. We demonstrated that regressed cells, although phenotypically similar to their normal counterparts and despite the absence of oncogenic signaling, they still maintained a metabolic phenotype largely mnemonic of the tumorous population. The residual cells were mainly characterized by an enhanced glycolysis, deregulated TCA cycle and urea cycle activity. While the transcriptional profile of the residual cells was closer to the normal than the tumorous state, the transcriptional changes of the metabolic enzymes specific to the altered metabolic pathways were largely in agreement with the metabolic phenotype. The metabolic phenotype was further supported by reporter metabolite analysis based on the transcriptional data, and flux predictions based on a genome-scale reconstructed mouse model (Ksenija Radić Shechter and Katharina Zirngibl, PhD theses). Two of the main metabolic alterations, namely the deregulated glycolysis and urea cycle, present in the three dimensional cell cultures, were further verified in the *in vivo* mouse model. Finally, evidence of both metabolic phenotypes were detected in breast cancer disease patients after treatment (which approximate the *in vitro* regressed state), substantiating the potential relevance in a human setting (Ksenija Radić Shechter and Katharina Zirngibl, PhD theses).

Both MYC and HER2 oncoproteins have been connected to enhanced glycolysis through the upregulation of glucose transporters and several enzymes along the glycolytic axis, including hexokinase 2 (HK2) and lactate dehydrogenase A (LDHA)[140]. Apart from glycolysis, MYC oncoprotein is known to enhance glutamine catabolism by upregulating both glutamine transporters and glutaminase (GLS), the enzyme which converts glutamine to glutamate[74,141], and to promote proline accumulation by upregulating the enzymes involved in its biosynthesis and inhibiting the enzymes catalyzing its degradation[142]. Moreover, MYC is also implicated to urea deregulation by enhancing the expression of two enzymes of the cycle, namely argininosuccinate synthase 1 (ASS1) and ornithine decarboxylase (ODC)[137]. Concomitantly, deregulation of the aforementioned pathways has been connected to offer several advantages, mainly by supporting the biosynthetic requirements of the highly proliferative cancer cells. Consistent with the above, the presence of such metabolic alterations in the MYC/HER2-induced tumor state was in agreement with expectations. However, the persistence of this phenotype in the regressed cells, albeit with variations or differential responses along the different pathways, was unexpected considering the absence of the oncogenic signals and their non-proliferative nature. On the other hand, this persistence is suggestive that the metabolic alterations are not merely a consequence of the oncogenic inputs, but continue to operate and possibly mediate additional modulatory functions sustaining a potential in the dormant population. Along these lines, it has been shown that a high glucose flux does not always correlate with the proliferative status of the cells[143], and a growing body of evidence support the additional roles that glycolysis could serve. For instance, the biosynthesis of certain metabolites with epigenetic-related functions are dependent on glycolysis. Acetyl-CoA or ADP-ribose, whose carbons mainly derive from glucose, are used for acetylation and poly-(ADP-ribosyl)ation of histones, respectively. These histone modifications are not only connected to gene expression regulation[7], but their presence is essential for DNA damage repair mechanisms too[29]. It is thus interesting to speculate whether a high glycolytic flux in the residual population could be linked to sustainability of the genomic integrity.

Certain glycolytic related metabolites can additionally act as signaling molecules. Accumulation of lactate in the intracellular environment leads to its binding in specific proteins and their stabilization, a process which in turn results in downstream signaling cascades promoting cell survival and growth[144]. Intracellular lactate catabolism has also been connected to breast cancer stemness through the generation of acetyl-CoA and subsequent epigenetic modifications[145]. Interestingly, in our data, lactate intracellular pools were significantly higher only in the residual population compared to the tumor cells where lactate was only altered in the extracellular spent growth medium. Similarly to lactate, two metabolites of TCA cycle, namely succinate and fumarate, were significantly accumulated specifically in the intracellular space of residual cells, with the tumor population exhibiting higher levels only in the extracellular environment. Notably, both succinate and fumarate are implicated in epigenetic modifications through direct inhibition of the alpha-ketoglutarate dependent demethylases that confer histone and DNA demethylation[18]. Additionally, succinate, like lactate, can operate as a signaling molecule, binding and stabilizing proteins, including the hypoxia inducible factor 1 (HIF1α). This protein-metabolite interaction further promotes the glycolytic phenotype through HIF1α-mediated events[146]. Along the same direction, nitric oxide, produced by nitric oxide synthase in the urea cycle whose activity was significantly increased in the residual population, represents another metabolic intermediate with multiple roles including signaling functions that further promote oncogenic pathways[147].

Overall, the aforementioned paradigms highlight how metabolic alterations are not merely required to support the oncogene-induced anabolic requirements, but also actively participating in cancer robustness, fate and recurrence through epigenetic and signaling events amongst other functions. Why and how this tumorous-like metabolic phenotype is maintained in the regressed state is an important question to address. Preliminary results from our research group (Katharina Zirngibl and Ksenija Radić Shechter) show that the DNA methylation landscape is largely similar between tumor and regressed cells and distinct from the normal state. It would be thus interesting to explore whether progressive alterations in the epigenetic landscape proceed or parallels metabolic alterations. In addition, the identified metabolic changes offer promising avenues of (combinatorial) cancer treatment at the metabolic level - together with a disease progression mapping showing points of intervention could potentiate the therapeutic efficacy. Finally, our metabolic results highlight another perspective in cancer monitoring schemes after treatment, setting the question whether the metabolic physiology, rather than phenotypic markers, could be utilized to assess the results of cancer elimination treatments and monitor cancer recurrence.

Contributions

This project is a collaboration with the Martin Jechlinger group, Cell Biology and Biophysics Unit, EMBL Heidelberg, Germany

- Eleni Kafkia (Structural and Computational Biology Unit, EMBL Heidelberg, Germany) was responsible for the metabolomics part and data interpretation: optimization of the three dimensional cell cultures harvesting protocol for metabolomics analysis (together with Federico Villa), experimental metabolic designing schemes, intracellular and extracellular GC-MS data acquisition pipeline optimization and data analysis, *ex vivo* metabolomics and ^{13}C-labeling experiments (together with Ksenija Radić Shechter) and respective data analysis, NOS enzymatic assay (together with Ksenija Radić Shechter)
- Ksenija Radić Shechter, (Jechlinger group, Cell Biology and Biophysics Unit, EMBL Heidelberg, Germany) was responsible for most experiments unless written otherwise, and contributed to all data interpretation
- Katharina Zirngibl (Structural and Computational Biology Unit, EMBL Heidelberg, Germany) was responsible for transcriptomics data analysis, reporter metabolites analysis, flux predictions, correlation of human datasets
- Martin Jechlinger and Kiran Raosaheb Patil (Structural and Computational Biology Unit, EMBL Heidelberg, Germany) supervised the project and designed experiments

Chapter III: Metabolic response of cancer cells to inhibition of lactate dehydrogenase and monocarboxylate transport activities

As mentioned in the previous chapters, cancer cells aberrant metabolic nature is a major transformative effector of their complex pathophysiology. As a consequence of this, efforts are made over the past years to target important facilitators of deregulated metabolic pathways. Enhanced glycolysis, which is characterized by a high glucose consumption towards lactate production, is the most commonly recurring alteration across many cancer types[41]. Considering the pleiotropic roles of lactate, including as a carbon source and signalling molecule enhancing proliferation and metastatic potential, of the most attractive metabolic targets are the enzyme lactate dehydrogenase A (LDH) and the monocarboxylate transporters (MCTs) involved in lactate production and transport, respectively. However, the development of potent inhibitors has proven a challenging task - none of the developed LDH inhibitors has progressed into clinical trials so far, either for reasons of being *in vivo* inactive or due to unfavourable pharmacokinetic properties[57,148,149]. MCT1 offers a promising target and is currently in phase I clinical trials in UK with potential progression to phase II[150] (www.clinicaltrials.gov). Towards a mechanistic understanding of the perturbations of lactate node, the current study, carried out in collaboration with Merck Serono GmbH, aimed at investigating the network-wide metabolic response upon inhibition of LDHA and MCT1 and 2 in cancer cell lines. Metabolic profiling with Gas Chromatography – Mass Spectrometry was utilized to identify the metabolite concentration changes, while [U-^{13}C]-glucose and [U-^{13}C]-glutamine labelling were employed to assess flux changes in the central carbon metabolism.

Treatment with LDHA and MCT1 and 2 inhibitors results in distinctive metabolic profiles characterized by opposite alterations in glycolysis and TCA cycle

To track the metabolic effects derived from the inhibition of LDHA and MCT1 and 2 (hereafter denoted as MCT1/2), SNU398 cells were treated for 1 and 4 hours with either of the inhibitors, and intracellular and extracellular samples were collected for metabolomics analysis via Gas Chromatography - Mass Spectrometry. Control samples included untreated cells and cells treated with DMSO. In total, we were able to reliably quantify 57 and 30 known metabolites in the intracellular and extracellular samples, respectively. The treatments affected the metabolic physiology of the cells, resulted in distinctive intracellular metabolic profiles between the two treated and the control states (Figure 23b). As expected, treatment with the LDHA inhibitor, which targets the lactate production from pyruvate, induced significantly decreased lactate concentration pools in the intracellular environment, already at

1h post-inhibition (Figure 23d). Alterations in the extracellular medium were observed only at 4h post-treatment (Figure 23c). Even though intracellular lactate was decreased, endogenous pyruvate remained largely unchanged. However, in the extracellular environment pyruvate was significantly increased, being in fact the most significantly altered extracellular metabolite (Figure 23c). Apart from the aforementioned changes, cells treated with the LDHA inhibitor exhibited significant alterations for 18 intracellular metabolites, mapped mainly to increased upper glycolytic metabolites (i.e. fructose 1.6-bisphosphate), and decreased intermediates of lower glycolysis (i.e. phosphoenolpyruvate) and TCA cycle, expect citrate, whose levels although significantly changed at 1h, approximated the levels of the control cells after 4 hours of inhibition (Figure 23d). Additional intracellular changes included an overall increase in unknown sugar phospho-compounds. Interestingly, we also observed significant increases in intracellular fructose and sorbitol levels (Figure 23d).

The MCT1 and 2 are involved in the bidirectional transportation of lactate, as well as pyruvate and ketone bodies, mainly inside and outside of the cells, with the direction of the transportation being dependent on the concentration gradients. Consistent with expectations, inhibition of the MCT1/2 resulted in an almost five-fold accumulation of the intracellular lactate from the first hour of treatment (Figure 23e). No greater accumulation was observed in the last time point. In the extracellular environment, lactate levels started to reduce only at four hours post-inhibition (Figure 23c). In addition to lactate, we observed that pyruvate extracellular levels were largely depleted in the extracellular medium, suggesting a strong inhibition in the export of pyruvate too (Figure 23c). Despite this, intracellular pyruvate levels remained relatively unchanged. Interestingly, regarding the downstream intracellular metabolic effects, the MCT1/2 inhibitor induced an almost reversed metabolic phenotype compared to the one induced after treatment with the LDHA inhibitor. More specifically, there was a significant decrease of the upper glycolytic metabolites, and an accumulation of TCA cycle intermediates (Figure 23e and 24). Finally, similar alterations between the two treatments were detected for the following metabolites: N-acetylserine, N-acetylglutamate, gluconate and taurine. These metabolites displayed decreased intracellular levels in the cells treated with either of the two inhibitors compared to the control state (Figure 24d and 24e). In the extracellular space, glutamine levels were significantly decreased at 4h post-treatment with either of the inhibitors (Figure 23c).

Overall, the initial metabolic profiling matched the expectations regarding the metabolites directly affected from the inhibitors treatment (i.e. lactate and pyruvate). The opposite concentration pools of upper glycolysis and TCA cycle intermediates (Figure 24) pointed to interesting flux re-organizations following the inhibitors treatment. Along these lines, as a next step, we were interested to compare the relative pathway activities and the relative contributions of the two major carbon sources, glucose and glutamine, that feed glycolysis and TCA cycle, upon treatment with the two inhibitors.

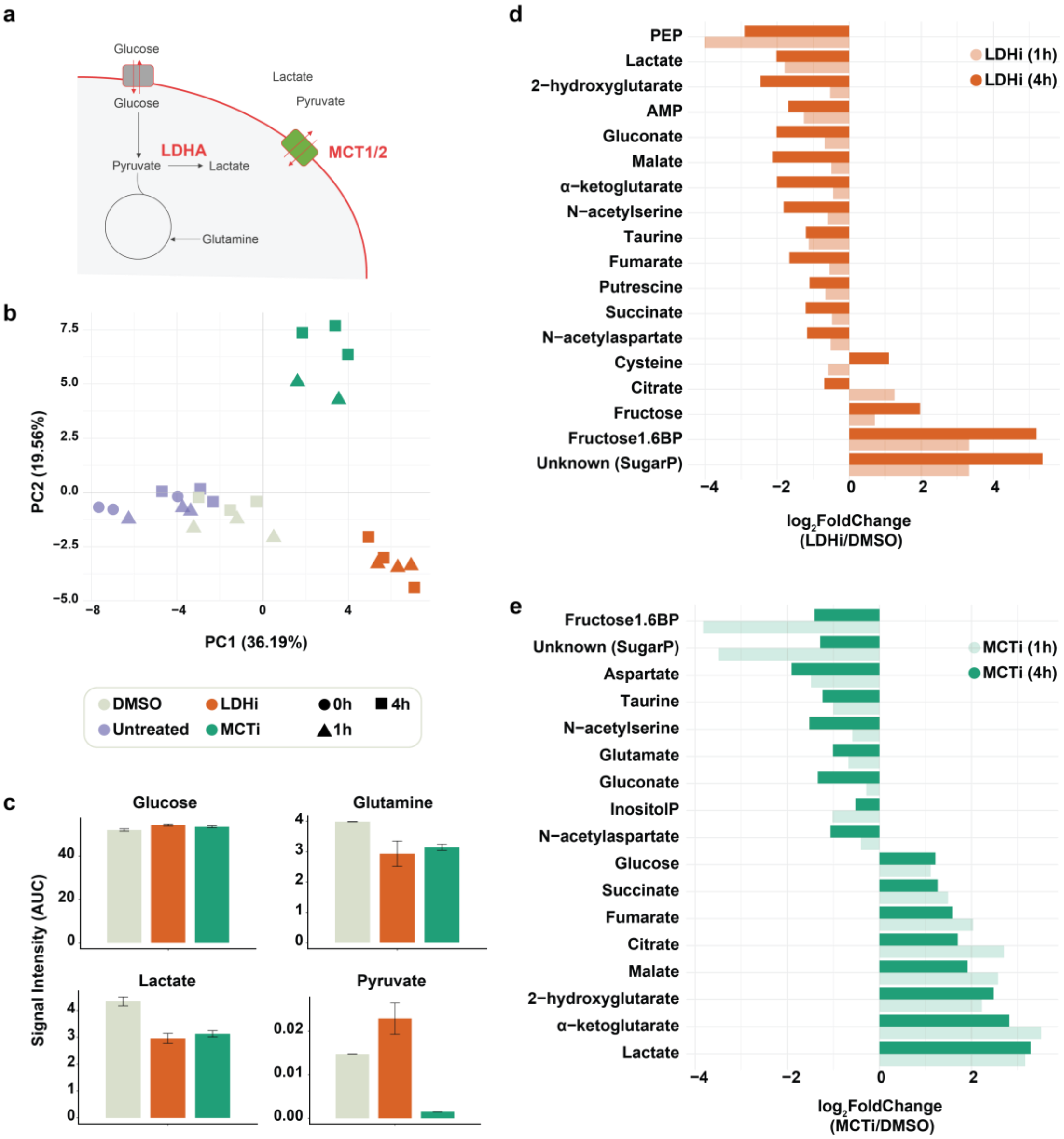

Figure 23: Metabolic alterations after inhibition of lactate dehydrogenase and monocarboxylate transporters 1 and 2. (a) Abstract schematic representation of the main metabolic targets and the respective metabolic pathways. (b) Principle component analysis based on the intracellular metabolic profiles of control cells (treated with DMSO and untreated), cells treated with the LDHA inhibitor (LDHi) and cells treated with the MCT1/2 inhibitor (MCTi), for 1h and 4h. Untreated cells prior treatment (0h) were also analyzed. (c) Extracellular metabolite abundances (after 4h of treatment) (abundances are measured as area under the curve, AUC). Color scheme is the same as in panel b. Error bars denote standard error of the mean (sem) for 3 biological replicates. (c) and (d) Barplots for the statistically significant intracellular metabolites depicting the fold change in treated cells relative to control DMSO samples (based on limma analysis, see methods).

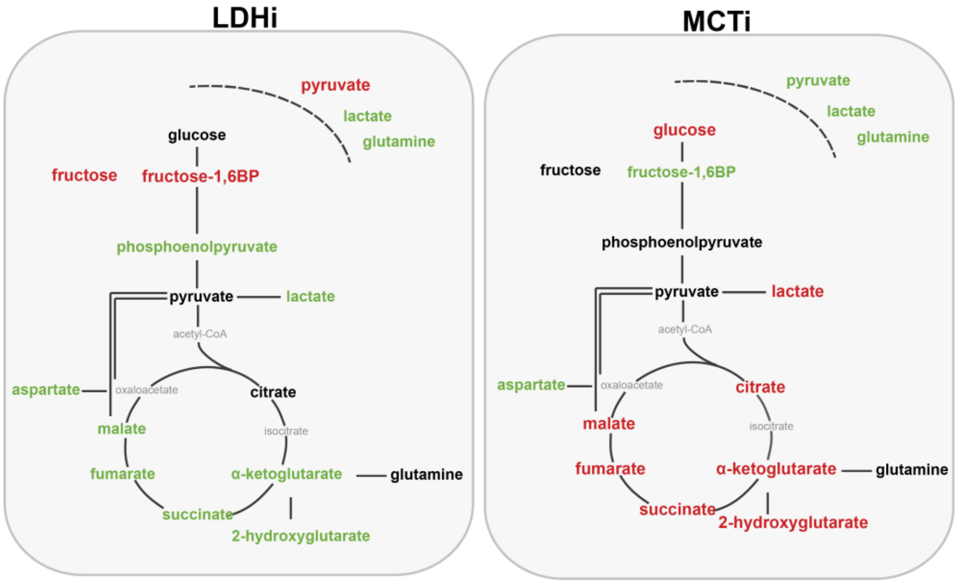

Figure 24: Overview of metabolic effects on glycolysis and TCA cycle upon inhibition of lactate dehydrogenase and monocarboxylate transporters 1 and 2. An abstract network of glycolysis and TCA cycle is depicted along with the significant changes in the concentration pools of cells treated with the lactate dehydrogenase inhibitor (LDHi) or with the monocarboxylate transporters 1 and 2 inhibitor (MCTi) compared to the control cells (treated with DMSO). Red and green corresponds to increased and decreased concentrations in the inhibitors treated samples compared to the control cells.

Inhibition of MCT1 and 2 ablates pyruvate export but not lactate

To examine glucose metabolism, SNU398 cells were treated for 1, 2, 4, 6 and 8 hours with either of the inhibitors in growth medium supplemented with [U-^{13}C]-glucose, and both intracellular and extracellular samples were collected to monitor ^{13}C carbon incorporation in glycolytic and TCA cycle intermediates. Following treatment with the MCT1/2 inhibitor, we did not detect any labelled pyruvate in the extracellular environment, indicating a stable inhibition of its export (Figure 25). On the other hand, extracellular lactate labeled in all three carbons (M3 isotopomer) started to increase after 4h of treatment (Figure 25) which was suggestive of the induction of compensatory mechanisms contributing to its export. Recently published data showed that MCT1 inhibition in breast cancer cell lines does not eventually affect lactate export owing to induced upregulation of the monocarboxylate transporter MCT4 which facilitates compensatory lactate export[151]. When cells were treated with the LDHA inhibitor, both intracellular and extracellular M3 lactate isotopomer levels started to increase after 4h of inhibition pointing to a possible loss of inhibitor's activity (Figure 25). Overall, these results indicated that the MCT1/2 inhibitor strongly affects pyruvate export and has mediocre effects on lactate export, which instead starts to increase after 4h of treatment through compensatory export mechanisms mediated by other MCT transporters, including MCT4. LDHA inhibitor showed an activity loss after 4h.

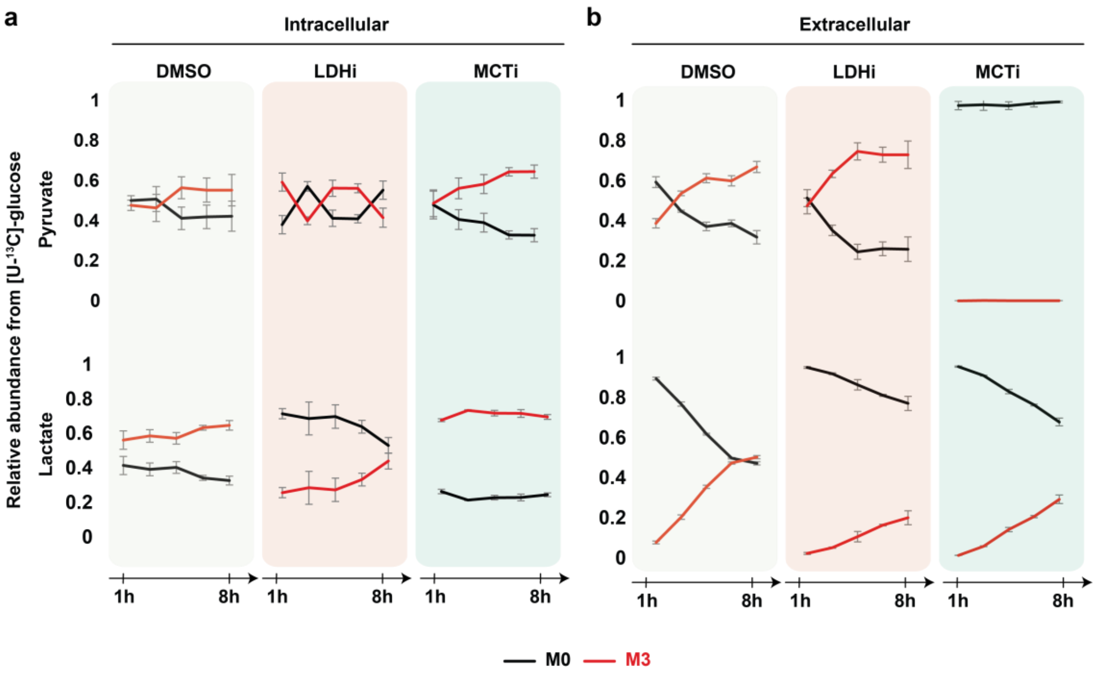

Figure 25: Labeling kinetics of pyruvate and lactate upon treatment with the lactate dehydrogenase inhibitor or the monocarboxylate transporters 1 and 2 inhibitor. Labeling of pyruvate (upper panels) and lactate (lower panels) from [U-^{13}C]-glucose in intracellular (a) and extracellular (b) samples after treatment with the LDHA inhibitor (LDHi) or the MCT1/2 inhibitor (MCTi), compared to control cells (DMSO-treated). The measurements were acquired over time, for five points corresponding to 1h, 2h, 4h, 6h and 8h after treatment with the inhibitors or DMSO and [U-^{13}C]-glucose addition. M0 to Mn indicates the different mass isotopomers (M0 denotes that zero carbons are labeled, and Mn denotes that n number of carbons are labeled). Error bars denote standard error of the mean (sem) for three biological replicates.

Inhibition of MCT1/2 or LDHA induces a decreased lipid biosynthesis flux

As a next step, we aimed to monitor whether there were alterations in glucose contribution to TCA cycle metabolites following inhibitors treatment. To estimate glucose carbon entry to TCA cycle, we calculated the ratio of total labeled carbon contribution in citrate (FC citrate) to total labeled carbon contribution in pyruvate (FC pyruvate) (Figure 26). Of note, we estimated the ratio to labeled pyruvate in order to exclude upstream changes in glycolysis. We found that the ratio was similar in both inhibitors and no different from the control cells (Figure 26). However, for the case of MCT1/2 inhibitor, the isotopomers of TCA cycle intermediates generated from multiple rounds of the cycle were increased (calculated as the ratio of M4 citrate to M2 citrate) (Figure 26). Same trend was observed for LDHA inhibition, however to a lesser extent (Figure 26). This increase indicated that less carbon leaves the TCA cycle for lipid biosynthesis at the level of citrate, and hence less anaplerosis of the TCA cycle takes place. The reduced biosynthetic flux was further confirmed by the ratio of N-acetyl-aspartate M2 isotopomer to pyruvate M3 isotopomer (again to exclude any upstream changes in glycolysis), which can be used as an estimate for the acetyl-moiety derived from acetyl-CoA (Figure 26).

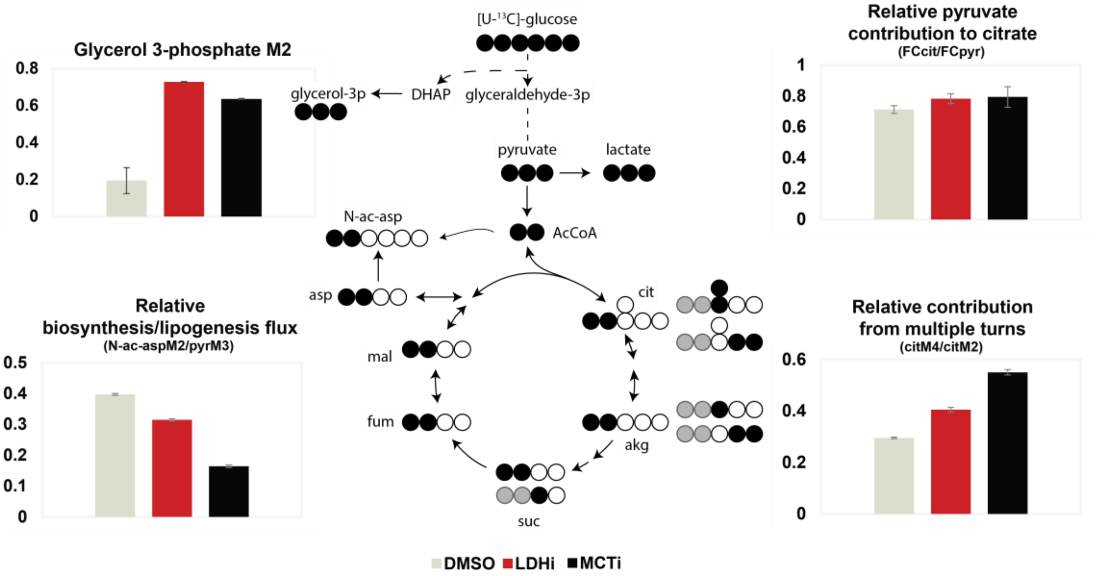

Figure 26: Effects on glucose's contribution to the central carbon metabolic pathways glycolysis and TCA cycle upon treatment with the lactate dehydrogenase inhibitor or the monocarboxylate transporters 1 and 2 inhibitor. The effects are depicted together with a schematic representation of glycolysis and TCA cycle pathways, and the respective carbon atom (circles) transitions for specific metabolites derived using [U-13C]-glucose as a tracer. Filled black and grey circles denote 13C carbons derived from the first and second round of oxidative TCA cycle, respectively. 12C carbons are depicted as empty circles. M0 to Mn indicates the different mass isotopomers (M0 denotes that zero carbons are labeled, and Mn denotes that n number of carbons are labeled). Note that not all possible isotopomers are depicted. For glycerol 3-phosphate only the ion that corresponds to the two carbon fragment was detected and quantified. Error bars denote standard error of the mean (sem) for three biological replicates. Abbreviations: N-ac-asp, N-acetyl-aspartate; AcCoA, Acetyl-Coenzyme A; cit, citrate; akg, alpha-ketoglutarate; suc, succinate; fum, fumarate; mal, malate; asp, aspartate; citM4, citrate M4 isotopomer; citM2, citrate M2 isotopomer; N-ac-aspM2, N-acetyl-aspartate M2 isotopomer; pyrM3, pyruvate M3 isotopomer, FCcit, total carbon contribution in citrate; FCpyr, total carbon contribution in pyruvate (see texts for details).

Both inhibitors, albeit more strongly the one targeting LDHA, had an increase in the glycerol 3-phosphate M2 isotopomer, possibly reflecting a compensatory route for NAD^+/NADH regeneration which is disturbed by the inhibition of pyruvate to lactate (Figure 26). As a final step, we estimated the glucose uptake rates, and we found an increase only for the LDHA inhibitor (appendix, Figure 32). Taken together, these data revealed that although there was no alteration in glucose uptake rates and glucose's entry in TCA cycle, both inhibitors induced a decrease in the biosynthetic flux, stronger for the case of MCT1/2 inhibitor, which is suggestive of a reduction in the proliferative status of the cells.

Contribution of glutamine to TCA cycle is altered after treatment with LDHA or MCT1/2 inhibitors

Glutamine is a major carbon source for TCA cycle anaplerosis, and as a next step we aimed to study whether the inhibitors induce changes in glutamine contribution to TCA cycle metabolites. For this purpose, the same experimental scheme as before was performed using uniformly labelled glutamine, [U-^{13}C]-glutamine. Following treatment with the inhibitors, we identified a significant decrease in the contribution of glutamine to the TCA cycle for both cases, but even more pronounced in the inhibition of LDHA (Figure 27). Glutamine enters TCA cycle at the level of alpha-ketoglutarate which subsequently is converted to succinyl-CoA and then to succinate. Interestingly, in the control cells and in cells treated with the LDHA inhibitor, we observed a substantially diluted labelling in succinate (Figure 27). We calculated that only 40% originated from alpha-ketoglutarate, with the rest deriving from an unlabeled source. However, in the case of MCT1/2 inhibition most of the succinate originated from alpha-ketoglutarate (Figure 29).

Anaplerosis of the TCA cycle flux can take place in different intermediates which results in label dilution if the source molecules are not labelled. Possible sources that could contribute to succinate dilution are odd fatty acids and branched chain amino acid catabolism (i.e. valine, isoleucine, leucine) that feed the TCA cycle at the level of succinyl-CoA. Branched chain amino acids catabolism starts in the cytoplasm and continues in the mitochondria. Amongst the produced metabolites are small organic compounds (i.e. 2-hydroxybutyric acid, acetoacetate), and studies have shown that their transportation, at least inside and outside of the cell, can be facilitated by MCTs. Our succinate labeling dilution results rather indicated that MCTs can also contribute to an intracellular trafficking between the different compartments, and more specifically between the cytoplasm and mitochondria. This is in agreement with a recent study where immunofluorescence indicated an intracellular localization signal for MCT1[152]. As such, succinate increased labeled isotopomers after treatment with the MCT1/2 inhibitor could be attributed to an inhibitory effect within the cell, targeting the transportation of small organic acids between the different compartments. Analysis of the isotopomers for the rest of the TCA cycle intermediates indicated that, despite the substantial dilution of succinate in the control cells and cells treated with the LDHA inhibitor, there was no dilution in the labeling of malate (Figure 27). In addition, the labelled fraction of malate was higher than that of succinate (~90%), which indicated the existence of two separate pools of succinate that are not in complete exchange. Moreover, we calculated the fraction of malate originating from alpha-ketoglutarate to quantify the label dilution in malate and we found it similar in all cases (Figure 27). There was a marginal increase and decrease in the reductive glutamine metabolism after MCT1/2 and LDHA inhibition, respectively, albeit not significant. Finally, glutamine uptake rates were decreased only for the case of LDHA inhibitor (appendix, Figure 32). Taken together, these results indicated less glutamine contribution in the TCA cycle after inhibition of LDHA. Additionally, we identified an interesting point where MCT1/2 inhibitory effects could be facilitated by targeting intra-organelle transportation of small organic acids but further validation of this mechanism is required.

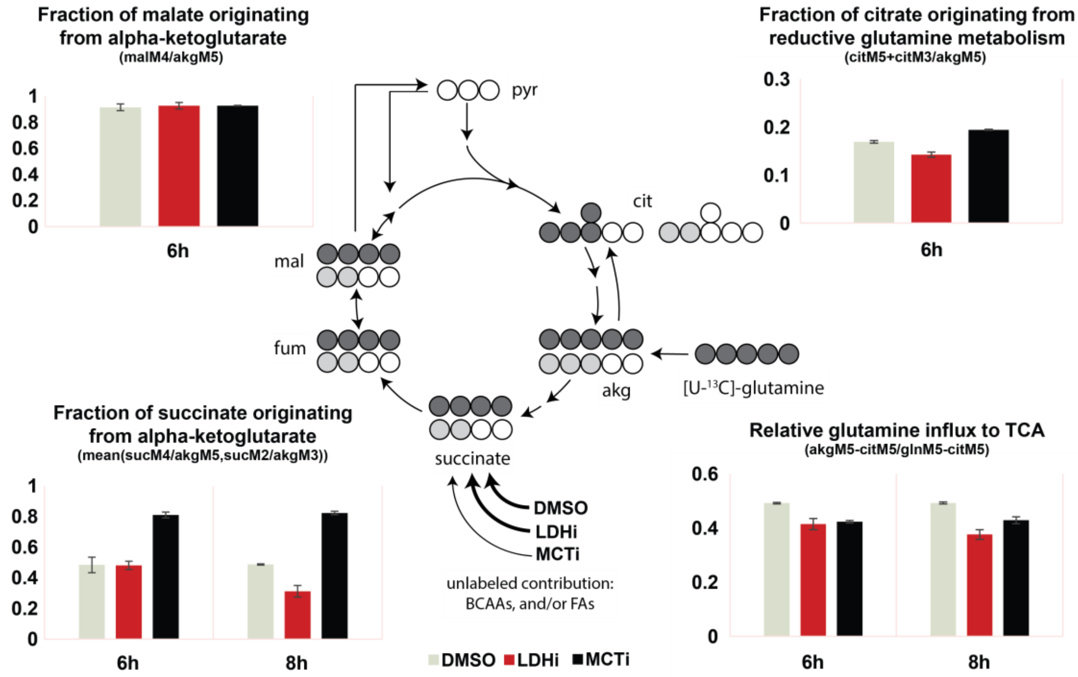

Figure 27: Effects on glutamine's contribution to TCA cycle after treatment with the lactate dehydrogenase inhibitor or the monocarboxylate transporters 1 and 2 inhibitor. The effects are depicted together with a schematic representation of the TCA cycle pathways, and the respective carbon atom (circles) transitions for specific metabolites derived using [U-13C]-glutamine as a tracer. Filled dark grey and grey circles denote 13C carbons derived from the first and second round of oxidative TCA cycle, respectively. 12C carbons are depicted as empty circles. M0 to Mn indicates the different mass isotopomers (M0 denotes that zero carbons are labeled, and Mn denotes that n number of carbons are labeled). Note that not all possible isotopomers are depicted. Error bars denote standard error of the mean (sem) for three biological replicates. Abbreviations: cit, citrate; akg, alpha-ketoglutarate; fum, fumarate; mal, malate; citM5, citrate M5 isotopomer; citM3, citrate M3 isotopomer; akgM5, alpha-ketoglutarate M5 isotopomer; akgM3, alpha-ketoglutarate M3 isotopomer; glnM5, glutamine M5 isotopomer; sucM4, succinate M4 isotopomer; sucM2, succinate M2 isotopomer; malM4, malate M4 isotopomer.

LDHA inhibition points to alterations in fructose metabolism

Owing to the intracellular opposite metabolic phenotype following LDHA and MCT1/2 inhibition, we decided to treat cells with both inhibitors simultaneously and measure the intracellular and extracellular metabolic profiles. Of note, calculation of the proliferation rates after the treatment did not result in significant changes compared to the control cells (Merck data). In the combinatorial treatment, the MCT1/2 inhibitor appeared as the dominant metabolic effector (Figure 28a), and most of the induced changes in the metabolic concentration pools resembled the ones found after treatment with the MCT1/2 inhibitor alone (Figure 28a and 28b). Despite this, we observed an interesting pattern in fructose 1.6-bisphosphate levels, which, although decreased after 1h of treatment, they displayed an opposite change at 4h (Figure 28b). In addition, the combinatorial treatment resulted in an increase in fructose and sorbitol levels, and several unknown sugar phopho-compounds, similarly to the metabolic phenotype observed after treatment with the LDHA inhibitor alone (Figure 28c). Overall, the combination treatment pointed to interesting alterations around fructose metabolism that seem to be induced from the inhibition of LDHA. This phenotype could point to an interesting node connected to the polyol pathway and the conversion of glucose to sorbitol and then fructose which can further be degraded (fructolysis) and enter glycolysis by skipping the heavily allosterically controlled step of PFK-1[153,154].

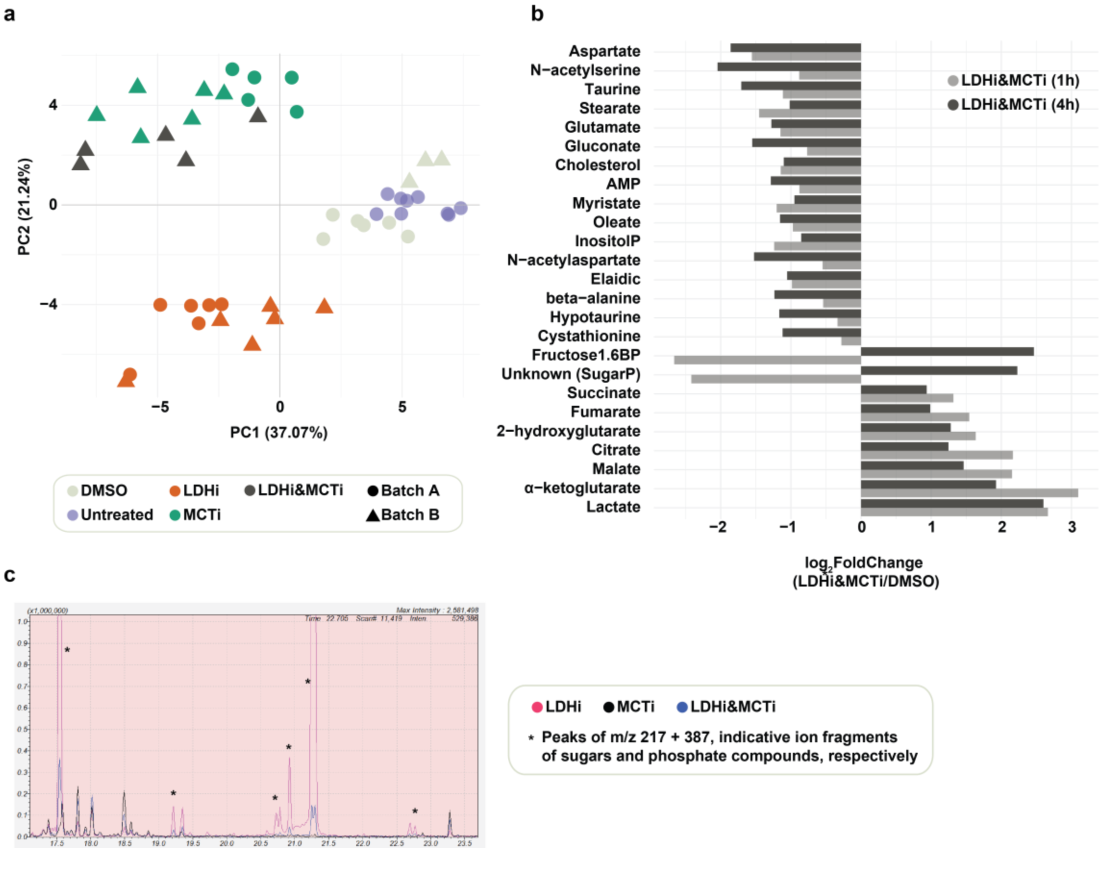

Figure 28: Metabolic alterations upon dual inhibition of lactate dehydrogenase and monocarboxylate transporters 1 and 2 activities. (a) Principle component analysis based on the intracellular metabolic profiles of control cells (treated with DMSO and untreated), cells treated with the LDHA inhibitor (LDHi), cells treated with the MCT1/2 inhibitor (MCTi), and cells treated with both inhibitors (LDHi&MCTi), for 1h and 4h. Untreated cells prior treatment (0h) were also analyzed. Batch A and Batch B denote two distinct biological experiments with three replicate per condition each. (b) Barplots for the statistically significant intracellular metabolites depicting the fold change in treated cells relative to control DMSO samples (based on limma analysis, see methods). (c) Overlay for the three treated conditions of the raw mass ion chromatographs of m/z 217 and 387, which are representative ion fragments for sugar phosphate compounds.

Summary and Outlook

Collectively, in the present study we comprehensively mapped at the metabolic level the effects exhibited after inhibition of LDHA and MCT1/2, alone and in combination. We detected opposite alterations in the concentration pools of metabolites in the glycolysis and TCA cycle pathways, reflective of the metabolic rewiring that takes place as a response to decreased and increased endogenous lactate levels occurring after the inhibitors treatments. Moreover, we identified an interesting deregulation around fructose metabolism induced by LDHA inhibition, which could indicate a compensatory mechanism of the cell to continue glycolysis by avoiding heavily regulated metabolic nodes. As a consequence of this, inhibitors that target the fructolytic pathway (i.e. inhibitors against aldolase b or fructokinase) could be evaluated in the future in a combinatorial scheme that could offer better therapeutic potential. By employing ^{13}C-labeling, we identified that MCT1/2 inhibitor strongly inhibits pyruvate export, but not lactate, which is in agreement with recently published results from other groups[151]. Regarding overall fluxes, for the MCT1/2 inhibitor, even though the glutamine and glucose uptake levels were similar to DMSO control cells, the significantly decreased lipid biosynthesis in combination with the increased concentrations of TCA compounds could argue to an overall lower flux. For the LDHA inhibition, the higher glucose and glutamine uptake rates in combination with the lower glutamine contribution to TCA cycle and the decreased TCA intracellular concentrations could point to an overall higher flux. Finally, we identified a so-far underestimated mechanism of action for the MCT1/2 inhibitor, rendering of further characterization, which suggests that its effects could be partially attributed to the attenuation of trafficking of small organic molecules within the cell between the different compartments.

Contributions

This project is a collaboration with Merck Serono Gmbh Germany

- Eleni Kafkia (Structural and Computational Biology Unit, EMBL Heidelberg, Germany) was responsible for the metabolomics and ^{13}C-labeling sample preparation, data acquisition, data analysis and data interpretation
- Paula Jouhten (VTT, Technical Research Center of Finland) was responsible for data analysis and interpretation
- Doreen Musch and Christine Siegl (Merck Serono Gmbh Germany) were responsible for cell culture treatments and sample harvesting
- Kiran Raosaheb Patil (Structural and Computational Biology Unit, EMBL Heidelberg, Germany) and Ralph Lindemann (Merck Serono Gmbh Germany) supervised the project and designed experiments

Concluding remarks

Concluding Remarks

Fundamental nuclear activities, ranging from DNA synthesis to gene regulation, are often dependent on small molecules, metabolites, commonly perceived to diffuse from other subcellular compartments into the nuclear sites of need. In the present PhD thesis, we demonstrate that the HeLa nucleus harbors a functional metabolic network that locally produces metabolites with key roles in nuclear processes, including chromatin and RNA chemical modifications, and nucleotide metabolism. This nuclear metabolic architecture presents hereafter important implications for the regulatory interplay between the metabolic state of a cell and gene expression. Considering that the intracellular environment is a rather viscous space where the free diffusion of even small molecules could be restricted, or metabolites could be siphoned to other metabolic processes until they reach the nucleus, our findings suggests how the efficiency of the epigenetic or epitranscriptomic reactions could be ensured regarding the accessibility and availability of these required metabolites. The detected intra-nuclear produced metabolites are known components of the aforementioned reactions, however given metabolites' increasingly recognized pleiotropic roles, future studies could further reveal an additional spectrum of nuclear processes they are associated with, as well as potential moonlighting functions of the respective metabolic enzymes. It is enticing to hypothesize the existence of additional nuclear metabolic networks apart from the one described here, or even smaller distinctive metabolic sub-niches on specific sites within the nuclear compartment, and the respective functionalities they could serve. Ultimately, the novel concept generated by this research has fundamental extensions relevant for understanding genome stability, embryonic development, stem cell differentiation, ageing and tumor progression. The complex and dynamic relationship between the latter and metabolism is further highlighted by our finding regarding the persistence in regressed cells of a metabolic phenotype mnemonic to the tumorous state. This in turn suggests that metabolic alterations are not merely a consequence of oncogenes, but could have a pivotal role in maintaining a tumorous potential that should further be explored in future studies. Overall these findings present broad and far-reaching implications in cancer monitoring schemes and therapeutics interventions.

Materials and Methods

Chapter I – Materials and Methods

Nuclei Purification

HeLa Kyoto cells were maintained under standard cultivation conditions (37°C, 5% CO_2) in high-glucose (5g/L) Dulbecco's modified Eagle medium (DMEM, GlutaMax, Thermo Fischer Scientific) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Thermo Fischer Scientific). For nuclei experiments, the cells were seeded in 245mm plates (2.0 x 10^6 cells/plate) and cultivated for 3 days prior to collection. Nuclei collection was adapted from Ori and colleagues[155]. Briefly, cells were washed twice with 10ml PBS (room temperature), followed by harvesting with trypsinization. Note that from this step and onwards, all steps and buffers were performed in cold conditions. After centrifugation at 500g for 5min at 4°C, the cell pellets were washed with 10ml ice cold PBS and centrifuged again using the aforementioned conditions. The cell pellet was subsequently resuspended in 7.5ml ice cold hypotonic buffer A (50mM Tris-HCl pH7.5, protease inhibitors cocktail, Sigma-Aldri) and incubated on ice for 30min. Once the cells were swollen, rapture was done utilizing a Dounce homogenizer with pestle B. Cell lysis efficiency was monitored via light microscopy. Once cells were efficiently lysed, the lysates were centrifuge at 1000g for 8min, at 4°C, the supernatant was removed, and the pelleted nuclei were resuspended in 10ml ice cold buffer B (0.25M sucrose, 50mM Tris-HCl pH7.5, 25mM KCl, 5mM $MgCl_2$, 2mM DTT, protease inhibitors cocktail) and transferred in round-bottom cell culture tubes. Nuclei were centrifuged at 1000g for 8min, at 4°C, and the supernatant was removed. The nuclei pellets were resuspended in 3ml of buffer 2C:1B (for recipe of buffer C, see below), and the samples were transferred to an ultracentrifuge tube. The sucrose cushion (buffer C, 2.3M sucrose, 50mM Tris-HCl pH7.5, 25mM KCl, 5mM $MgCl_2$, 2mM DTT, protease inhibitors cocktail) was slowly placed at the bottom of the tubes via a wide long needle of at least 18G attached to a luer-lock syringe. The samples were centrifuged in a Beckman ultracentrifuge using the SW55Ti rotor at 15000 rpm for 30min at 4°C. After centrifugation, the interphase and the supernatant were carefully aspirated via vacuum, and the pellet was resuspended in 1ml ice cold buffer B. The nuclei were centrifuged at 1000g for 8min, at 4°C, and the step was repeated once more. The pelleted nuclei were immediately utilized for 13C-labeling. Alternatively, samples were also saved at -80°C for immunoblot analysis.

13C-labeling experiments in isolated nuclei and cell lysates and metabolite extraction

For the 13C-labeling experiments, the following modified buffer B was prepared: 0.25M sucrose, 50mM Tris-HCl pH7.5, 25mM KCl, 5mM $MgCl_2$, 2mM DTT, protease inhibitors cocktail, 1mM ATP, 1mM ADP, 1mM FAD, 1mM NAD$^+$. The modified buffer B was supplemented with either of the following [U-^{13}C] substrates at a 10mM final concentration: [U-^{13}C]-pyruvate, [U-^{13}C]-citrate, [U-^{13}C]-glutamine, [U-^{13}C]-succinate (Cambridge Isotopes, Inc). Fluorocitrate (Sigma-Aldrich) was added at the indicated experiments at a final concentration of 0.5µM. Equal numbers of nuclei between the cell lysates labeling experiment and the nuclei labeling experiments were calculated based on counting using a Neubauer chamber. For the cell lysates experiment, the samples were collected immediately after the homogenization step (see section above) and the hypotonic buffer was adjusted to the modified buffer B by adding the missing compounds and the 13C-labeling substrates, as described above. For the nuclei experiments, isolated nuclei were re-suspended in 500 µl modified buffer B with the 13C-labeling substrates. The nuclei or the cell lysates were placed in 12-well cell culture plates and incubated in the dark, at 37^{o}C, for 5 hours. After incubation, the nuclei were collected and centrifuged at 1000g for 8min, at 4^{o}C. The pellets were resuspended in 500 µl of ice cold PBS and centrifuged at 1000g for 8min, at 4^{o}C. The supernatant was carefully aspirated and the nuclei were resuspended in 300µl of ice cold HPLC-grade methanol (Biosolve). The methanol/nuclei mixture was stored at -80^{o}C until metabolite extraction (performed on the next day). For metabolite extraction, 10µl of ribitol (50µg/ml) (Alfa Aesar, UK) was added to the nuclei/methanol mixture and the samples were vortexed for 2 minutes. Subsequently, the samples were incubated at 72^{o}C for 15min, followed by addition of 300 µl ice cold MilliQ water and vortex for 2 minutes. The samples were centrifuged at 15000rpm for 10 minutes, at 4^{o}C, the supernatants were carefully transferred in glass vials (Agilent, Santa Clara, CA), dried in a speed-vac, and stored at -80^{o}C until GC-MS analysis.

Gas Chromatography - Mass Spectrometry (GC-MS) data acquisition and analysis[1]

Dried polar metabolites were derivatized with 40 µl of 20 mg/mL methoxyamine hydrochloride (Alfa Aesar, UK) solution in pyridine for 90 min at 37^{o}C, followed by reaction with 80 µL N-methyl-trimethylsilyl-trifluoroacetamide (MSTFA) (Alfa Aesar, UK) for 10 hours at room temperature, as justified in. GC-MS analysis was performed using a Shimadzu TQ8040 GC-(triple quadrupole) MS system (Shimadzu Corp.) equipped with a 30m x 0.25 mm x 0.25 µm DB-50MS capillary column (Phenomenex, USA). 1 µl of sample was injected in split mode (split ratio 1:5) at 250^{o}C using helium as a carrier gas with a flow rate of 1 ml/min. GC oven temperature was held at 100^{o}C for 4 min followed by an increase to 320^{o}C with a rate of 10^{o}C/min, and a final constant temperature period at 320^{o}C for 11 min. The interface and the ion source were held at 280^{o}C and 230^{o}C, respectively. The detector was operated both in

---

[1] The following method part is based on the methods described on Strucko and colleagues[160], and Blasche and colleagues[169] where it was originally written by myself

scanning mode recording in the range of 50-600 *m/z*, as well as in MRM mode for specified metabolites. The metabolite quantification was carried out by calculating the area under the curve (AUC) of the identified marker ions of each metabolite. Natural abundance isotopes were corrected using the Isotope Correction Toolbox (ICT)[156].

Generation of Constructs

Constructs containing the gene of interest fused to BioID2 and a FLAG epitope were generated via Gateway cloning technology (Invitrogen). Firstly, two destination vectors containing the BioID2 and FLAG for N-terminus fusion and C-terminus fusion were created. To achieve this, we used two destination vectors with the older version of BioID (pcDNA5-pDEST-BioID-FLAG-N-term and pcDNA5-pDEST-BioID-FLAG-C-term, donated from Martin Beck, EMBL Heidelberg) and replaced it with the new and smaller version of BioID2 via Gibson assembly (New England Biolabs). Briefly, BioID2 in pcDNA3.1-BioID2-HA plasmid was PCR amplified with primers containing the FLAG epitope for N- or C-terminus integration (FLAG-BioID2 or BioID2-FLAG). pcDNA5-pDEST-BioID-FLAG-N-term and pcDNA5-pDEST-BioID-FLAG-C-term plasmids were amplified by PCR, excluding the BioID-FLAG region, and creating appropriate overlapping to the BioID2-FLAG or FLAG-BioID2 end fragments. One µl of the PCR products (15ng BioID2-FLAG or 15ng FLAG-BioID2, 7.5ng pcDNA5-pDEST---N-term or 10ng pcDNA5-pDEST---C-term) were incubated with 2 µl Gibson assembly reaction mix at 50^{o}C for 60 min. Two µl of the reaction were diluted with 8 µl water, and 4 µl of this dilution were used to transform One Shot *ccd*B Survival 2 T1^R competent cells (Invitrogen). Transformants were selected with chloramphenicol (33 µg/ml) (Sigma-Aldrich) and ampicillin (100 µg/ml) (Sigma-Aldrich). The entry clones with the genes of interests (PDHB, ACO2, IDH2, IDH3G, IDH1 and OGDH) were purchased from the Human ORFeome Collection (Dharmacon). The final constructs were created by performing the LR recombination reaction between the destination vectors and the entry clones according to manufacturer's instructions. Briefly, 1 µl of entry clone (~100ng) was combined with 1 µl of destination vector (~150ng) and 2 µl of LR Clonase II reaction buffer into a final volume of 10 µl with TE buffer. The reaction was incubated at 25^{o}C for 90 min followed by addition of 1 µl proteinase K and incubation at 37^{o}C for 10 min. One µl of the LR reaction was used to transform DH5α competent E. coli (Invitrogen) and transformants were selected with the appropriate antibiotics.

All PCR conditions were defined using the http://webpcr.appspot.com/. All constructs were extracted using the QIAprep Spin Miniprep Kit (Qiagen). The final constructs were sequenced-verified by Sanger sequencing at Eurofins Genomics. An overview of all vectors and selection antibiotics, as well as primers can be found in Tables 1, 2 and 3.

Table 1: Destination vectors and entry clones used for the generation of the engineered constructs

Constructs	Type	Selection Antibiotics	Catalogue Number	Source
pcDNA3.1-BioID2-HA		Ampicillin	#74224	Addgene
pcDNA5-pDEST-BioID-FLAG-N-term	Destination Vector	Ampicillin		Martin Beck, EMBL Heidelberg [157]
pcDNA5-pDEST-BioID-FLAG-C-term	Destination Vector	Ampicillin		Martin Beck, EMBL Heidelberg [157]
pcDNA5-pDEST-BioID2-FLAG-N-term	Destination Vector	Ampicillin		Created for this study (see text)
pcDNA5-pDEST-BioID2-FLAG-C-term	Destination Vector	Ampicillin		Created for this study (see text)
PDHB	Entry clone (pENTR221)	Kanamycin	OHS5894-202496027	Human ORFeome (Dharmacon)
ACO2	Entry clone (pENTR221)	Kanamycin	OHS5894-202497317	Human ORFeome (Dharmacon)
IDH1	Entry clone (pENTR221)	Kanamycin	OHS5894-202498644	Human ORFeome (Dharmacon)
IDH2	Entry clone (pENTR221)	Kanamycin	OHS5894-202497544	Human ORFeome (Dharmacon)
IDH3G	Entry clone (pENTR201)	Kanamycin	OHS5894-202501324	Human ORFeome (Dharmacon)
OGDH	Entry clone (pENTR221)	Kanamycin	OHS5894-202498466	Human ORFeome (Dharmacon)

Cell Culture and generation of stable inducible cell lines

Flp-In T-Rex HeLa cells were maintained under standard cultivation conditions (37°C, 5% CO_2) in high-glucose (5g/L) Dulbecco's modified Eagle medium (DMEM, GlutaMax, Thermo Fischer Scientific) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Thermo Fischer Scientific). The parental cell line was grown with the addition of zeocin (50µg/ml) (Invitrogen) and blasticidin (5µg/ml) (Thermo Fisher Scientific), while the stable cell lines were grown under the presence of blasticidin (5µg/ml) and hygromycin (200µg/ml) (Invitrogen).

The stable Flp-In T-Rex HeLa cell lines were generated using the X-tremeGENE 9 DNA Transfection Reagent (Sigma-Aldrich) following the manufacturer's instructions. Briefly, Flp-In T-Rex HeLa cells (passage 1) were seeded in 6-well plates (1.8 x 10^4 cells/well) in complete DMEM without antibiotics. After 24h, the transfection mixture was prepared [consisted of 100 µl Opti-MEM (Gibco), 3 µl of X-tremeGENE 9 DNA Transfection Reagent, 100 ng of plasmid with the construct of interest and 900 ng of pOG44 (Thermo Fisher Scientific)], incubated for 15 min at room temperature, and added in a drop-wise manner to the cells under gentle shaking. On the third day, the cells were expanded via trypsinization (Trypsin-EDTA, Gibco) into 150 mm plates, and the following day the medium was changed into complete DMEM with the addition of the selection antibiotics blasticidin (5µg/ml) and hygromycin (200µg/ml). The following weeks, the medium was replaced with fresh complete DMEM and antibiotics thrice per week. When large colonies were formed, the cells were trypsinized into the same plates and when 80% confluency was reached they were prepared for cryopreservation. For the cryopreservation, the cells were washed with PBS (Thermo Fischer Scientific), trypsinized, centrifuged at 500g for 5 min and re-suspended in DMEM supplemented with 20% fetal bovine serum and 10% DMSO (Sigma-Aldrich). The vials were placed into a freezing container (Mr. Frosty™, Thermo Fischer Scientific) for 24h at -80^{o}C followed by transfer into a liquid nitrogen tank for long term storage.

Table *2:* Primers for generation of the engineered constructs and oligos

Template	Primers (5’-3’)	Final Construct/Oligo
pcDNA3.1-BioID2-HA	5’ACAAAGTGGTTGATTCTAGATTCAAGAACCTGATCTGGCT3’ 5’TCGTCATCGTCTTTGTAGTCGCTTCTTCTCAGGCTGAACT3’	BioID2-FLAG
pcDNA3.1-BioID2-HA	5’CGGATCCAGCTAGCGCCACCATGGACTACAAAGACGATGACGATAAATTCAAGAACCTGATCTGG3’ 5’GCTTTTTTGTACAAACTTGTGCTTCTTCTCAGGCTGAACTCGCCG3’	FLAG-BioID2
pcDNA5-pDEST-BioID-FLAG-C-term	5’AGTTCAGCCTGAGAAGAAGCGACTACAAAGACGATGACGA3’ 5’AGCCAGATCAGGTTCTTGAATCTAGAATCAACCACTTTGTACA3’	pcDNA5-pDEST---C-term
pcDNA5-pDEST-BioID-FLAG-N-term	5’AGTTCAGCCTGAGAAGAAGCACAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAAAATG3’ 5’AGTTCAGCCTGAGAAGAAGCACAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAAAATG3’	pcDNA5-pDEST---N-term

BioID labeling and cell harvesting

Flp-In T-Rex HeLa cells stably transfected with BioID2-FLAG-tagged constructs and wild type parental cells were seeded in 245mm plates (2.0 x 10^6 cells/plate) in complete DMEM without antibiotics. After 24h, the medium was supplemented with 0.13 µg/ml tetracycline (Sigma-Aldrich) for induction of protein expression. The following day, 13 µM biotin (Sigma-Aldrich) was added for protein biotinylation, and after 24h the cells were harvested. For the harvesting, the cells were washed with 20 ml PBS, detached with 10ml trypsin-EDTA and collected with 20ml complete medium. The collected cells were centrifuged at 500g for 5min at 4ºC, re-suspended in 15ml PBS (4ºC) and centrifuged again under the previous conditions. Once the supernatant was aspirated, the cell pellets were frozen on dry ice and stored at -80ºC until further processing. The final number of collected cells was 2x10^7 per plate. For each stable cell line, four replicates were collected.

Table 3: Sequencing primers for the engineered constructs

Final Constructs	Sequencing Primers
PDHB-BioID2-FLAG, IDH1- BioID2-FLAG, IDH2- BioID2-FLAG, IDH3G- BioID2-FLAG	5'CGCAAATGGGCGGTAGGCGTG3' 5'CTCACGTTCCACTCCTTCAG3'
FLAG-BioID2-PDHB, FLAG-BioID2-IDH1, FLAG-BioID2-IDH2, FLAG-BioID2-IDH3G,	5'GCATCAAGGAGATCCTGAGC3' 5'GGGCAAACAACAGATGGCTG3'
ACO2-BioID2-FLAG	5'CGCAAATGGGCGGTAGGCGTG3' 5'CTCACGTTCCACTCCTTCAG3' 5'CCTGAACTTCTTGCCATCCG3'
FLAG-BioID2-ACO2	5'GCATCAAGGAGATCCTGAGC3' 5'GGGCAAACAACAGATGGCTG3' 5'CCTGAACTTCTTGCCATCCG3'
OGDH-BioID2-FLAG	5'CGCAAATGGGCGGTAGGCGTG3' 5'CTCACGTTCCACTCCTTCAG3' 5'GCCACTTTGCACACGTACATG3' 5'GCAGATGAACTGGTCGATG3'
FLAG-BioID2-OGDH	5'GCATCAAGGAGATCCTGAGC3' 5'GGGCAAACAACAGATGGCTG3' 5'GCCACTTTGCACACGTACATG3' 5'GCAGATGAACTGGTCGATG3'

Affinity purification of biotinylated proteins

The affinity purification of the biotinylated proteins was adapted from the method developed by Mackmull et al 2017[157] with slight modifications. The frozen cell pellets were re-suspended in 8ml of lysis buffer (4ºC) [50 mM Tris pH 7.5, 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 1 mM EDTA, 1mM EGTA, 1 mM PMFS (Sigma-Aldrich), 1 mg/ml Aprotinin (Carl-Roth), 0.5 mg/ml Leupeptin (Carl-Roth), 250 Units HS-Nuclease (Accelagen)] and incubated for 1h at 4ºC under constant mild rotation (30 rpm). The cell lysates were sonicated at 4ºC for 30 sec x 5 times with 30 sec rest in-between, and subsequently centrifuged at 17000g for 30 min at 4ºC to remove any insoluble material. 80 µl of Streptavidin Sepharose High Performance beads (GE Healthcare) were equilibrated in 1ml of lysis buffer for 30 min at 4ºC under constant mild rotation (30 rpm). The equilibrated beads were centrifuged at 2000g for 5 min at 4ºC, then transferred to the lysed supernatants and incubated for 3h at 4ºC under constant mild rotation (30 rpm). After the incubation, the beads were centrifuged at 2000g for 5 min at 4ºC and 7.5 ml of the supernatant were discarded. The remaining beads-lysates were transferred to a Spin Column (Pierce, Thermo Fischer Scientific), washed once with 800 µl of lysis buffer and then five times with 700 µl of 50 mM ammonium bicarbonate pH 8.3. After the washing, the column was plugged and the beads were transferred to a fresh eppendorf tube with 300 µl of 50 mM ammonium bicarbonate pH 8.3. The same procedure was repeated two more times to ensure all beads were collected. One µg of Sequencing Modified Trypsin (Promega) was added and the samples were incubated at 37 °C for 16 hr under constant shaking (500 rpm). The following day, 0.5 µg of trypsin were added and the beads were incubated for 2 additional hours under the same conditions. Following incubation, the beads were transferred to a new Spin Column and the digested peptides were eluted with two times with 150 µl of 50 mM ammonium bicarbonate pH8. The eluted peptides were dried in a speed-vac and stored at -80ºC until further processing. All steps were performed using low protein binding tips (Starlab) and low protein binding eppendorf tubes (Thermo Fischer Scientific).

Sample preparation and TMT labeling for mass spectrometry[2]

Dried samples were dissolved in 1% formic acid with 4% acetonitrile and subjected to OASIS® HLB µElution Plate (Waters) for desalting according to manufacturer's instructions. Desalted peptides were reconstituted in 50 mM HEPES (pH 8.5) and labelled with TMT10plex[158] Isobaric Label Reagent (Thermo Fischer Scientific) according the manufacturer's instructions. For further sample clean up an OASIS® HLB µElution Plate (Waters) was used. Offline high pH reverse phase fractionation was carried out on an Agilent 1200 Infinity high-performance liquid chromatography system, equipped with a Gemini C18 column (3 µm, 110 Å, 100 x 1.0 mm, Phenomenex)[159].

Mass spectrometry data acquisition[2]

An UltiMate 3000 RSLC nano LC system (Dionex) fitted with a trapping cartridge (µ-Precolumn C18 PepMap 100, 5µm, 300 µm i.d. x 5 mm, 100 Å) and an analytical column (nanoEase™ M/Z HSS T3 column 75 µm x 250 mm C18, 1.8 µm, 100 Å, Waters). Trapping was carried out with a constant flow of solvent A (0.1% formic acid in water) at 30 µL/min onto the trapping column for 6 minutes. Subsequently, peptides were eluted via the analytical column with a constant flow of 0.3 µL/min with increasing percentage of solvent B (0.1% formic acid in acetonitrile) from 2% to 4% in 4 min, from 4% to 8% in 2 min, then 8% to 28% for a further 96 min, and finally from 28% to 40% in another 10 min. The outlet of the analytical

---

[2] The following method part was provided from Mandy Rettel and Frank Stein, Proteomics Core Facility, EMBL, Heidelberg, Germany

column was coupled directly to a QExactive plus (Thermo Fischer Scientific) mass spectrometer using the proxeon nanoflow source in positive ion mode.

The peptides were introduced into the QExactive plus via a Pico-Tip Emitter 360 µm OD x 20 µm ID; 10 µm tip (New Objective) and an applied spray voltage of 2.3 kV. The capillary temperature was set at 320°C. Full mass scan was acquired with mass range 350-1400 m/z in profile mode in the FT with resolution of 70000. The filling time was set at maximum of 100 ms with a limitation of $3x10^6$ ions. Data dependent acquisition (DDA) was performed with the resolution of the Orbitrap set to 35000, with a fill time of 120 ms and a limitation of $2x10^5$ ions. A normalized collision energy of 32 was applied. A loop count of 10 with count 1 was used and a minimum AGC trigger of $2e^2$ was set. Dynamic exclusion time of 30 s was used. The peptide match algorithm was set to ‘preferred’ and charge exclusion ‘unassigned’, charge states 1, 5 - 8 were excluded. MS^2 data were acquired in profile mode[160].

Mass spectrometry data analysis[2]

IsobarQuant[161] and Mascot (v2.2.07) were used to process the acquired data, which was searched against a Uniprot Homo sapiens proteome database (UP000005640) containing common contaminants and reversed sequences. The following modifications were included into the search parameters: Carbamidomethyl (C) and TMT10 (K) (fixed modification), Acetyl (N-term), Oxidation (M) and TMT10 (N-term) (variable modifications). For the full scan (MS1) a mass error tolerance of 10 ppm and for MS/MS (MS2) spectra of 0.02 Da was set. Further parameters were set: Trypsin as protease with an allowance of maximum two missed cleavages: a minimum peptide length of seven amino acids; at least two unique peptides were required for a protein identification. The false discovery rate on peptide and protein level was set to 0.01.

Statistical Analysis and Identification of proximal interactors[2]

Raw data of IsobarQuant were loaded into R (ISBN 3-900051-07-0). As a quality criteria, only proteins which were quantified with at least two different unique peptides were used for downstream analysis. The “signal_sum” columns of the “proteins”-output sheet from IsobarQuant were cleaned for potential batch-effects with limma[162] and subsequently normalized with vsn (variance stabilization)[163]. Missing values were imputed with the impute function (method = “knn”) from the MSNBase package[164]. Limma was employed again to test for differential abundance. Each engineered cell line (stable cells with inducible expression of the protein of interest fused to biotin ligase) was compared to wild type HeLa and a protein was considered a proximal interactor with a $\log_2$ fold change ≥ 0.6 and a (Benjamini-Hochberg) adjusted p value ≤ 0.1. These results were further refined by removing known contaminants from the CRAPome[165] database detected in experiments performed on HeLa cell lines with a cutoff frequency ≥ 20% - proteins measured with a $\log_2$ fold change ≥ 1.32 (engineered HeLa to wild type HeLa) were kept as proximal interactors. To further exclude unspecific proteins as a result of the exogenous expression of the biotin ligase, the engineered cells were compared to the engineered cells with N-fused IDH2, and only the proteins detected as significant (using the same criteria as above) in this comparison in addition to the comparison with wild type HeLa were kept in the list of proximal interactors. To define the higher specificity interactome each engineered cell line was compared to wild type parental HeLa, N-fused IDH2 and C-fused IDH2 or C-fused PDHB engineered cell lines, and the common significant proteins in all comparisons were included in the list of proximal interactors.

Gene ontology enrichment analysis

For the gene ontology enrichment analysis, the g:Profiler[112,113] (version e94_eg41_p11_ff2a873) was used with the following parameters: significance threshold, Benjamini-Hochberg FDR; user threshold, 0.05; statistical domain scope, custom, using the entire proteomics dataset as a background list.

Dotplot generation

For the dotplot generation, the ProHits-viz[166] platform was used for the higher specificity interactome. The data contained the $\log_2$fold changes of each detected biotinylated protein as derived from the comparisons between each engineered cell line to the wild type HeLa cells along with the respective adjusted p values estimated by limma[162] analysis (see above). The following parameters were used: score type, adjusted p value defined by limma analysis; primary filter, 0.01; secondary filter, 0.1; no clustering was utilized.

Cytoscape visualization

Cytoscape[123] 3.6.1 was used for the visualization of the higher specificity interactome and GO enrichment analysis was performed utilizing the ClueGO[122] plug-in. The uploaded data file contained the $\log_2$fold changes of each detected biotinylated protein as derived from the comparisons between each engineered cell line to the wild type HeLa cells along with the respective adjusted p values estimated by limma[162] analysis (see above).

Immunofluorescence and Imaging

For immunofluorescence analysis, cells were seeded in 8-well Nunc Lab-Tek II coverglass chambers (Sigma-Aldrich) (2700 cells/well in 300 µl medium/well) and the induction of protein expression and biotinylation was performed as described before. The cells were washed thrice with PBS, fixed with 4% formaldehyde (Image-iT, Invitrogen) for 10 min, followed by three washing steps with PBS. Subsequently, the cells were permeabilized with 0.2% Triton-X100 for 20 min, washed thrice with PBS and blocked (0.1% BSA, 0.3M glycine, 0.1% Tween-20 in PBS) for 1h at room temperature. Incubation with primary antibodies in blocking solution was performed overnight in a humidified chamber at 4ºC: the biotinylated proteins were detected with the streptavidin Alexa Fluor 488 (Thermo Fischer Scientific, S11223), the bait proteins with mouse anti-FLAG (Sigma-Aldrich, F1804) as well as with specific antibodies that bound to the endogenous proteins. The following day, the cells were washed three times with PBS and incubated with secondary antibodies in blocking solution for 1hr at room temperature in a dark room. Following three washing steps in PBS, the cells were stained with Hoechst (2.6 µg/ml) (Thermo Fischer Scientific, 62249) for 20 min at room temperature in the dark, then washed three times with PBS and finally saved in PBS with 0.02% sodium azide (Sigma-Aldrich) at 4ºC. Sample images were acquired at the Advanced Light Microscopy Facility at EMBL, Heidelberg with a Leica TCS SP8 DLS system using a 63x oil objective, and analyzed with LAS X (Leica) and ImageJ software. A list of primary and secondary antibodies along with the respective dilutions and combinations is presented in Table 4.

Table 4: Immunofluorescence Antibodies Conditions

Cell line	Primary Incubation	Secondary Incubation
PDHB-BioID2, OGDH-BioID2, IDH3G-BioID2, Parental HeLa	streptavidin Alexa Fluor 488 (1:100) and mouse anti-FLAG (1:100)	anti-mouse Alexa Fluor 647 (1:200) (Thermo Fischer Scientific, A28181)
	streptavidin Alexa Fluor 488 (1:100) and rabbit anti-OGDH (1:100) (Sigma-Aldrich, HPA020347)	anti-rabbit Alexa Fluor 647 (1:200) (Thermo Fisher Scientific, A27040)
	streptavidin Alexa Fluor 488 (1:100) and rabbit anti-IDH3G (1:400) (Sigma-Aldrich, HPA002017)	anti-rabbit Alexa Fluor 647 (1:800)
ACO2-BioID2, IDH2-BioID2, Parental HeLa	streptavidin Alexa Fluor 488 (1:500) and mouse anti-FLAG (1:100)	anti-mouse Alexa Fluor 647 (1:200)
	streptavidin Alexa Fluor 488 (1:500) and rabbit anti-Aconitase 2 (1:500) (Abcam, ab129069)	anti-rabbit Alexa Fluor 647 (1:1000)
IDH1-BioID2, Parental HeLa	streptavidin Alexa Fluor 488 (1:1000) and mouse anti-FLAG (1:100)	anti-mouse Alexa Fluor 647 (1:200)
	streptavidin Alexa Fluor 488 (1:1000)	-

Immunoblotting

For immunoblotting, cell samples or isolated nuclei were collected as described in the respective sections above. The samples were lysed in sample buffer (62.5 mM Tris pH 6.8, 2% SDS, 10% glycerol, 0.0006% Bromophenol Blue, 5% β-Mercaptoethanol), incubated at 95ºC for 5 min and sonicated to shear the DNA. The proteins were separated on 4–20%

gradient gels (Mini-PROTEAN® TGX™ Precast Protein Gels, Biorad). Prior transfer, the gel, the nitrocellulose membrane (Sigma-Aldrich) and the blot filter papers (Biorad) were equilibrated in transfer buffer (25 mM Tris, 192 mM glycine pH 8.3, 20% methanol, 0.04 % SDS) for 15 min with agitation. For the transfer, the Trans-Blot Turbo System (Biorad) with the STANDARD SD protocol followed by the Mixed MW program was used. The membrane was then blocked in 1% BSA in PBST (PBS, 0.02% Tween-20) for 1h at room temperature followed by overnight incubation at 4ºC with primary antibodies diluted in 1% BSA in PBST under agitation. The following day, the membrane was washed thrice with PBST, incubated for 1h at room temperature with the appropriate secondary antibody in 1% BSA in PBST, and then washed thrice with PBST. Finally, the membrane was developed using the Pierce ECL Plus Western Blotting Substrate (Thermo Fischer Scientific) and imaged in a ChemiDoc MP Imaging System (Biorad). The antibodies used for immunoblot of the stable cell lines were the following: anti-Biotin HRP-linked (1:3000) (Cell Signaling Technology, #7075), streptavidin HRP-linked (1:3000) (Thermo Fischer Scientific, 21130), mouse-anti FLAG (1:1000) (Sigma-Aldrich, F1804) and anti-mouse HRP-linked (1:2000) (Abcam, ab205719). For the nuclei immunoblots, the following antibodies were used: rabbit anti-IDH2 (1:1000) (Abcam, ab129180), rabbit anti-COXIV (1:1000) (Cell Signaling Technology, #8674), rabbit anti-cytochrome c (1:1000) (Cell Signaling Technology, #8674), rabbit anti-beta III tubulin (1:1000) (Abcam, ab18207) and anti-rabbit HRP-linked (1:2000) (Abcam, ab205718).

Chapter II and Chapter III – Materials and Methods

Three dimensional cell cultures[3]

Three-dimensional cell cultures were established according to the published protocol[167] 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 37oC in 5 % (vol/vol) CO2 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 37oC, 5 % CO2 in loosely capped tubes,

---

[3] The following method part was written by Ksenija Radić Shechter

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 37oC, 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.

Three dimensional cell cultures experiments - Cell and Spent Growth Media Collection and Metabolite Extraction[4]

Organoid structures were 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 at 1000 rpm for 2 min, at room temperature, and quenched with 200 µl cold (-80°C) HPLC-grade methanol (Biosolve). Samples were stored at -80°C until metabolite extraction. For metabolite extraction, 10µl of ribitol (50µg/ml) (Alfa Aesar, UK) was added to the cell/methanol mixture and the samples were vortexed for 2 minutes. Subsequently, the samples were incubated at 72°C for 15min, followed by addition of 200 µl ice cold MilliQ water and vortex for 2 minutes. The samples were centrifuged at 15000rpm for 10 minutes at 4°C, the supernatants were carefully transferred in

---

[4] The following method part is based on the method described in Havas and colleagues[131] where it was originally written by me

glass vials (Agilent, Santa Clara, CA), dried in a speed-vac, and stored at -80^{o}C until GC-MS analysis. 50µl of the spent growth medium of each well and the last washing solution of each sample were also collected and quenched with 100µl of cold (-80^{o}C) HPLC-grade methanol (Biosolve). Metabolite extraction from the spent growth media was performed as described above with the following difference: instead of 200 µl MilliQ water, 100µl were added.

In vivo mammary glands experiments[5]

For glucose labeling experiment, mammary glands were dissected, minced and digested for 2 hours at 37^{o}C using collagenase and liberase enzymes, then were cultured for 8 hours at 37^{o}C in 5 % (vol/vol) CO2 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 37oC 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 the method described in the above section.

Gas Chromatography - Mass Spectrometry (GC-MS) data acquisition and analysis[6]

Unlabelled samples. Dried polar metabolites were derivatized with 50 µl of 20 mg/mL methoxyamine hydrochloride (Alfa Aesar, UK) solution in pyridine for 90 min at 37^{o}C, followed by reaction with 100 µL N-methyl-trimethylsilyl-trifluoroacetamide (MSTFA) (Alfa Aesar, UK) for 10 hours at room temperature, as justified in[168]. GC-MS analysis was performed using a Shimadzu TQ8040 GC-(triple quadrupole) MS system (Shimadzu Corp.) equipped with a 30m x 0.25 mm x 0.25 µm DB-50MS capillary column (Phenomenex, USA). 1 µl of sample was injected in split mode (split ratio 1:20) at 250^{o}C using helium as a carrier gas with a flow rate of 1 ml/min. GC oven temperature was held at 100^{o}C for 4 min followed by an increase to 320^{o}C with a rate of 10^{o}C/min, and a final constant temperature period at 320^{o}C for 11 min. The interface and the ion source were held at 280^{o}C and 230^{o}C, respectively. The detector

---

[5] The following method part is based on Ksenija Radić Shechter, PhD thesis

[6] The following method part is based on the methods described on Strucko and colleagues[160], and Blasche and colleagues[169] where it was originally written by me

was operated both in scanning mode recording in the range of 50-600 *m/z*, as well as in MRM mode for specified metabolites. The metabolite identification was based in an in-house database were standards were used to define the retention time and the mass spectrum for all mentioned metabolites. Both information were subsequently used to identify the metabolites in the biological samples. The metabolite quantification was carried out by calculating the area under the curve (AUC) of the identified marker ion of each metabolite normalized to the area of ribitol's marker ion 319. Subsequently, areas were normalized to total metabolite levels. To identify the statistically significant altered metabolites limma[162] analysis was performed, and the statistically significant metabolites were defined using the following criteria: Benjamini-Hochberg FDR adjusted p value ≤ 0.01 and a $\log_2$ fold change (condition/control) ≥ 1or ≤ -1.

13C labelled samples. For the 13C isotope tracing experiments, the dried polar metabolites were derivatized with 50 µl of 20 mg/mL methoxyamine hydrochloride (Alfa Aesar, UK) solution in pyridine for 90 min at 37°C, followed by reaction with 100 µL N-tert-Butyldimethylsilyl-N-methyltrifluoroacetamide with 1 % tert- Butyldimethylchlorosilane (TBDMS) (Sigma-Aldrich) for 1h at 60°C. The samples remained at room temperature until GC-MS analysis. The GC-MS was operated using the same conditions as described in the section above, with the following difference: GC oven temperature was held at 100°C for 3 min followed by an increase to 300°C with a rate of 3.5°C/min, and a final constant temperature period at 300°C for 10 min. The metabolite quantification was carried out by calculating the area under the curve (AUC) of the identified marker ions of each metabolite, and the natural abundance isotopes were corrected using the Isotope Correction Toolbox (ICT)[156].

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).

Two dimensional cell cultures experiments - Cell and Spent Growth Media Collection and Metabolite Extraction

SNU-389 cells seeded on 10 cm dishes (90000 c/cm^2), after 6 h they were starved overnight in RPMI 1640 (Gibco #31870), 2 mM L-Glutamine, 1% FCS. On the next day medium was changed to assay medium RPMI 1640 (Gibco #31870), 2 mM L-Glutamine, 5 % FCS and cells were treated with 0.3% DMSO, 2.5 µM LDHi, 1 µM MCTi, 2.5 µM LDHi+1 µM MCTi, for 1 and 4 h. For isotope tracing experiments, the same conditions were used as before utilizing [U-^{13}C]-glucose and the [U-^{13}C]-glutamine labeled media, and cells were collected for

1h, 2h, 4h, 6h and 8h after inhibitors treatment. Prior to cell collection, 2ml of spent growth medium was collected for extracellular metabolomics. Subsequently, the growth medium was removed, the cells were washed thrice with 10 ml ice cold PBS. The cells were then quenched using 2 ml ice-cold HPLC-grade methanol (Biosolve). Next, the cells were scraped off the plate, another 2 ml of ice-cold HPLC-grade methanol was added to collect any remaining cells and the samples were stored at -80^oC until metabolite extraction. For metabolite extraction, 10µl of ribitol (50µg/ml) (Alfa Aesar, UK) was added to the cells/methanol mixture and the samples were vortexed for 2 minutes. Subsequently, the samples were incubated at 72^oC for 15min, followed by addition of 4 ml ice cold MilliQ water and vortex for 2 minutes. The samples were centrifuged at 15000rpm for 10 minutes, at 4^oC, the supernatants were carefully transferred in glass vials (Agilent, Santa Clara, CA), dried in a speed-vac, and stored at -80^oC until GC-MS analysis. For the metabolomics acquisition and analysis, see the sections before. The contribution of glucose to TCA cycle was calculated as the ratio of total carbon contribution (FC) in citrate to total carbon contribution (FC) in pyruvate. Total carbon contribution (FC) was calculated using the following equation: $\sum_{i=0}^{n} i * m\,i / (n * \sum_{i=0}^{n} mi)$, where n is the number of carbons in a metabolite, i the different mass isotopomers and m the abundance of a certain mass.

List of figures

Figure 1: Paradigms of subcellular localization of metabolic enzymes related to the production of epigenetic relevant metabolites. 24
Figure 2: The nuclear metabolic proteome. 36
Figure 3: Nuclear Proteome and Nuclear Localization Signal motifs suggest the potentiality of a nuclear subnetwork of TCA cycle enzymes.. 37
Figure 4: Experimental design of the nuclei labeling with ^{13}C metabolic substrates. 39
Figure 5: Nuclei purity assessment 40
Figure 6: Glutamine is a precursor of key TCA intermediates in the HeLa nucleus. 42
Figure 7: [U-13C]-citrate supplementation reconstructs a citrate to succinate axis actively present in HeLa nuclei.. 45
Figure 8: Proximity biotinylation mass spectrometry based on the BioID method. 47
Figure 9: Exploration of the proximal interactomes of ACO2, OGDH and IDH3G.. 49
Figure 10: Subcellular distribution of the proximal interactors. 51
Figure 11: Proximity biotinylation mass spectrometry reveals a potential nuclear niche for ACO2, OGDH and IDH3G coupled to histones, chromatin binding and RNA modifications. 52
Figure 12: Immunofluorescence microscopy of the engineered cell lines expressing the protein of interest fused to biotin ligase 2. 53
Figure 13: Immunofluorescence microscopy of the engineered cell lines expressing the protein of interest fused to biotin ligase 2. 54
Figure 14: Proposed nuclear metabolic network and relevant functions. 56
Figure 15: Residual cells are phenotypically similar to their normal counterparts. 60
Figure 16: Tumor and residual cells exhibit divergent response at the transcriptional and metabolic level. 61
Figure 17: The metabolic profile of residual cells is more similar to the tumor than the control state.. 62
Figure 18: The metabolic profile of residual cells is more similar to the tumor than the control state. 64
Figure 19: Metabolic alterations in the residual and tumor population. 65
Figure 20: Global overview of central carbon metabolic alterations of the residual population compared to the control state. 68
Figure 21: Global overview of central carbon metabolic alterations of the tumorous population compared to the control state. 69
Figure 22: Ex vivo and in vivo validation supports the altered glycolysis and urea cycle metabolic signatures in the residual population. 71
Figure 23: Metabolic alterations after inhibition of lactate dehydrogenase and monocarboxylate transporters 1 and 2. 79
Figure 24: Overview of metabolic effects on glycolysis and TCA cycle upon inhibition of lactate dehydrogenase and monocarboxylate transporters 1 and 2. 80
Figure 25: Labeling kinetics of pyruvate and lactate upon treatment with the lactate dehydrogenase inhibitor or the monocarboxylate transporters 1 and 2 inhibitor. 81
Figure 26: Effects on glucose's contribution to the central carbon metabolic pathways glycolysis and TCA cycle upon treatment with the lactate dehydrogenase inhibitor or the monocarboxylate transporters 1 and 2 inhibitor.. 82
Figure 27: Effects on glutamine's contribution to TCA cycle after treatment with the lactate dehydrogenase inhibitor or the monocarboxylate transporters 1 and 2 inhibitor. 84
Figure 28: Metabolic alterations upon dual inhibition of lactate dehydrogenase and monocarboxylate transporters 1 and 2 activities.. 85

Figure 30: Mass isotopomers analysis for metabolic intermediates in cell lysates incubated with [U-^{13}C]-pyruvate for 5 hours. 119
Figure 31: Pairwise pearson correlation analysis of the proximal interactome of TCA cycle enzymes. 120
Figure 32: Metabolic alterations in control (never induced and wild-type), tumor and residual cells. 121
Figure 33: Glutamine and glucose uptake changes. 122

List of Tables

Table 1: Destination vectors and entry clones used for the generation of the engineered constructs 98
Table 2: Primers for generation of the engineered constructs and oligos 99
Table 3: Sequencing primers for the engineered constructs 100
Table 4: Immunofluorescence Antibodies Conditions 104
Table 5: Top GO term for cellular compartments of the proximal interactors of the engineered proteins 123
Table 6: Overview of higher specificity proximal interactors (Chapter I) 124

Abbreviations

ACAA2	acetyl-Coenzyme A acyltransferase 2
ACAT2	acetoacetyl-CoA thiolase
AcCoA	acetyl-coenzyme A
ACLY	ATP citrate lyase
ACO1	aconitase 1
ACO2	aconitase 2
akg	alpha-ketoglutarate
akgM3	alpha-ketoglutarate M3 isotopomer
akgM5	alpha-ketoglutarate M5 isotopomer
asp	aspartate
cit	citrate
citM2	citrate M2 isotopomer
citM3	citrate M3 isotopomer
citM4	citrate M4 isotopomer
citM5	citrate M5 isotopomer
CS	citrate synthase
FCcit	total carbon contribution in citrate;
FCpyr	total carbon contribution in pyruvate;
FH	fumarate hydratase
fum	fumarate
GC-MS	Gas Chromatography – Mass Spectrometry
glnM5	glutamine M5 isotopomer

glu	glutamate
GOT1	glutamic-oxaloacetic transaminase 1
GOT2	glutamic-oxaloacetic transaminase 2
HADHB	trifunctional enzyme subunit beta
IDH1	isocitrate dehydrogenase 1 (NADP+)
IDH2	isocitrate dehydrogenase 2 [NADP+]
IDH3A	isocitrate dehydrogenase 3 [NAD+] subunit alpha
IDH3B	isocitrate dehydrogenase 3 [NAD+] subunit beta
IDH3G	isocitrate dehydrogenase 3 [NAD+] subunit gamma
lac	lactate
mal	malate
malM4	malate M4 isotopomer
MDH1	malate dehydrogenase 1
MDH2	malate dehydrogenase 2
ME1	malic enzyme 1
ME2	malic enzyme 2
ME3	malic enzyme 3
N-ac-asp	N-acetyl-aspartate
N-ac-aspM2	N-acetyl-aspartate M2 isotopomer
oaa	oxaloacetate
OGDC	oxoglutarate dehydrogenase complex
PDC	pyruvate dehydrogenase complex
PYC	pyruvate carboxylase
pyr	pyruvate
pyrM3	pyruvate M3 isotopomer

suc	succinate
SDHA	Succinate dehydrogenase subunit A
SDHB	succinate dehydrogenase subunit B
SUCLA2	succinyl-CoA ligase [ADP-forming] subunit beta
SUCLG1	succinyl-CoA ligase [GDP-forming] subunit alpha
SUCLG2	succinyl-CoA ligase [GDP-forming] subunit beta
sucM2	succinate M2 isotopomer
sucM4	succinate M4 isotopomer
TCA cycle	tricarboxylic acid cycle

Publications

Blasche S, Kim Y, Mars R, **Kafkia E**, Maansson M, Machado D, Teusink B, Nielsen J, Benes V, Neves R, Sauer U, Patil KR. (2019) **Emergence of stable coexistence in a complex microbial community through metabolic cooperation and spatio-temporal niche partitioning.** *BioRxiv,* p. 541870.

Strucko T, Zirngibl K, Pereira F, **Kafkia E**, Mohamed ET, Rettel M, Stein F, Feist AM, Jouhten P, Patil KR, Forster J. (2018) **Laboratory evolution reveals regulatory and metabolic trade-offs of glycerol utilization in Saccharomyces cerevisiae.** *Metab Eng* 47:73-82.

Ponomarova O, Gabrielli N, Sévin DC, Mülleder M, Zirngibl K, Bulyha K, Andrejev S, **Kafkia E**, Typas A, Sauer U, Ralser M, Patil KR. (2017) **Yeast Creates a Niche for Symbiotic Lactic Acid Bacteria through Nitrogen Overflow.** *Cell Syst* S2405-4712(17):30390-30393.

Havas KM, Milchevskaya V, Radic K, Alladin A, **Kafkia E**, Garcia M, Stolte J, Klaus B, Rotmensz N, Gibson TJ, Burwinkel B, Schneeweiss A, Pruneri G, Patil KR, Sotillo R, Jechlinger M. (2017) **Metabolic shifts in residual breast cancer drive tumor recurrence.** *J Clin Invest* 127(6):2091-2105.

Appendix

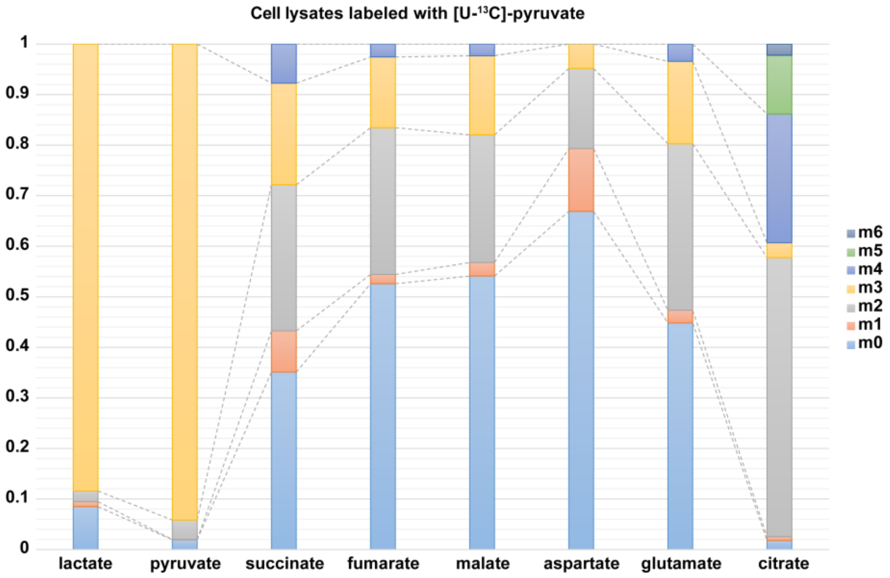

Figure 29: Mass isotopomers analysis for metabolic intermediates in cell lysates incubated with [U-^{13}C]-pyruvate for 5 hours. The y axis indicate the percent of labeling. M0 to Mn indicates the different mass isotopomers (M0 denotes that zero carbons are labeled, and Mn denotes that n number of carbons are labeled) (The figure is related to Chapter I).

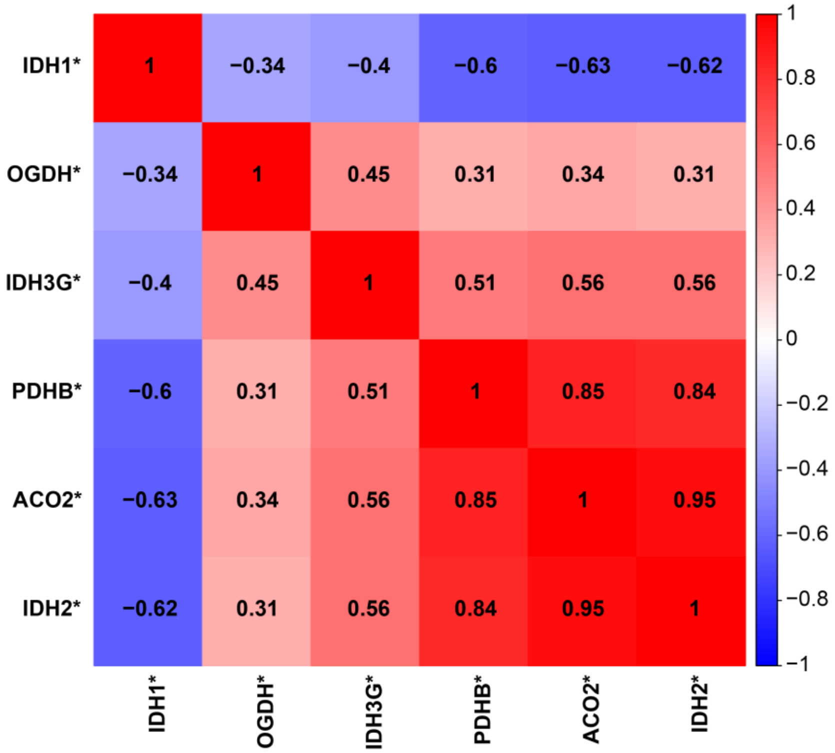

Figure 30: Pairwise Pearson correlation analysis of the proximal interactome of TCA cycle enzymes (the figure is related to Chapter I).

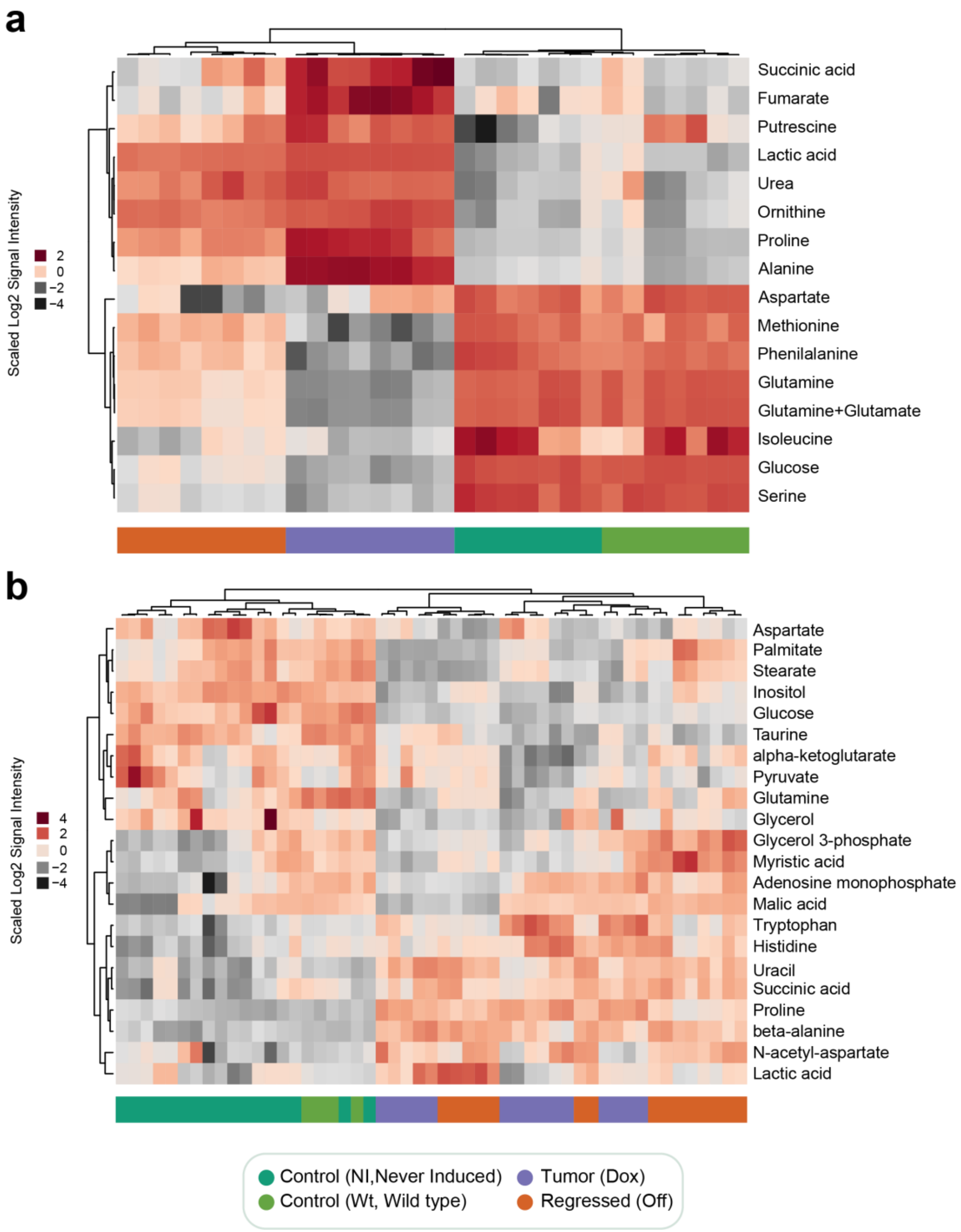

Figure 31: Metabolic alterations in control (never induced and wild-type), tumor and residual cells. Hierarchical clustering (Pearson correlation) on significantly altered metabolites in residual and tumor populations compared to the control state. (a) extracellular spent growth media samples, and (b) intracellular samples. The results are based on targeted GC-MS measurements (the figure is related to Chapter II).

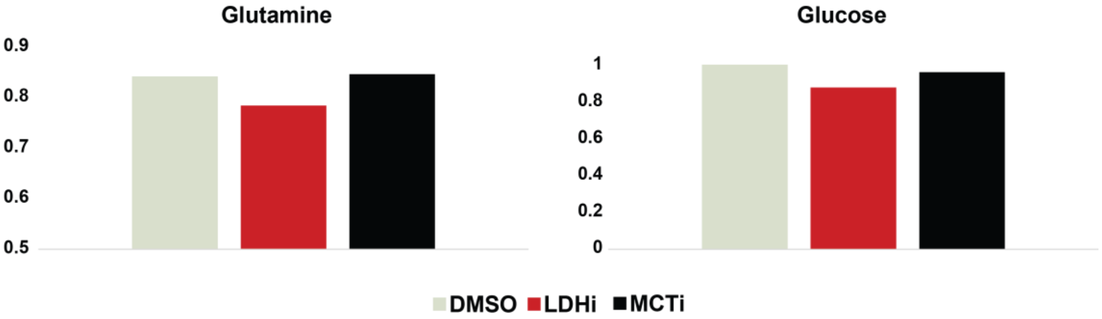

Figure 32: Glutamine and glucose uptake changes after treatment with the LDHA inhibitor (LDHi) or MCT1/2 inhibitor (MCTi) compared to the control cells (DMSO-treated). The barplots represent the fold change of glutamine or glucose extracellular concentration levels at 6h to 1h & 2h time points. One representative experiment is depicted (the figure is related to Chapter III).

Table 5: Top GO term for cellular compartments of the proximal interactors of the engineered proteins

Cellular Compartment	GO term ID	adjusted p value	Constructs
mitochondrion	GO:0005739	2.31E-142	PDHB*
mitochondrion	GO:0005739	1.07E-144	ACO2*
mitochondrial part	GO:0044429	1.31E-59	IDH3G*
mitochondrial protein complex	GO:0098798	3.56E-34	OGDH*
mitochondrion	GO:0005739	2.86E-153	IDH2*
cytosol	GO:0005829	1.19E-16	IDH1*
chaperone complex	GO:0101031	5.50E-10	*PDHB
cytoskeleton	GO:0005856	4.41E-11	*OGDH
outer mitochondrial membrane protein complex	GO:0098799	0.0004273	*IDH2

Table 6: Overview of higher specificity proximal interactors (Chapter I)

IDH3G*	OGDH*	ACO2*	IDH3G* & OGDH*	OGDH* & ACO2*	IDH3G* & OGDH* & ACO2*
IDH3G-bait	MRPS36	ACO2-bait	SRSF1	DLD	RBMX
IDH3A	SDF4	PDK2	SSRP1	DLST	TOP2B
IDH3B	SNRPD3	ATP5D	SUPT16H		HMGA1
RUVBL1	TMOD3	PDK3	GAR1		
RUVBL2	NOL11	PDK1	TRA2B		
NXN	RPL7L1	MGME1	DDX27		
CCT8	NOP58	MDH2	UBTF		
IDH1-bait	SMARCA5	PREPL	HNRNPDL		
UTP14A	KRR1	HSPE1	PARP1		
PSMB6	DKC1	MRPL28	FTSJ3		
PABPN1	EBNA1BP2	MRPS16	HNRNPA1		
	PUM3	MRPL57	NOP2		
	PDCD11		RRS1		
	CHD1		HNRNPA2B1		
	DNTTIP2		NIFK		
	RPF2		NOP56		
	YTHDC1		CBX5		
	MYO1C		DHX9		
	NSA2		HNRNPUL2		
	HEATR1		TOP2A		
	TBL3		HIST1H1B		
	DDX18		RALY		
	SAP18		SNU13		
	WDR12		SRSF10		
	BRIX1		HNRNPA3		
	FBL		MKI67		
	MYH9		BAZ1B		
	MYL6		HNRNPC		
	ACTC1		H2AFZ|H2AFV		
	HIST1H1E		PSIP1		
	RBM12B		SAFB		
	UTP18		HIST1H4A		
	WDR46		SLTM		
			RSL1D1		
			H2AFY		
			PHIP		
			SAFB2		

Bibliography

1. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961).
2. Chantranupong, L., Wolfson, R. L. & Sabatini, D. M. Nutrient-sensing mechanisms across evolution. *Cell* **161**, 67–83 (2015).
3. Sabari, B. R., Zhang, D., Allis, C. D. & Zhao, Y. Metabolic regulation of gene expression through histone acylations. *Nat. Rev. Mol. Cell Biol.* **18**, 90–101 (2017).
4. Farrelly, L. A. *et al.* Histone serotonylation is a permissive modification that enhances TFIID binding to H3K4me3. *Nature* (2019). doi:10.1038/s41586-019-1024-7
5. Kaelin, W. G. & McKnight, S. L. Influence of metabolism on epigenetics and disease. *Cell* **153**, 56–69 (2013).
6. Gut, P. & Verdin, E. The nexus of chromatin regulation and intermediary metabolism. *Nature* **502**, 489–498 (2013).
7. Wellen, K. E. *et al.* ATP-Citrate Lyase Links Cellular Metabolism to Histone Acetylation. *Science (80-.).* **324**, 1076–1080 (2009).
8. Lee, J. V *et al.* Akt-Dependent Metabolic Reprogramming Regulates Tumor Cell Histone Acetylation. *Cell Metab.* **20**, 306–319 (2014).
9. Zhao, Y. *et al.* The rate of glycolysis quantitatively mediates specific histone acetylation sites. *Cancer Metab.* **3**, 1–12 (2015).
10. Schug, Z. T. *et al.* Acetyl-CoA synthetase 2 promotes acetate utilization and maintains cancer cell growth under metabolic stress. *Cancer Cell* **27**, 57–71 (2015).
11. Bulusu, V. *et al.* Acetate Recapturing by Nuclear Acetyl-CoA Synthetase 2 Prevents Loss of Histone Acetylation during Oxygen and Serum Limitation. *Cell Rep.* **18**, 647–658 (2017).
12. Mews, P. *et al.* Acetyl-CoA synthetase regulates histone acetylation and hippocampal memory. *Nature* **546**, 381–386 (2017).
13. Cai, L., Sutter, B. M., Li, B. & Tu, B. P. Acetyl-CoA Induces Cell Growth and Proliferation by Promoting the Acetylation of Histones at Growth Genes. *Mol. Cell* **42**, 426–437 (2011).
14. Shiraki, N. *et al.* Methionine Metabolism Regulates Maintenance and Differentiation of Human Pluripotent Stem Cells. *Cell Metab.* **19**, 780–794 (2014).
15. Shyh-Chang, N. *et al.* Influence of Threonine Metabolism on S-Adenosylmethionine and Histone Methylation. *Science (80-.).* **339**, 222–226 (2013).
16. Carey, B. W., Finley, L. W. S., Cross, J. R., Allis, C. D. & Thompson, C. B. Intracellular α-ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* **518**, 413–416 (2015).
17. TeSlaa, T. *et al.* α-Ketoglutarate Accelerates the Initial Differentiation of Primed Human Pluripotent Stem Cells. *Cell Metab.* **24**, 485–493 (2016).
18. Xiao, M. *et al.* Inhibition of α-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev.* **26**, 1326–1338 (2012).

19. Shah, H. *et al.* L-2-Hydroxyglutarate production arises from noncanonical enzyme function at acidic pH. *Nat. Chem. Biol.* **13**, 494–500 (2017).

20. Roundtree, I. A., Evans, M. E., Pan, T. & He, C. Dynamic RNA Modifications in Gene Expression Regulation. *Cell* **169**, 1187–1200 (2017).

21. Thomas, J. M., Batista, P. J. & Meier, J. L. Metabolic Regulation of the Epitranscriptome. *ACS Chem. Biol.* (2019). doi:10.1021/acschembio.8b00951

22. Su, R. *et al.* R-2HG Exhibits Anti-tumor Activity by Targeting FTO/m6A/MYC/CEBPA Signaling. *Cell* **172**, 90-105.e23 (2018).

23. Ito, S. *et al.* A single acetylation of 18 S rRNA is essential for biogenesis of the small ribosomal subunit in Saccharomyces cerevisiae. *J. Biol. Chem.* **289**, 26201–12 (2014).

24. van Delft, P. *et al.* The Profile and Dynamics of RNA Modifications in Animals. *Chembiochem* **18**, 979–984 (2017).

25. Ito, S. *et al.* Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). *J. Biol. Chem.* **289**, 35724–30 (2014).

26. Montgomery, D. C. *et al.* Global Profiling of Acetyltransferase Feedback Regulation. *J. Am. Chem. Soc.* **138**, 6388–6391 (2016).

27. Katoh, Y. *et al.* Methionine Adenosyltransferase II Serves as a Transcriptional Corepressor of Maf Oncoprotein. *Mol. Cell* **41**, 554–566 (2011).

28. Li, S. *et al.* Serine and SAM Responsive Complex SESAME Regulates Histone Modification Crosstalk by Sensing Cellular Metabolism. *Mol. Cell* **60**, 408–421 (2015).

29. Sivanand, S. *et al.* Nuclear Acetyl-CoA Production by ACLY Promotes Homologous Recombination. *Mol. Cell* **67**, 252-265.e6 (2017).

30. Sutendra, G. *et al.* A Nuclear Pyruvate Dehydrogenase Complex Is Important for the Generation of Acetyl-CoA and Histone Acetylation. *Cell* **158**, 84–97 (2014).

31. Nagaraj, R. *et al.* Nuclear Localization of Mitochondrial TCA Cycle Enzymes as a Critical Step in Mammalian Zygotic Genome Activation. *Cell* **168**, 210-223.e11 (2017).

32. Matsuda, S. *et al.* Nuclear pyruvate kinase M2 complex serves as a transcriptional coactivator of arylhydrocarbon receptor. *Nucleic Acids Res.* **44**, 636–647 (2016).

33. Lee, D., Lee, K.-C., Jung, S.-J., Roe, J.-H. & Seo, Y. Essential function of Aco2, a fusion protein of aconitase and mitochondrial ribosomal protein bL21, in mitochondrial translation in fission yeast. *FEBS Lett.* **589**, 822–828 (2015).

34. Wang, Y. *et al.* KAT2A coupled with the α-KGDH complex acts as a histone H3 succinyltransferase. *Nature* **552**, 273–277 (2017).

35. De, P. & Chatterjee, R. Nucleolar localization of succinic dehydrogenase in human malignant cells with MTT. *Experientia* **18**, 562–562 (1962).

36. De, P. & Chatterjee, R. Evidence of nucleolar succinic dehydrogenase activity. *Exp. Cell Res.* **27**, 172–173 (1962).

37. Yogev, O. *et al.* Fumarase: A Mitochondrial Metabolic Enzyme and a Cytosolic/Nuclear Component of the DNA Damage Response. *PLoS Biol.* **8**, e1000328 (2010).

38. Jiang, Y. *et al.* Local generation of fumarate promotes DNA repair through inhibition of histone H3 demethylation. *Nat. Cell Biol.* **17**, 1158–1168 (2015).

39. Pavlova, N. N. & Thompson, C. B. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab.* **23**, 27–47 (2016).

40. Vander Heiden, M. G. & DeBerardinis, R. J. Understanding the Intersections between Metabolism and Cancer Biology. *Cell* **168**, 657–669 (2017).

41. Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem. Sci.* **41**, 211–218 (2016).

42. Warburg, O. On the origin of cancer cells. *Science* **123**, 309–14 (1956).

43. Lunt, S. Y. & Vander Heiden, M. G. Aerobic Glycolysis: Meeting the Metabolic Requirements of Cell Proliferation. *Annu. Rev. Cell Dev. Biol.* **27**, 441–464 (2011).

44. Gaglio, D. *et al.* Oncogenic K-Ras decouples glucose and glutamine metabolism to support cancer cell growth. *Mol. Syst. Biol.* **7**, 523 (2011).

45. Hu, H. *et al.* Phosphoinositide 3-Kinase Regulates Glycolysis through Mobilization of Aldolase from the Actin Cytoskeleton. *Cell* **164**, 433–446 (2016).

46. Yu, P. *et al.* FGF-dependent metabolic control of vascular development. *Nature* **545**, 224–228 (2017).

47. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–33 (2009).

48. Locasale, J. W. *et al.* Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nat. Genet.* **43**, 869–874 (2011).

49. DeBerardinis, R. J. *et al.* Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19345–50 (2007).

50. Riemann, A. *et al.* Acidic priming enhances metastatic potential of cancer cells. *Pflügers Arch. - Eur. J. Physiol.* **466**, 2127–2138 (2014).

51. Colegio, O. R. *et al.* Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. *Nature* **513**, (2014).

52. Carmona-Fontaine, C. *et al.* Metabolic origins of spatial organization in the tumor microenvironment. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2934–2939 (2017).

53. Hui, S. *et al.* Glucose feeds the TCA cycle via circulating lactate. *Nature* **551**, 1–12 (2017).

54. Faubert, B. *et al.* Lactate Metabolism in Human Lung Tumors. *Cell* **171**, 358-371.e9 (2017).

55. Chen, Y. J. *et al.* Lactate metabolism is associated with mammalian mitochondria. *Nat. Chem. Biol.* **12**, 937–943 (2016).

56. Sonveaux, P. *et al.* Targeting lactate-fueled respiration selectively kills hypoxic tumor cells in mice. *J. Clin. Invest.* **118**, 3930–42 (2008).

57. Boudreau, A. *et al.* Metabolic plasticity underpins innate and acquired resistance to LDHA inhibition. *Nat. Chem. Biol.* **12**, 779–786 (2016).

58. Tan, A. S. *et al.* Mitochondrial Genome Acquisition Restores Respiratory Function and Tumorigenic Potential of Cancer Cells without Mitochondrial DNA. *Cell Metab.* **21**, 81–94 (2015).

59. Dong, L.-F. *et al.* Horizontal transfer of whole mitochondria restores tumorigenic

potential in mitochondrial DNA-deficient cancer cells. *Elife* **6**, (2017).

60. Wallace, D. C. Mitochondria and cancer. *Nat. Rev. Cancer* **12**, 685–698 (2012).

61. Bajzikova, M. *et al.* Reactivation of Dihydroorotate Dehydrogenase-Driven Pyrimidine Biosynthesis Restores Tumor Growth of Respiration-Deficient Cancer Cells. *Cell Metab.* **29**, 399-416.e10 (2019).

62. Sullivan, L. B. *et al.* Supporting Aspartate Biosynthesis Is an Essential Function of Respiration in Proliferating Cells. *Cell* **162**, 552–563 (2015).

63. Birsoy, K. *et al.* An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. *Cell* **162**, 540–551 (2015).

64. Wheaton, W. W. *et al.* Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. *Elife* **3**, e02242 (2014).

65. Griss, T. *et al.* Metformin Antagonizes Cancer Cell Proliferation by Suppressing Mitochondrial-Dependent Biosynthesis. *PLOS Biol.* **13**, e1002309 (2015).

66. Porporato, P. E., Filigheddu, N., Pedro, J. M. B.-S., Kroemer, G. & Galluzzi, L. Mitochondrial metabolism and cancer. *Cell Res.* **28**, 265–280 (2018).

67. Beloribi-Djefaflia, S., Vasseur, S. & Guillaumond, F. Lipid metabolic reprogramming in cancer cells. *Oncogenesis* **5**, e189–e189 (2016).

68. Yoo, H., Stephanopoulos, G. & Kelleher, J. K. Quantifying carbon sources for de novo lipogenesis in wild-type and IRS-1 knockout brown adipocytes. *J. Lipid Res.* **45**, 1324–1332 (2004).

69. Chajès, V., Cambot, M., Moreau, K., Lenoir, G. M. & Joulin, V. Acetyl-CoA Carboxylase α Is Essential to Breast Cancer Cell Survival. *Cancer Res.* **66**, 5287–5294 (2006).

70. Migita, T. *et al.* ATP Citrate Lyase: Activation and Therapeutic Implications in Non-Small Cell Lung Cancer. *Cancer Res.* **68**, 8547–8554 (2008).

71. Wise, D. R. *et al.* Hypoxia promotes isocitrate dehydrogenase-dependent carboxylation of α-ketoglutarate to citrate to support cell growth and viability. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19611–6 (2011).

72. Metallo, C. M. *et al.* Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature* **481**, 380–384 (2012).

73. Fendt, S.-M. *et al.* Reductive glutamine metabolism is a function of the α-ketoglutarate to citrate ratio in cells. *Nat. Commun.* **4**, 2236 (2013).

74. Wise, D. R. *et al.* Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc. Natl. Acad. Sci.* **105**, 18782–18787 (2008).

75. Gross, M. I. *et al.* Antitumor Activity of the Glutaminase Inhibitor CB-839 in Triple-Negative Breast Cancer. *Mol. Cancer Ther.* **13**, 890–901 (2014).

76. Vander Heiden, M. G. & DeBerardinis, R. J. Understanding the Intersections between Metabolism and Cancer Biology. *Cell* **168**, 657–669 (2017).

77. Yuneva, M. O. *et al.* The metabolic profile of tumors depends on both the responsible genetic lesion and tissue type. *Cell Metab.* **15**, 157–170 (2012).

78. Mayers, J. R. *et al.* Tissue of origin dictates branched-chain amino acid metabolism in mutant Kras-driven cancers. *Science* **353**, 1161–5 (2016).

79. Davidson, S. M. *et al.* Direct evidence for cancer-cell-autonomous extracellular protein catabolism in pancreatic tumors. *Nat. Med.* **23**, 235–241 (2017).

80. Davidson, S. M. *et al.* Environment Impacts the Metabolic Dependencies of Ras-Driven Non-Small Cell Lung Cancer. *Cell Metab.* **23**, 517–528 (2016).

81. Jechlinger, M. Organotypic Culture of Untransformed and Tumorigenic Primary Mammary Epithelial Cells. *Cold Spring Harb. Protoc.* **2015**, pdb.prot078295 (2015).

82. Kinnaird, A., Zhao, S., Wellen, K. E. & Michelakis, E. D. Metabolic control of epigenetics in cancer. *Nat. Rev. Cancer* **16**, 694–707 (2016).

83. Katada, S., Imhof, A. & Sassone-Corsi, P. Connecting threads: Epigenetics and metabolism. *Cell* **148**, 24–28 (2012).

84. Beck, M. *et al.* The quantitative proteome of a human cell line. *Mol. Syst. Biol.* **7**, 549–549 (2014).

85. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

86. Kanehisa, M. *et al.* KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).

87. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).

88. Nagaraj, R. *et al.* Nuclear Localization of Mitochondrial TCA Cycle Enzymes as a Critical Step in Mammalian Zygotic Genome Activation: Cell. *Cell* **168**, 210-223.e11 (2016).

89. Robinson, J. B. & Srere, P. A. Organization of Krebs tricarboxylic acid cycle enzymes. *Biochem. Med.* **33**, 149–157 (1985).

90. Robinson, J. B., Inman, L., Sumegi, B. & Srere, P. A. Further characterization of the Krebs tricarboxylic acid cycle metabolon. *J. Biol. Chem.* **262**, 1786–1790 (1987).

91. Wu, F. & Minteer, S. Krebs cycle metabolon: Structural evidence of substrate channeling revealed by cross-linking and mass spectrometry. *Angew. Chemie - Int. Ed.* **54**, 1851–1854 (2015).

92. Wu, F., Pelster, L. N. & Minteer, S. D. Krebs cycle metabolon formation: Metabolite concentration gradient enhanced compartmentation of sequential enzymes. *Chem. Commun.* **51**, 1244–1247 (2015).

93. Zhang, Y. *et al.* Protein-protein interactions and metabolite channelling in the plant tricarboxylic acid cycle. *Nat. Commun.* **8**, (2017).

94. Thul, P. J. *et al.* A subcellular map of the human proteome. *Science (80-.).* **356**, (2017).

95. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science (80-.).* **347**, (2015).

96. Maddocks, O. D. *et al.* A roadmap for interpreting 13 C metabolite labeling patterns from cells. *Curr. Opin. Biotechnol.* **34**, 189–201 (2015).

97. Jang, C., Chen, L. & Rabinowitz, J. D. Metabolomics and Isotope Tracing. *Cell* **173**, 822–837 (2018).

98. Lemarie, A., Huc, L., Pazarentzos, E., Mahul-Mellier, A. L. & Grimm, S. Specific disintegration of complex II succinate:Ubiquinone oxidoreductase links pH changes to

oxidative stress for apoptosis induction. *Cell Death Differ.* **18**, 338–349 (2011).

99. Schwall, C. T., Greenwood, V. L. & Alder, N. N. The stability and activity of respiratory Complex II is cardiolipin-dependent. *Biochim. Biophys. Acta - Bioenerg.* **1817**, 1588–1596 (2012).

100. Lane, A. N. & Fan, T. W. M. Regulation of mammalian nucleotide metabolism and biosynthesis. *Nucleic Acids Res.* **43**, 2466–2485 (2015).

101. Gui, D. Y. *et al.* Environment Dictates Dependence on Mitochondrial Complex I for NAD+ and Aspartate Production and Determines Cancer Cell Sensitivity to Metformin. *Cell Metab.* **24**, 716–727 (2016).

102. Irvine, D. J. *et al.* Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature* **481**, 380–384 (2011).

103. Mullen, A. R. *et al.* Reductive carboxylation supports growth in tumour cells with defective mitochondria. *Nature* **481**, 385–388 (2012).

104. Iacobazzi, V. & Infantino, V. Citrate – new functions for an old metabolite. *Biol. Chem.* **395**, 387–99 (2014).

105. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801–810 (2012).

106. Kim, D. I. *et al.* An improved smaller biotin ligase for BioID proximity labeling. *Mol. Biol. Cell* **27**, 1188–1196 (2016).

107. Liu, X. *et al.* An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. *Nat. Commun.* **9**, (2018).

108. Ma, T., Peng, Y., Huang, W., Liu, Y. & Ding, J. The β and γ subunits play distinct functional roles in the α 2 βγ heterotetramer of human NAD-dependent isocitrate dehydrogenase. *Sci. Rep.* **7**, 1–12 (2017).

109. Mellacheruvu, D. *et al.* The CRAPome : a contaminant repository for affinity purification – mass spectrometry data. **10**, (2013).

110. Dudek, J., Rehling, P. & van der Laan, M. Mitochondrial protein import: Common principles and physiological networks. *Biochim. Biophys. Acta - Mol. Cell Res.* **1833**, 274–285 (2013).

111. Hernandez-Verdun, D., Roussel, P., Thiry, M., Sirri, V. & Lafontaine, D. L. J. The nucleolus: Structure/function relationship in RNA metabolism. *Wiley Interdiscip. Rev. RNA* **1**, 415–431 (2010).

112. Uri Reimand, J. ¨ *et al.* g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkw199

113. Ri Reimand, J., Kull, M., Peterson, H., Hansen, J. & Vilo, J. g:Profiler-a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.* 1–8 (2007). doi:10.1093/nar/gkm226

114. Binder, J. X. *et al.* COMPARTMENTS: Unification and visualization of protein subcellular localization evidence. *Database* **2014**, 1–9 (2014).

115. Sgarra, R. *et al.* HMGA molecular network: From transcriptional regulation to chromatin remodeling. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1799**, 37–47 (2010).

116. Matsunaga, S. *et al.* RBMX: A Regulator for Maintenance and Centromeric Protection

of Sister Chromatid Cohesion. *Cell Rep.* **1**, 299–308 (2012).

117. Chen, S. H., Chan, N.-L. & Hsieh, T. New Mechanistic and Functional Insights into DNA Topoisomerases. *Annu. Rev. Biochem.* **82**, 139–170 (2013).

118. Heinrich, B. *et al.* Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regionsin pre-mRNA. *J. Biol. Chem.* **284**, 14303–14315 (2009).

119. Fusco, A. & Fedele, M. Roles of HMGA proteins in cancer. *Nat. Rev. Cancer* **7**, 899–910 (2007).

120. Tiwari, V. K. *et al.* Target genes of Topoisomerase II regulate neuronal survival and are defined by their chromatin state. *Proc. Natl. Acad. Sci.* **109**, E934–E943 (2012).

121. Conte, A. *et al.* High mobility group A1 protein modulates autophagy in cancer cells. *Cell Death Differ.* **24**, 1948–1962 (2017).

122. Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).

123. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13**, 2498–2504 (2003).

124. Knight, J. D. R. *et al.* A web-tool for visualizing quantitative protein-protein interaction data. *Proteomics* **15**, 1432–1436 (2015).

125. Doyon, Y., Selleck, W., Lane, W. S., Tan, S. & Côté, J. Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Mol. Cell. Biol.* **24**, 1884–96 (2004).

126. Sigoillot, F. D., Berkowski, J. A., Sigoillot, S. M., Kotsis, D. H. & Guy, H. I. Cell cycle-dependent regulation of pyrimidine biosynthesis. *J. Biol. Chem.* **278**, 3403–3409 (2003).

127. Evans, D. R. & Guy, H. I. Mammalian Pyrimidine Biosynthesis: Fresh Insights into an Ancient Pathway. *J. Biol. Chem.* **279**, 33035–33038 (2004).

128. Liao, Y. *et al.* The Cardiomyocyte RNA-Binding Proteome: Links to Intermediary Metabolism and Heart Disease. *Cell Rep.* **16**, 1456–1469 (2016).

129. Malty, R. H. *et al.* A Map of Human Mitochondrial Protein Interactions Linked to Neurodegeneration Reveals New Mechanisms of Redox Homeostasis and NF-κB Signaling. *Cell Syst.* **5**, 564-577.e12 (2017).

130. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat. Publ. Gr.* **19**, 327–341 (2018).

131. Havas, K. M. *et al.* Metabolic shifts in residual breast cancer drive tumor recurrence. *J. Clin. Invest.* **127**, 2091–2105 (2017).

132. Slamon, D. J. *et al.* Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. *N. Engl. J. Med.* **344**, 783–792 (2001).

133. 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.* **83**, 7074–7080 (2011).

134. Murphy, M. P. & O’Neill, L. A. J. Krebs Cycle Reimagined: The Emerging Roles of Succinate and Itaconate as Signal Transducers. *Cell* **174**, 780–784 (2018).

135. Chouchani, E. T. *et al.* Ischaemic accumulation of succinate controls reperfusion injury through mitochondrial ROS. *Nature* **515**, 431–435 (2014).

136. Sullivan, L. B. *et al.* Aspartate is an endogenous metabolic limitation for tumour growth. *Nat. Cell Biol.* **20**, 782–788 (2018).

137. Keshet, R., Szlosarek, P., Carracedo, A. & Erez, A. Rewiring urea cycle metabolism in cancer to support anabolism. *Nat. Rev. Cancer* **18**, 634–645 (2018).

138. Cervelli, M., Pietropaoli, S., Signore, F., Amendola, R. & Mariottini, P. Polyamines metabolism and breast cancer: state of the art and perspectives. *Breast Cancer Res. Treat.* **148**, 233–248 (2014).

139. Husson, A., Brasse-Lagnel, C., Fairand, A., Renouf, S. & Lavoinne, A. Argininosuccinate synthetase from the urea cycle to the citrulline-NO cycle. *Eur. J. Biochem.* **270**, 1887–99 (2003).

140. Camarda, R., Williams, J. & Goga, A. In vivo Reprogramming of Cancer Metabolism by MYC. *Front. Cell Dev. Biol.* **5**, 1–13 (2017).

141. Gao, P. *et al.* c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. *Nature* **458**, 762–765 (2009).

142. Liu, W. *et al.* Reprogramming of proline and glutamine metabolism contributes to the proliferative and metabolic responses regulated by oncogenic transcription factor c-MYC. *Proc. Natl. Acad. Sci.* **109**, 8983–8988 (2012).

143. Jones, W. & Bianchi, K. Aerobic glycolysis: Beyond proliferation. *Front. Immunol.* **6**, 1–5 (2015).

144. Lee, D. C. *et al.* A lactate-induced response to hypoxia. *Cell* **161**, 595–609 (2015).

145. Martinez-Outschoorn, U. E. *et al.* Ketones and lactate increase cancer cell “stemness,” driving recurrence, metastasis and poor clinical outcome in breast cancer. *Cell Cycle* **10**, 1271–1286 (2011).

146. Tannahill, G. M. *et al.* Succinate is an inflammatory signal that induces IL-1β through HIF-1α. *Nature* **496**, 238–242 (2013).

147. Basudhar, D. *et al.* Coexpression of NOS2 and COX2 accelerates tumor growth and reduces survival in estrogen receptor-negative breast cancer. *Proc. Natl. Acad. Sci.* **114**, 13030–13035 (2017).

148. Billiard, J. *et al.* Quinoline 3-sulfonamides inhibit lactate dehydrogenase A and reverse aerobic glycolysis in cancer cells. *Cancer Metab.* **1**, 1–17 (2013).

149. Flores, A. *et al.* Increased lactate dehydrogenase activity is dispensable in squamous carcinoma cells of origin. *Nat. Commun.* **10**, 91 (2019).

150. Beloueche-Babari, M. *et al.* MCT1 inhibitor AZD3965 increases mitochondrial metabolism, facilitating combination therapy and noninvasive magnetic resonance spectroscopy. *Cancer Res.* **77**, 5913–5924 (2017).

151. Hong, C. S. *et al.* MCT1 Modulates Cancer Cell Pyruvate Export and Growth of Tumors that Co-express MCT1 and MCT4. *Cell Rep.* **14**, 1590–1601 (2016).

152. Silva, L. S. *et al.* Branched-chain ketoacids secreted by glioblastoma cells via MCT1 modulate macrophage phenotype. *EMBO Rep.* **18**, 2172–2185 (2017).

153. Chang, Y.-C., Yang, Y.-C., Tien, C.-P., Yang, C.-J. & Hsiao, M. Roles of Aldolase Family Genes in Human Cancers and Diseases. *Trends Endocrinol. Metab.* **29**, 549–

559 (2018).

154. Schwab, A. *et al.* Polyol Pathway Links Glucose Metabolism to the Aggressiveness of Cancer Cells. *Cancer Res.* **78**, 1604–1618 (2018).

155. Ori, A., Andrés-Pons, A. & Beck, M. The use of targeted proteomics to determine the stoichiometry of large macromolecular assemblies. *Methods Cell Biol.* **122**, 117–46 (2014).

156. Jungreuthmayer, C., Neubauer, S., Mairinger, T., Zanghellini, J. & Hann, S. *ICT* : isotope correction toolbox. *Bioinformatics* **32**, btv514 (2015).

157. Mackmull, M. *et al.* Landscape of nuclear transport receptor cargo specificity. *Mol. Syst. Biol.* **13**, 962 (2017).

158. Werner, T. *et al.* Ion Coalescence of Neutron Encoded TMT 10-Plex Reporter Ions. *Anal. Chem.* **86**, 3594–3601 (2014).

159. Reichel, M. *et al.* In Planta Determination of the mRNA-Binding Proteome of Arabidopsis Etiolated Seedlings. *Plant Cell* **28**, 2435–2452 (2016).

160. Strucko, T. *et al.* Laboratory evolution reveals regulatory and metabolic trade-offs of glycerol utilization in Saccharomyces cerevisiae. *Metab. Eng.* **47**, 73–82 (2018).

161. Franken, H. *et al.* Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat. Protoc.* **10**, 1567–1593 (2015).

162. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

163. Huber, W., von Heydebreck, A., Sültmann, H., Poustka, A. & Vingron, M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18 Suppl 1**, S96-104 (2002).

164. Gatto, L. & Lilley, K. S. MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics* **28**, 288–9 (2012).

165. Mellacheruvu, D. *et al.* The CRAPome: A contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* (2013). doi:10.1038/nmeth.2557

166. Knight, J. D. R. *et al.* ProHits-viz: a suite of web tools for visualizing interaction proteomics data. *Nat. Methods* **14**, 645–646 (2017).

167. 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.* **23**, 1677–1688 (2009).

168. Kanani, H. H. & Klapa, M. I. Data correction strategy for metabolomics analysis using gas chromatography–mass spectrometry. *Metab. Eng.* **9**, 39–51 (2007).

169. Blasche, S. *et al.* Emergence of stable coexistence in a complex microbial community through metabolic cooperation and spatio-temporal niche partitioning. *bioRxiv* 541870 (2019). doi:10.1101/541870