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

presented by

Master in Medical Biotechnologies and Molecular Medicine – Anna Chiara Pirona
born in: San Daniele del Friuli
Oral-examination: 12.03.2020

Unraveling metabolic mechanisms underlying the folate pathway and targeting antifolate resistance via synthetic lethality

Referees:

Prof. Dr. Matthias Mayer
Prof. Dr. Stefan Wiemann

“Maybe the journey isn’t so much about becoming anything.
Maybe it’s about unbecoming everything that isn’t you
so you can be who you were meant to be in the first place.”

Contents

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

In this thesis we investigate acquired antifolate-resistance, with particular interest to folypolyglutamate synthase (FPGS) -dependent drug resistance. Our study model is T-cell acute lymphoid leukemia (T-ALL), already known to show methotrexate FPGS-dependent resistance[1]. In particular, we use two different cell lines: wild type CCRF-CEM and resistant MTA-C3, the latter one showing afunctional FPGS. Key of our study is the one carbon metabolism pathway, that takes part in the processing of (anti)folates in the cell. We first tried to get a better picture of one carbon metabolism and folate pathway in our model of study. We especially focused on the effect that different folate concentrations would have on the cytosolic vs mitochondrial flux usage, and our findings show that the capacity of retaining folates, dependent on RFC levels, is at the bases of it. Moreover, further investigations proved SHMT1 to be synthetic lethal to RFC. This is a novel discovery that can have considerable impact on clinical level. This part of the work contributed in establishing proper culturing conditions for our cell line models. This was essential in order to perform an unbiased CRISPR screening with the aim of finding synthetic lethal candidates to FPGS, downregulated in MTA-C3. To this aim we first established a new efficient process for cell transduction, that was necessary to not introduce subpopulation selection bias during the screening. In fact, T-ALL cell lines are well known for being quite difficult to manipulate and initial transduction trials showed 3-5% efficiency. The CRISPR library used in this screening had 260.000 sgRNA, therefore, not only to avoid experimental bias, but also to be able to handle practically wise such experiment, it was essential to establish a more efficient method. We aimed to reach 30% transduction rate, and we were able to achieve such result by using a new mixture of lipofectamines combined with protein columns for virus production, and cell pelleting for the transduction process. We then performed a CRISPR Cas9 screening and combined the results with three different protein screenings that were also conducted to better understand pathways involved in antifolate resistance. Out of these screenings, we selected 6 candidates based on common hits and possible relation with one carbon metabolism pathway. AHCY showed the most promising phenotype during validation, and our metabolomics analysis confirm the hypothesis that AHCY could act synthetic lethal to FPGS. However, further experiments need to be performed to unravel the complex puzzle and get a complete view.

תקציר המחקר

מחקר זה עוסק בעמידות נרכשת לאנטיפולאטים עם דגש מיוחד על עמידות לתרופות התלויה בפוליגלוטמאט סינתאז (FPGS). מודל המחקר שבו השתמשנו, לוקמיה לימפוציטית חריפה של תאי T (T-ALL), כבר הראה בעבר עמידות למטוטראקסאט התלויה ב FPGS[1]. באופן ספציפי, השתמשנו בשני קווי תאים: תאי CCRF-CEM מזן הבר, ותאי MTA-C3 העמידים לתרופה, כאשר האחרונים מציגים חוסר פעילות של FPGS. חלק חשוב במחקר שלנו הינו המסלול למטבוליזם של יחידת פחמן אחת, המשתתף בעיבוד ה(אנטי)פולאטים בתא. ראשית, ניסינו לקבל תמונה ברורה יותר של מטבוליזם של יחידת פחמן אחת ומסלול הפולאטים במודל המחקר שלנו. התמקדנו בעיקר על האפקט של ריכוזי פולאטים שונים על השטף המטבולי בציטוזול לעומת השטף במיטוכונדריה, ותוצאותינו מעידות שבבסיסו של אפקט זה טמונה היכולת של התאים להחזיק את הפולאטים, כתלות ברמות RFC. יתרה מכך, חקירה נוספת הוכיחה ש SHMT1 ו RFC יוצרים synthetic lethality. זהו גילוי חדש אשר יכול להשפיע משמעותית ברמה הקלינית. חלק זה של העבודה תרם לביסוס תנאי תרבית נאותים למודלים שלנו. דבר זה היה חיוני על-מנת שנוכל לבצע סריקת CRISPR לא מוטה, במטרה למצוא מועמדים שיוכלו להיות synthetic lethal עם FPGS, שמבוטא בצורה פחותה בתאי MTA-C3. לצורך כך ביססנו, כצעד ראשון, תהליך חדש ויעיל להדבקת תאים, כדי שנוכל להימנע מבחירת תת-אוכלוסיה באופן מוטה בזמן הסריקה. למעשה, קווי התאים של T-ALL ידועים בכך שהינם קשים לתמרון, ובדיקות הדבקה ראשוניות הראו יעילות של 3-5%. ספרית ה CRISPR שבנינו לסריקה הכילה 260,000 רנ"א מובילים (sgRNAs), ולכן, פיתוח שיטה יעילה יותר היה נחוץ עד מאד, לא רק על מנת להימנע מהטיית המחקר, אלא גם על מנת שנוכל להתנהל ברמה הפרקטית עם ניסוי בסדר גודל שכזה. המטרה שהצבנו היתה להגיע לקצב הדבקה של 30%, ומטרה זו הושגה באמצעות תרכובת חדשה של ליפופקטאמינים משולבים עם קולונות חלבון לייצור וירוסים, ושיקוע תאים עבור תהליך ההדבקה. לאחר מכן, ביצענו סריקה עם CRISPR Cas9 ושילבנו את התוצאות עם שלוש סריקות חלבונים שונות שנעשו גם הן על מנת להגיע להבנה טובה יותר של מסלולים המעורבים בעמידות לאנטיפולאטים. מתוך סריקות אלו, בחרנו 6 מועמדים על בסיס זיהוי משותף בסריקות והקשר שלהם עם המסלול למטבוליזם של יחידת פחמן אחת. במהלך שלב הואלידציה, AHCY הראה את הפנוטיפ המבטיח ביותר, והניתוח המטבולי אישש את ההיפותזה ש AHCY יכול לגרום ל synthetic lethality בשילוב עם FPGS. עם כל זאת, נדרשים ניסויים נוספים כדי להשלים את התצרף ולקבל את התמונה המלאה.

ZUSAMMENFASSUNG

Diese Arbeit untersucht die erworbene Resistenz gegen Antifolat, mit besonderem Interesse an der Folypolyglutamatsynthase (FPGS)-bedingten Resistenz gegen Arzneimittel. Unser Studienmodell ist die akute lymphatische T-Zell-Leukämie (T-ALL), deren Methotrexat-FPGS-abhängige Resistenzen bereits nachgewiesen wurden. Zwei unterschiedliche Zelllinien wurden als Modell verwendet: die wildtyp-CCRF-CEM- und die resistente MTA-C3-Zelllinie, wobei die letzte eine erhaltene FPGS zeigt. Der Schwerpunkt unserer Studie ist der Kohlenstoff-Stoffwechselweg, der an der Verarbeitung von (Anti) Folaten in der Zelle beteiligt ist. In unserem Studienmodell haben wir zunächst versucht, ein besseres Bild von dem Kohlenstoff-Metabolismus und Folat-Signalweg zu erhalten. Wir haben uns insbesondere auf die Auswirkungen von unterschiedlichen Folatkonzentrationen auf den Verbrauch vom zytosolischen und mitochondrialen Fluss konzentriert und gezeigt, dass die Fähigkeit, Folate in Abhängigkeit von den RFC-Spiegeln zu speichern, für dieses Prozess grundlegend ist. Darüber hinaus haben weitere Untersuchungen gezeigt, dass SHMT1 für RFC synthetisch tödlich ist. Diese Erkenntnis kann in Zukunft erhebliche Auswirkungen auf die Patientenversorgung haben. Dieser Teil der Arbeit trug dazu bei, die richtigen Kulturbedingungen für unsere Zelllinienmodelle zu etablieren. Dies war unerlässlich, um ein unvoreingenommenes CRISPR-Screening mit dem Ziel durchzuführen, synthetische letale Kandidaten für FPGS zu finden, die in MTA-C3 herunterreguliert sind. Zu diesem Zweck haben wir zunächst einen neuen effizienten Prozess für die Zelltransduktion etabliert, der notwendig ist, um während des Screenings keine Subpopulationsselektionsverzerrung zu verursachen. T-ALL-Zelllinien gelten als ziemlich schwierig zu manipulieren, und anfängliche Transduktionsversuche zeigten eine Effizienz von 3-5%. Die CRISPR-Bibliothek, die für dieses Screening verwendet wurde, wies 260.000 sgRNA auf. Um nicht nur experimentelle Verzerrungen zu vermeiden, sondern auch solche Experimente praktisch durchführen zu können, war uns wichtig, eine effizientere Methode zu etablieren. Wir strebten eine Transduktionsrate von 30% an und konnten ein solches Ergebnis erzielen, indem wir eine neue Mischung aus Lipofectaminen in Kombination mit Proteinsäulen für die Virusproduktion und Zellpelletierung für den Transduktionsprozess verwendeten. Anschließend führten wir ein CRISPR Cas9-Screening durch und kombinierten die Ergebnisse mit drei verschiedenen Protein-Screenings, um die mit der Resistenz gegen Antifolat verbundenen Signalwege besser zu verstehen. Aus diesen Screenings haben wir 6 Kandidaten ausgewählt, die auf gemeinsamen Treffern und einer möglichen Beteiligung am Stoffwechselweg für Kohlenstoff beruhen. AHCY zeigte den vielversprechendsten Phänotyp während der Validierung und unsere Metabolomics-Analyse bestätigt die Hypothese, dass AHCY für FPGS synthetisch letal wirken könnte. Es müssen jedoch weitere Experimente durchgeführt werden, um das komplexe Rätsel zu lösen und einen vollständigen Überblick zu erhalten.

1. INTRODUCTION

1.1 WHAT IS FOLATE

Folate, also known as folic acid or Vitamine B9, is an essential metabolite for the synthesis of purines and pyrimidines [2] , and it is therefore involved in the production of DNA and RNA. As a consequence, it is easy to understand how folate plays a major role in cell growth and proliferation. As our body is not able to produce folate, we must take it with the diet. The recommended intake is 400 ug daily. Physiological concentrations of folate range between 150-450 nM in serum and get 10-20 fold higher inside the cells [34]. We can find folate in different forms that share a common structure constituted of 3 chemical moieties: a pteridine ring, a para-aminobenzoic acid (PABA) linker and a polyglutamate tail of variable length [5]. The pteridine ring together with PABA, serves to bind 1C units. The number of glutamic groups gives an information about the localization of the folate within the cell. The folate form introduced with the diet is tetrahydrofolate (THF), which presents a reduced pteridine ring and is also the biologically active form. Other forms of folate are 5-methyl THF, which is folate in the monoglutamyl form, and folic acid, which is the fully oxidized monoglutamate form. Folic acid and folinic acid (also known as Leucovorin) are usually administered together with antifolates to the patients in order to reduce the side effects, and to pregnant women to help prevent birth defects of the brain and spinal cord [678]. Folic acid must be reduced first to dihydrofolate (DHF) and then to THF by dihydrofolate reductase (DHFR) before it can enter the folate cycle [9], while folinic acid does not require activation by DHFR and it is therefore ready to use[10] . Insufficient dietary folate can cause anemia in adults (John Hopkins Medicine, https://www.hopkinsmedicine.org/health/conditions-and-diseases/folate-deficiency-anemia), while in developing fetuses it can lead to neural tube defects, resulting in anencephaly, fetal loss, spina bifida [11]. Once folate enters our system, there is a number of transporters and enzymes that process it in

order to obtain its active form, this whole pathway is called one-carbon metabolism[9,12] .

1.1.2 One Carbon Metabolism

One-carbon metabolism is a complex pathway, taking place in both the cytoplasm and mitochondria [13], supporting the biosynthesis of RNA and DNA, as well as transmethylation of DNA and proteins [14] . It is tightly interconnected to that of the methyl group donors, serine, glycine, and methionine, whose major tumorigenic role has been demonstrated recently [15,16].

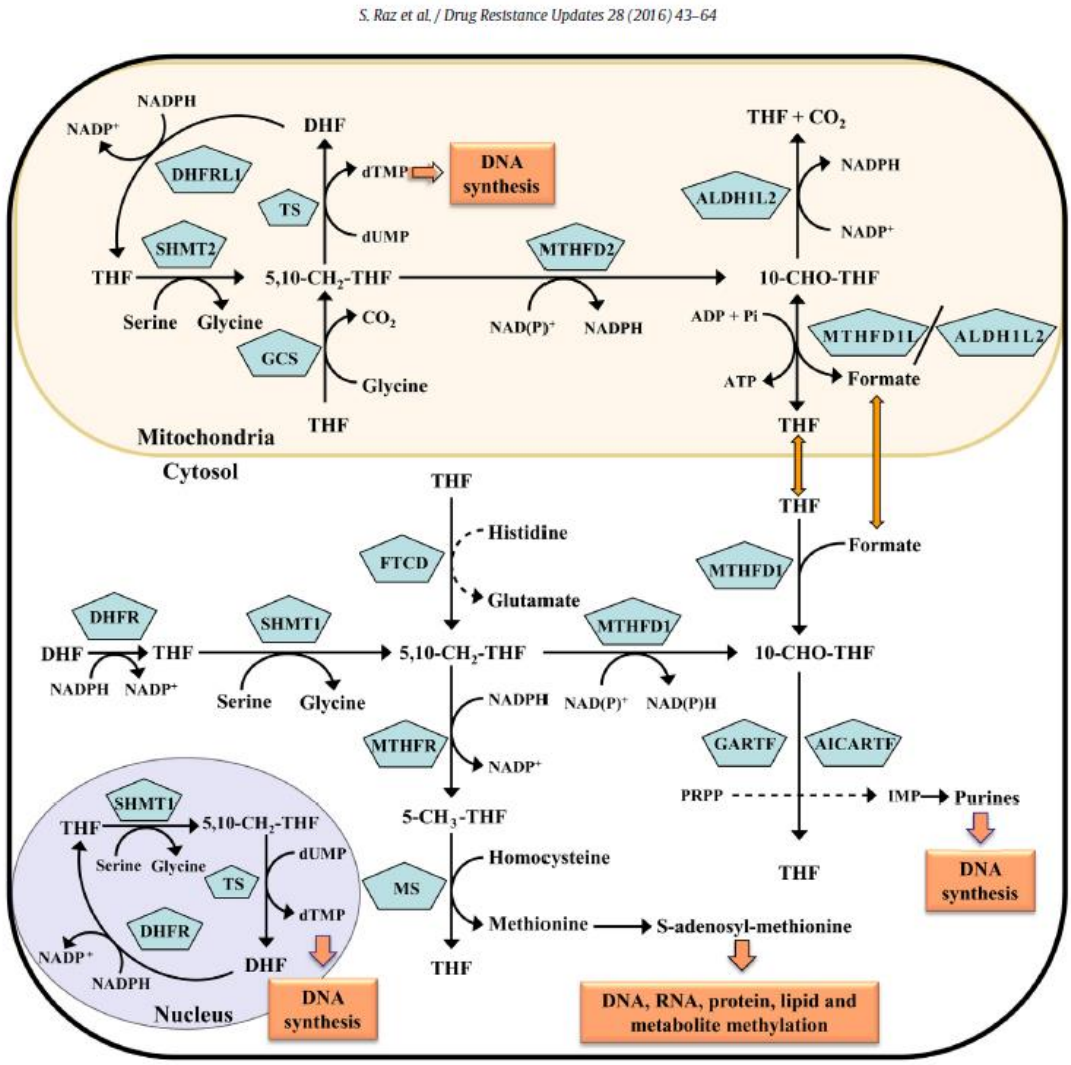


Figure 1: Illustration of one carbon metabolism pathway and it's compartimentalization in cytosol, mitochondria and nucleous. Raz et al., 2016.

Together with various enzymes and transporters, folate constitutes the one carbon metabolism network (*Figure 1*). Through one carbon metabolism, 1C units are activated and transferred for the purpose of purine and thymidine biosynthesis, amino acid homeostasis and for methylation processes [9]. In particular, folate enters the cell via 3 different transporters: Proton Coupled Folate Transporter (PCFT), Reduced Folate Carrier (RFC) and Folate Receptor (FR) [12]. PCFT is mainly involved in the intestinal folate absorption and distribution of it to the central nervous system, RFC is the main transport system for reduced folates, and its physiologic substrate is 5-methyl THF, which is also the most abundant circulating folate form. FR is a high affinity receptor that transports folate via endocytosis. Folate functions as a carrier for 1C units, that get covalently bound to the 5-N atom of the pteridine ring and the 10-N atom of the PABA. Because of different types of binding, 1C units can have 3 different carbon oxidation states: 5,10 – methylene – THF, 5- methyl –THF, and 10-formyl-THF. Each of these forms correlates with different biosynthetic processes. We can also find 5-formyl- THF which simply serves as a 1C reservoir [9,17]. Serine is the main provider of 1C units, though glycine can also be converted into serine through serine hydroxymethyltransferase (SHMT) and serve as a 1C units source. Main products of the 1C metabolism pathway are deoxythymidine monophosphate (dTMP), methionine, and purines. Deoxythymidine monophosphate derives from the conversion of deoxyuridine monophosphate (dUTP) through thymidilate synthase (TYMS) in a 5,10-methylene-THF-dependent reaction.

Methionine is produced through the remethylation of homocysteine, for which 5-methyl-THF is the responsible folate form. Methylene tetrahydrofolate reductase (MTHFR) reduces 5,10 – methyl-THF to 5-methyl-THF, which is then used by methionine synthase (MTR) to obtain the final product. Homocysteine remethylation has a key role in epigenetics, since methionine is the substrate for S-adenosyl-methionine (SAM) which together with S-adenosyl-homocysteine (SAH), is responsible for histone methylation and other methylation reactions. SAH is hydrolyzed by S-adenosyl homocysteine

hydrolase (AHCY) into adenosine and homocysteine. The latter one is then remethylated by methionine synthase (MTR) to obtain methionine, or redirected to the transsulfuration pathway to form sulfur-containing aminoacids (*Figure 2*)[18]. SAM is the most common enzymatic cofactor after ATP and is involved in several biosynthetic processes, including phosphatidylcholine, creatine and polyamine synthesis.

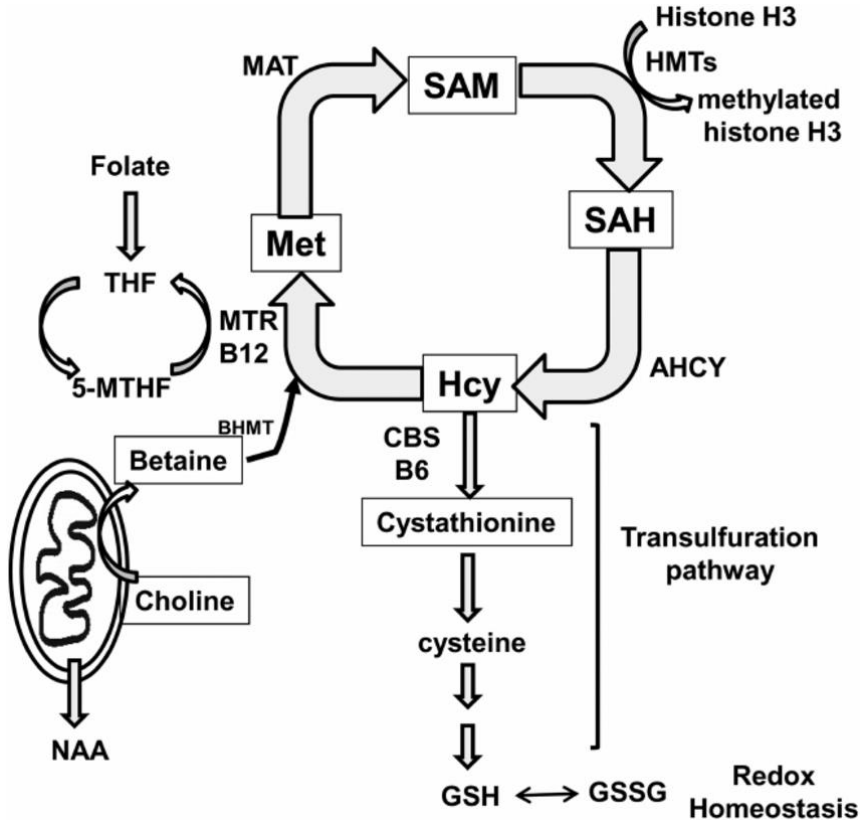


Figure 2: Representation of methionine metabolism reactions, Singhal et at., 2015

Purine synthesis is the metabolic process with the largest demand for 1C units. It occurs on the outside of mitochondria in a complex called purinosome. Mitochondrial 10-formyl-THF is required for the starting of the process, and at the end two carbon units derived from 10-formyl-THF will be incorporated into the purine backbone. This process requires 11 reactions and is illustrated in *Figure 3*[19].

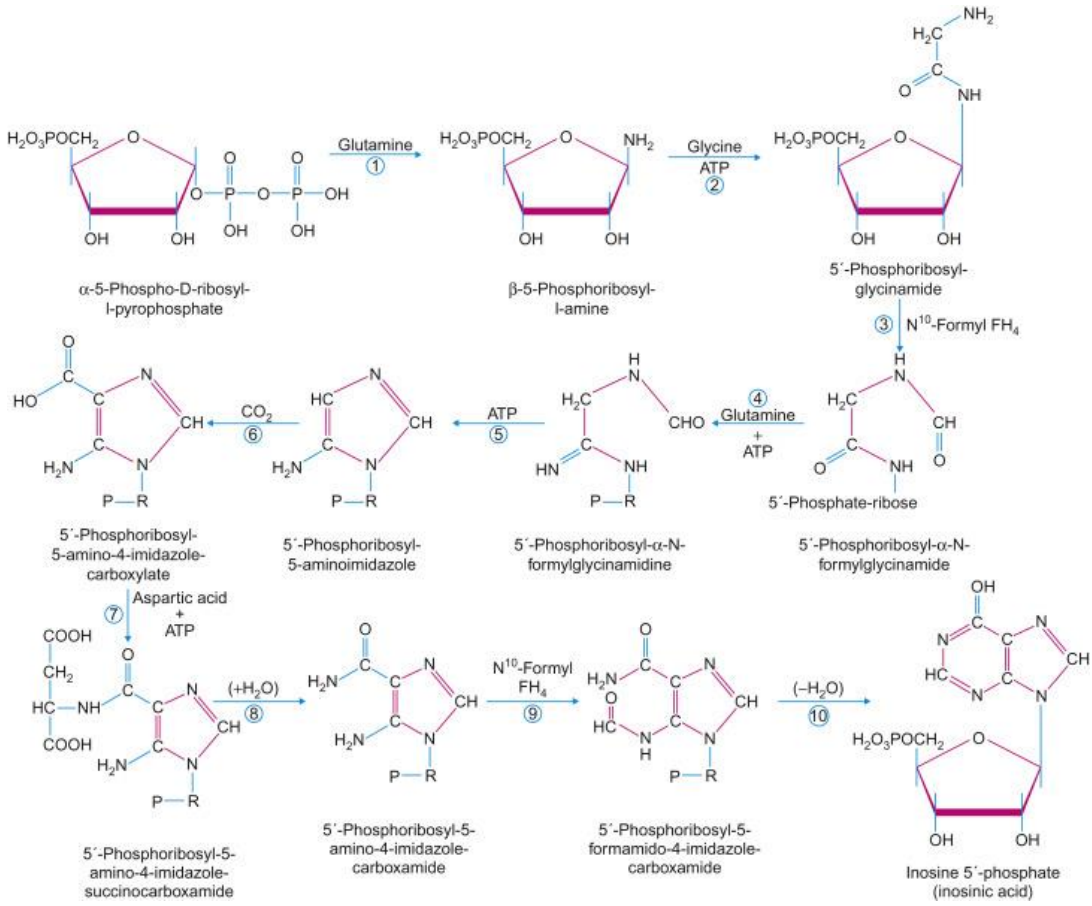


Figure 3: An overview of purine biosynthesis, Bhagavan et al., 2015.

Many of the 1C metabolism reactions are reversible, however the oxidation of 10-formyl-THF with production of NADPH and release of CO_2 is not. One carbon metabolism also contributes to the synthesis of glycine when needed. In fact, when endogenous glycine is unavailable, it can get synthetized from serine. This is very important for several biosynthetic pathways where glycine is required, such as glutathione, purine, creatine and heme synthesis [20].

1.1.2.1 SHMT1 SHMT2 Pathway

SHMT is an enzyme present in the cells in two isoforms: SHMT1 and SHMT2, localized in cytoplasm and mitochondria respectively [21]. Mitochondrial one carbon metabolism generates 1C units through the activity of SHMT2. The 1C

units obtained from this pathway are exported to the cytoplasm as formate, that will then contribute to the cytoplasmic one carbon metabolism. The cytoplasmic enzyme, SHMT1, generates THF (1C donor) that will contribute to methylene and thymidylate biosynthesis. SHMT1 together with TYMS and DHFR is also present in the nucleus, contributing to de novo thymidylate biosynthesis [22]. Because of the role of SHMT in producing available 1C units, this enzyme has lately gained interest under a clinical perspective [23,24]. *Figure 4* illustrates a schematic view of the pathway[25]

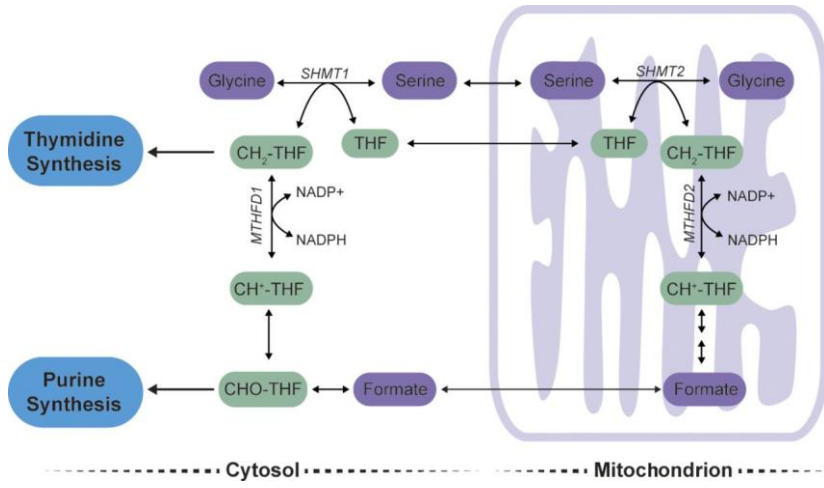


Figure 4: An overview of SHMT1 SHMT2 pathway and serine metabolism. Mattaini et al., 2016.

1.1.3 Culturing Medias

Since this study wants to unravel possible mechanisms of antifolate resistance and folate is an essential nutrient, it is important to establish proper cell culturing conditions for *in vitro* experiments and consider proper diet for *in vivo* studies. Folate is generally present in basic commercial cell culturing media at a concentration over 10-fold higher than the physiological one. The standard RPMI culturing media used for leukemic cells has a folate concentration of circa 2 uM, while serum levels of folate range between 150-450 nM, meaning that the actual concentration of folic acid used for cell culture purposes is at least 10 folds higher that the physiological one. Due to this reason, for our studies we

used two customized media formulation that allowed us to compare cell behaviors and respective activated pathways in low (200 nM) and high folate (2 uM).

1.2 CANCER AND FOLATE PATHWAY

Cancer, together with cardiovascular diseases (CVD), is the main cause of death among middle aged adults, recently even overtaking CVDs in high income countries (HIC) [26,27]. There are over 100 types of cancers classified [28] considering different signs typical of the disease. Researchers defined 12 hallmarks of cancer (*Figure 5*) based on genomic instability, inflammation, dysregulated growth and cell proliferation, evasion of the immune system, metabolic changes, suppression of apoptosis, upregulation of oncogenes, angiogenesis and metastatic potential [29].

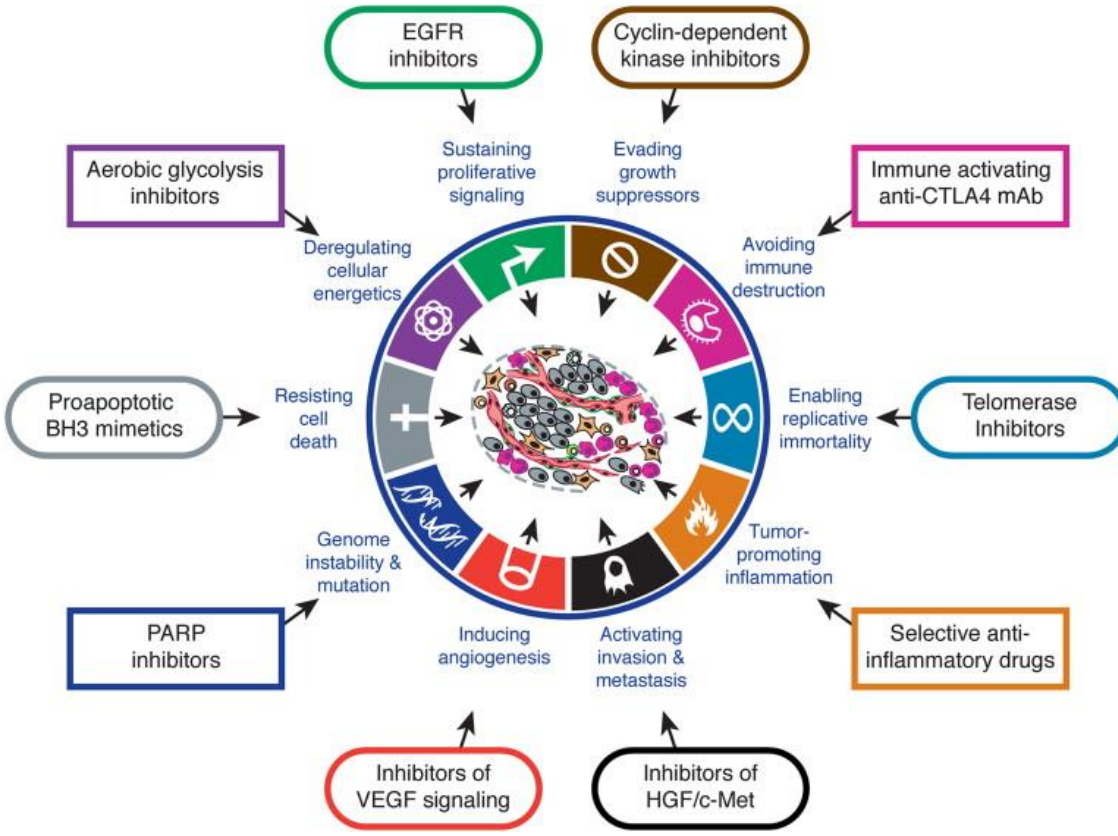


Figure 5: Hallmarks of cancer, Hanahan et al., 2011.

The recent resurgence of interest in the field of cancer metabolism followed several intriguing findings [30,31]: (i) The metabolic rewriting in cancer cells spans many pathways and goes much beyond the sole induction of aerobic glycolysis; (ii) tumor suppressors and oncogenes within growth-factor signaling pathways directly regulate the activity of a variety of metabolic pathways; and (iii) oncogenic mutations were found in several metabolic enzymes in specific cancers. Cancer cells develop a remarkably distinct metabolism compared to normal cells due to major anabolic and energetic requirements necessary to sustain their abnormal proliferation [32]. Evidence for alterations in metabolic activity in malignant cells goes back almost a century ago, where classical studies on cancer bioenergetics demonstrated enhanced glycolysis and relatively suppressed oxidative phosphorylation (a phenomenon known as the 'Warburg effect' [33]). The finding of an enhanced rate of DNA replication in cancer cells has led to the pioneer development of cytotoxic agents known as antimetabolites that directly interfere with various metabolic pathways involved in the biosynthesis of deoxynucleotides essential for DNA replication, which are widely used nowadays in the treatment of various human hematologic malignancies and solid tumors. One of these metabolic pathways is the one carbon metabolism, that sees the folate pathway at the center of it, providing 1C units for purine and pyrimidine biosynthesis, contributing therefore to DNA replication and cell proliferation [9]. *Figure 6* shows an historic of one-carbon metabolism in cancer[34].

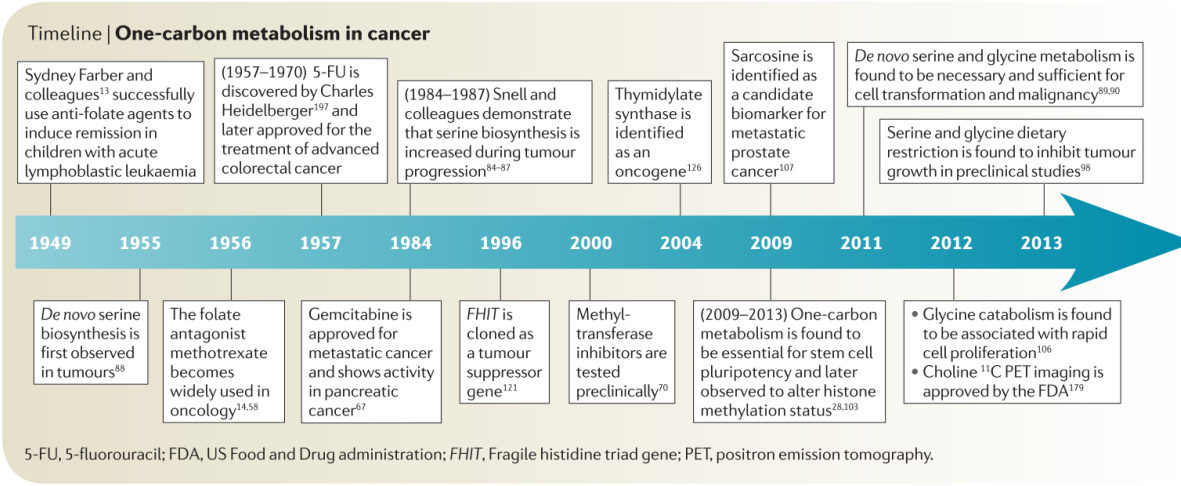


Figure 6: Clinical impact of one-carbon metabolism in cancer, Locasale et al., 2013.

For this reason, antifolates represent an excellent example for cytotoxic drugs targeting the metabolism of cancer cells. Being structural analogues of folic acid, these antimetabolite drugs specifically inhibit various crucial enzymes in folate metabolism and de novo biosynthesis of purine and pyrimidine nucleotides [35].

1.2.1 Antifolates And Drug Resistance: The Role Of FPGS

Antifolates have been extensively used in the past six decades, though their efficacy has been limited due to the frequent emergence of drug resistance phenomena [36–38]. Aminopterin was the first antifolate drug, initially used to treat children leukemia, subsequently replaced by Methotrexate (MTX)[39]. MTX is one of the classical antifolate drugs, and it has a high affinity for DHFR. This leads to defects in THF synthesis therefore affecting the de novo synthesis of thymidine and DNA replication. Other used antifolates in cancer treatment are pemetrexed, pralatrexate and raltitrexed (Tomudex)[35]. Most cancer treatments that involve the use of antifolates, require the administration of Leucovorin (folinic acid) to decrease the toxic effects of the chemotherapeutic agent [40]. Folinic acid can bypass DHFR and is therefore available for the cell to be used for biosynthetic processes. Antifolates enter the cell through folate transporters RFC and PCFT, and get polyglutamilated bt FPGS[12,41] . The addition of glutamic groups to the antifolate drug causes its retention in the cell with consequent

cytotoxic action and cell death[36]. Being folate mimics, antifolates mostly target metabolic enzymes of the folate pathway. The favourite targets are DHFR and TS, however, because of the role of SHMT in producing available 1C units, this enzyme has lately gained interest under a clinical perspective [23,24]. Inhibition of SHMT might slow down the progress of the disease affecting cell growth and proliferation. Antifolate drugs act as folate mimics (*Figure 7*), and get therefore processed by the same cellular pathways that serve to obtain the active folate form, such as the SHMT pathway. Pemetrexed has already been shown to be an effective SHMT inhibitor in cases of mesothelioma [42]. *Table 1* lists the main antifolate drugs currently use in clinic and their specific target.

Antifolate	**Class**	**Pharmacologic target**	**Indications**
Aminopterin	Antineoplastic	DHFR	Children leukemia
Methotrexate	Antineoplastic and immunosuppressant	DHFR	Hematologic malignancies and osteosarcoma
Pemetrexed (MTA, multitargeted antifolate)	Antineoplastic	DHFR, TS, GARFT, SHMT	Non-small cell lung carcinoma, mesothelioma, leukemia
Pralatrexate	Antineoplastic	DHFR, TS	PTCL
Raltitrexed (Tomudex)	Antineoplastic	TS	Colorectal cancer

Table 1: Principal antifolate drugs, targets and clinical indications.

A dominant obstacle towards selective drug targeting and curative cancer therapeutics is the frequent emergence of drug resistance phenomena that are often provoked by chemotherapy [43,44]. For drugs that target cancer metabolism, including antimetabolites, acquired drug resistance is associated with metabolic alterations at the metabolite level as well as fluxes that compensate for the inhibited target enzyme [45]. Such metabolic adaptations may spread across multiple pathways due to the interconnected nature of metabolic networks and the abundance of metabolic regulatory feedback mechanisms. Antifolate resistance involves a plethora of metabolic alterations, including the

mutational inactivation and down-regulation of (anti)folate transporters, such as the reduced folate carrier (RFC), amplified activity of the target enzymes, or acquired point mutation provoking lower affinity of the enzyme for the antifolates, alterations in specific metabolite concentrations, alterations in efflux of folates and antifolates via various transporter pathways [38,46], and defective polyglutamilation of drugs mostly related to afucntional or low levels of folypoluglutamate synthetase (FPGS)[1,36].

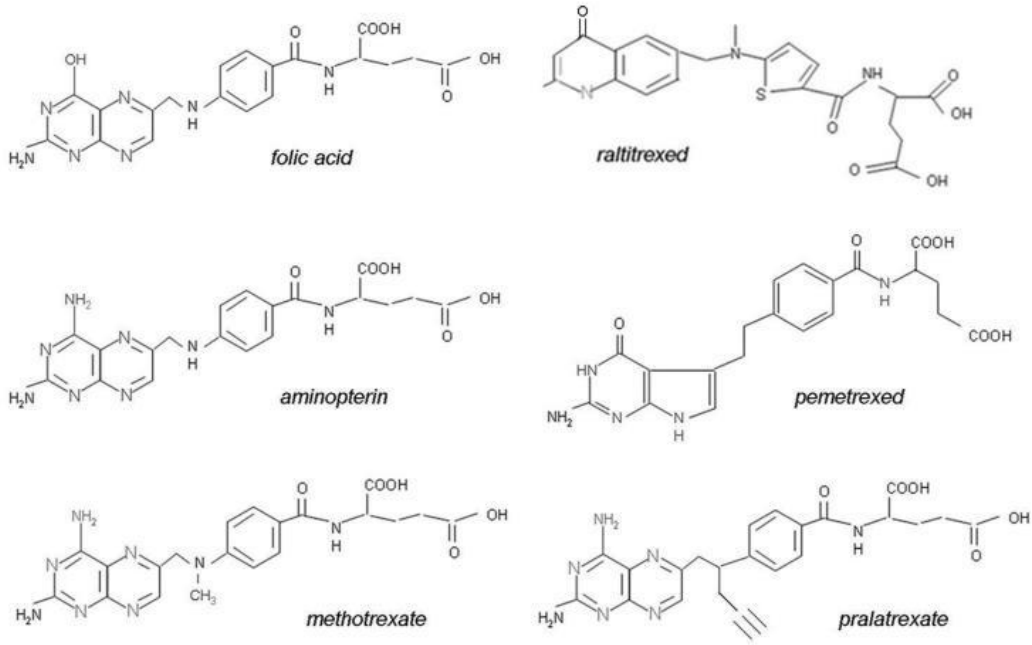


Figure 7: Comparison of the structures of folic acid and antifolates. Visentin et al., 2012.

In this thesis I mostly focus on antifolate resistance due to the impairment of polyglutamylation via loss of function of folylpoly-γ-glutamate synthetase (FPGS). Polyglutamylation of reduced folate cofactors and antifolates markedly enhances their cellular retention and renders them preferred substrates for several folate-dependent enzymes. Impaired FPGS activity was shown to modify the polyglutamylation profile of reduced folate intermediates (by decreasing their retention within the cell) [47] , and hence is expected to significantly alter fluxes in one-carbon metabolism. Considering the central role that reduced folate cofactors play in one-carbon metabolism, these specific metabolic alterations are expected to percolate to other metabolic pathways.

1.2.2 Leukemic Cell Models

For the study purpose of antifolate FPGS-dependent resistance, we decided to focus on leukemic cell line models. This choice was made considering previous literature showing that low levels of FPGS or enzyme mutation are associated with the treatment outcome in leukemic patients [1,48]. The antifolate preferred for the resistant model was Pemetrexed, considering its wide action on several folate pathway enzymes [49]. The fact that this drug has a broad spectra of action once inside the cell, highers the probabilities that the resistance mechanism would be based on its expulsion from the cell, by failing in its polyglutamilation. The cell line chosen was CCRF-CEM, a T-ALL leukemia cell line deriving from a four years old Caucasian female.

1.2.2.1 Resistant Cancer Cell Lines (MTA-C3, MTA-R15)

CCRF-CEM cell line was treated with 460 nM Pemetrexed single dose exposure to obtain MTA-C3 or 1.5 uM Pemetrexed stepwise exposure to obtain MTA-R15 [50] . Both cell lines show loss of over 97% activity of FPGS. In particular, MTA-C3 harbors a heterozygous missense mutation in exon 12 of FPGS (i.e. R363Q substitution, or G1088A), causing the loss of activity of the enzyme. While the allele that harbors this mutation results fully expressed but afunctional, the wild type (WT) allele is completely silenced, due to an epigenetic regulation through Smad4/Ets-1 complex [51]. MTA-R15 instead harbors missense mutations in both FPGS alleles, leading to FPGS loss of function. Because MTA-C3 cell line is a pretty well defined model, genetically more stable than MTA-R15, harboring a specific FPGS mutation and not a corollarium of alterations, we decided to focus on this cell line to study FPGS-dependent antifolate resistance.

1.2.2.2 FPGS Loss Of Function And G1088A Mutation

The genomic mutation, G1088A, present in exon 12 of FPGS in MTA-C3 model is predicted to abolish the binding of ATP to FPGS, thereby causing the loss of activity of the enzyme [51]. FPGS exon 12 can efficiently drive gene transcription and when looking at TFs binding sites, they result disrupted by

the G1088A mutation, which can add to the explanation of loss of activity of the enzyme.

1.3 OMIC-SCIENCES

The term "omic" derives from the Latin "ome", meaning complete, collective set of, or many. –omic sciences are therefore offering a picture of the whole interactions, features, and/or characteristics of the specimen in exam. These approaches usually generate a huge amount of data that need to be processed and analyzed to be able to answer specific scientific questions [52]. For this purpose, statistic becomes a very mandatory and useful science, allowing researchers to stratify results taking into account significance and relevance of the data. Complex bioinformatics models are used for instance to unravel pathway interactions or discover new potential genetic polymorphisms associated to the risk for a certain disease. An interdisciplinary approach combining biological knowledge with data analysis is crucial to be able to direct the next steps in the right direction. A very good example is given by the very high levels of keratin and trypsin when performing mass spectrometry analysis, which are simply contaminants due to the handling of the samples, and without the biological knowledge could be accidentally picked as top candidates of the results [53] . Omic sciences are proteomics (proteins), genomics (DNA), transcriptomics (RNA), metabolomics (metabolites), lipidomics (lipids).

1.3.1 Importance Of Wide Screenings

The impact of –omics is especially seen with the large diffusion of screenings, being these ones in a proteomic, genomic or metabolomics context. The advantage in conducting a screening is the wide overview obtained from the results when comparing two or more different conditions [54,55] . We could imagine ourselves taking a walk along some green fields just nearby the forest, and noticing all the different trees that are standing at the entrance of the woods. It is easy, from this perspective, to declare that the trees are the main difference between the fields and the forest, and are typical of the forest but

extremely absent in the fields. However, if we could take a helicopter and monitor from above the same area, we might be able to see that sometimes, in the middle of the fields, there are some groups of trees. We might also notice how there can be different fruits and plants growing distinctingly in the fields and in the forest. Trees in the forest might prefer a certain disposition in order to allow all of them to receive the needed amount of sun and nutriments. In the same way we can think about healthy and cancerogenic tissue for instance, and we can understand how a wide screening (looking from the helicopter) will provide us more information about similarities and differences between the two. Moreover, these characteristics can also be used to eventually compare out of context samples. Going back to the forest example, we could then compare different forests typical of different areas, i.e. tropical forests versus temperate ones, but not only! We could also relate the type of forest with the type of fields nearby, or compare forest and desert based on the trees/bushes dispositions and eventually find similarities between very far away realities. It is easy now to understand how wide screenings strongly contribute to our knowledge, allowing new insights into diseases and possibly leading to the discovery of new therapeutic targets, very often already known in other contexts for which clinical drugs are already available. This point is of key relevance, considering the long and hard way consisting of *in vitro*, *in vivo*, and clinical trials, necessary to put in commerce a new drug. Very often chemical drugs get approved for their usage 10-20 years after their discovery (cancer research UK, https://www.cancerresearchuk.org/), it is therefore extremely useful to discover clinical targets that could be addressed with already existing approved drugs. After having a broad picture of the situation, it is very important to come back to the close look approach and validate the depicted candidates or pathways [56,57] . This ensures a better understanding of the biology underlying specific mechanisms, and it can reveal small details that were lost with the screening approach [58]. In fact, a downside of wide screenings, is that not-outstanding candidates might get lost in the analysis, despite their function or role being extremely relevant. Again, going back to our metaphor, we can easily imagine how when looking from a helicopter we might not be able to see small

wild strawberries that are part of the forest ecosystem. However, once coming back to our walk to better study the typical trees (depicted from the screening), we might notice a little bed of wild strawberries covering the roots of the trees. Focusing our attention to the specific context, allows us to better study, define and characterize interactions and interactors.

1.3.2 Proteomics

Proteomics refers to the large-scale study of proteomes. A proteome is defined as the set of proteins produced in a biological context (i.e. organism, specific organ, system etc). For instance, we can talk about the proteome of a species (i.e., mouse) or of an organ (i.e., the intestine). The set of proteins is rather dynamic, it changes not only from organ to organ but also from cell to cell and over time [59,60]. The proteome is in part reflective of the transcriptome [61,62] , which refers to the set of mRNAs of that specific biological context. Protein activity is measured as the reaction rate of the processes directed by the protein itself and is dependent on many different factors beside the expression level of the related gene. Proteomics offers great insights about protein expression and cellular/tissue localization, but also modifications (i.e. post-translational modifications) and protein-protein interactions. It also provides knowledge about more complex pathways and the role of specific proteins in a particular context [63]. Through proteomics it is also possible to relate proteins to metabolic pathways of interest [64]. We apply high-throughput technologies to investigate proteomes. The most commonly used is mass spectrometry (MS), and to this regard, in this thesis, we will talk about Stable Isotope Labeling by Amino acids in Cell culture (SILAC) and Label Free Quantification (LFQ). Also protein arrays gained particular interest especially for their direct visual readout when comparing for instance healthy and cancer patient tissue. However, these high-throughput technologies generate huge amounts of data that need to be handled with specific databases, to allow researchers to relate their results with the existing knowledge [65]. The four major proteomic databases are

UniProtKB, IntAct, Reactome and PRIDE. These four databases are created from DNA sequence data and protein functional analysis information (i.e. InterPro).

1.3.2.1 SILAC

Stable isotope labeling by amino acid in cell culture is a quantitative proteomic method. It is best used to compare two or more different conditions (i.e. treatment 1, treatment 2, untreated), and the advantage of this technique is that all samples are run together at the same time in the mass spectrometer [66] . This allows a fair comparison of the peptide amounts found in the different samples, making it easier to depict the proteins differentially expressed in the conditions analyzed. The method is based on growing cells in a culture media containing heavy or light isotope-labeled Amino Acids (AA). In this way, the peptides produced by the cells, will be constituted of the same labeled amino acids present in the culture media.

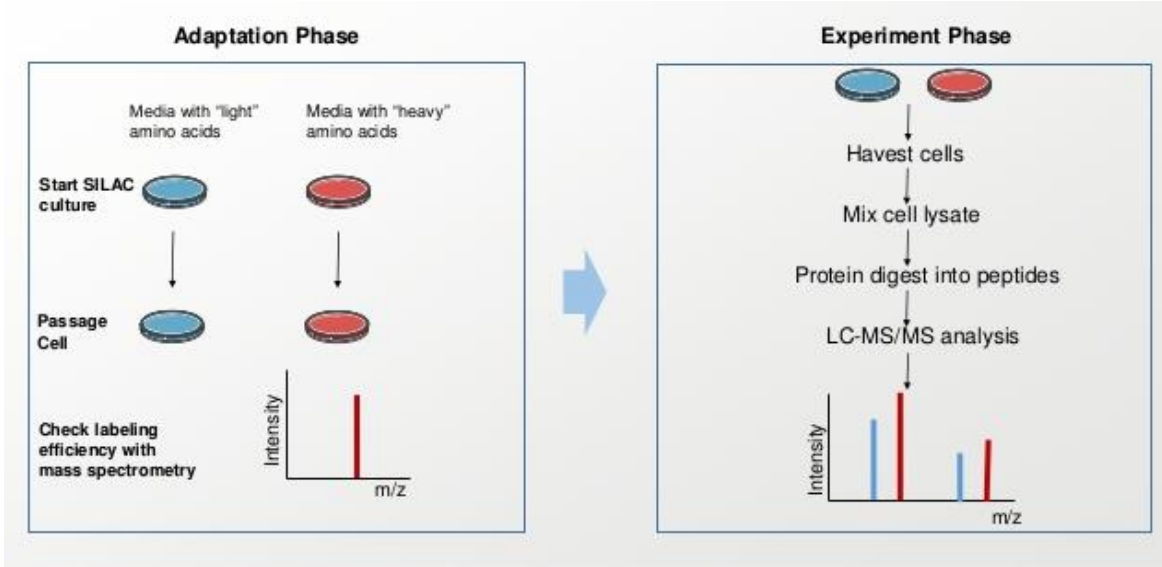


Figure 8: Schematic representation of SILAC analysis, creative proteomics.

We could for instance have three different conditions and grow cells undergoing treatment 1 with heavy istotope-labeled AA, cells of treatment 2 with light isotope labeled AA, and untreated cells in simply normal media without any specific labeling. When running these three samples together on the mass spectrometer, we will be able to quantify the relative amount of each protein for each sample just by looking at the AA labeling. If the protein contains

heavy isotope-labeled AA, it is for sure found in treatment 1 sample, while light-labeled will refer to treatment 2 and unlabeled will be reflective of the amount of protein in the control. This allows a relative quantitative estimation of a specific protein in different samples simultaneously. When comparing healthy and tumorigenic tissue, this method will immediately show which ones are those proteins overexpressed or extremely downregulated in the malignancy. It is important to keep in mind that SILAC gives as output the ratio of a specific protein when looking at different samples. For instance, if P53 value of tumorigenic/healthy tissue is 0.5, it is possible to deduce that P53 expression in the healthy tissue is twice as much compared to the tumor samples, however, it does not provide information about the absolute amount of protein present in any of the two tissues. The output of SILAC is therefore a relative comparison between the analyzed samples (*Figure 8*).

1.3.2.2 LFQ

Label Free Quantification is another mass spectrometry technology that provides information about the proteins present in a sample (qualitative) and their relative amount, usually expressed in intensities [67].

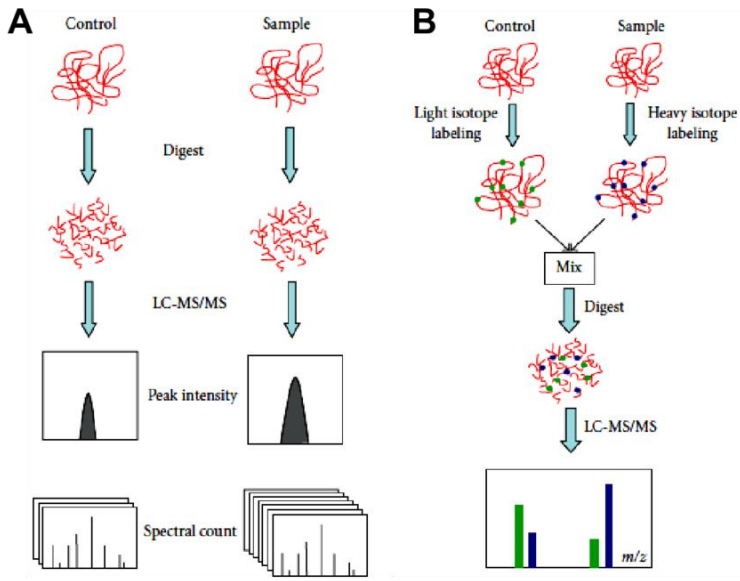


Figure 9: Comparison of LFQ (A) and SILAC (B) technologies. Deracinois et al., 2013.

Samples are simply run on the mass spectrometer without any special treatment, and peptides are identified through their specific isotopic pattern. The advantage of LFQ is that any sample can be compared to any other, while in labeled methods (such as SILAC), the only reliable comparison is between the samples that were actually mixed and run together (*Figure 9*)[68]. The disadvantage of label free experiments in general, is that conditions such as temperature, column conditions, preparation of the samples, are most likely different for each run, but it is possible to include a standard to be able to normalize the data. The comparison between different samples might be tricky on a computational level, however the advantage is that it is possible to analyze samples run on different days. Another disadvantage of LFQ is that peptides with very low abundances might get lost in the output of the run.

1.3.2.3 Antibody Microarrays

Antibody microarrays are another platform used for large-scale protein analysis. In particular, the antibody microarray used in this thesis was established in our laboratory at DKFZ and it currently consists of 2,000 well-characterized cancer-related binders. Basically, all proteins that are annotated in KEGG pathway description as cancer-associated are covered by the array, which permits quantitative analyses at a high level of accuracy and specificity [69,70] . The method is rather similar to the protein label techniques, in fact it is essential to properly label the samples with a fluorophore in order to be able to have a readout. Each sample is labeled with Cy5 and normalized using a pool mixture of different cell lines as a background, labeled with Cy3. Similar protein expression levels will appear as a yellow dot on the array, while upregulation or downregulation will be red or green respectively. It is possible to simply acquire protein data of a relative condition and compare it with other data acquired at different times, being the pool of cell lines used for the normalization the same in all the different experiments. Similarly to a SILAC experiment, it is also possible to label two different conditions with Cy5 and Cy3 and run them together as a mixture.

This will provide good insights about the outstanding protein profile differences of the two different samples [71]. *Figure 10* shows a typical antibody microarray workflow[72].

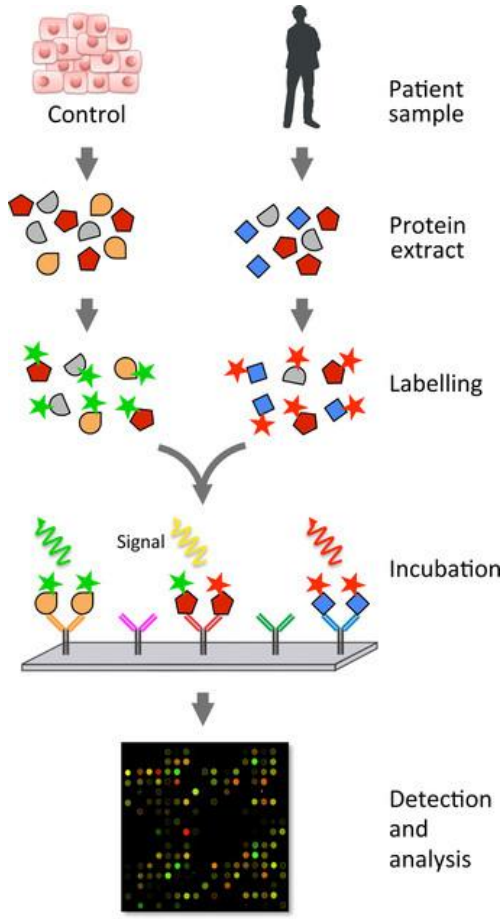


Figure 10: Schematic representation of antibody microarray. Betzen et al., 2015.

1.3.3 Metabolomics

Metabolites are small molecules, of very low molecular weight (less than 1kDa), usually obtained as an intermediate or end product of a metabolic reaction [73]. The large-scale science studying the metabolic profile of specific cell processes, is called metabolomics. Metabolomics offers an instant snapshot of the cell, tissue, organ, or sample in general, as a functional readout of a specific state. Hormones, metabolic intermediates, signaling molecules, are just some of

the categories constituting the metabolome. What we obtain through metabolomics is quantitative information about the different metabolites. In this study we used a Liquid Chromatography – Mass Spectrometry (LC-MS) system, that uses hydrophilic interaction chromatography coupled to an Orbitrap mass-spectrometer, enabling targeted metabolomics analysis of ~300 metabolites, including acids, amino acids, fatty acids, nucleotides, sugars, and sugar phosphates. METLIN was the first metabolomics database and at today it has over 500.000 molecular standards. When talking about metabolomics, it is important to have clear in mind the concept of „flux", which represents the rate in which an enzyme converts a substrate metabolite to a product. With metabolomics we try to decipher intracellular metabolic fluxes in living cells, which is rather a complicated task since flux is not a directly measurable quantity such as a metabolite concentration or enzyme levels. The only experimental technique for directly inferring intracellular flux rates is isotope tracing[74,75]. A computational method used to predict metabolic fluxes on a larger scale is constraint-based modeling (CBM). CBM allows the prediction of fluxes via genome-scale metabolic networks, consisting of thousands of metabolic enzymes, relying on simple physical-chemical characteristics (without taking in account enzyme kinetics)[76].

1.3.3.1 Metabolites And Media Preparation

Because metabolites are very small molecules, it is required to follow a strict media preparation for cell culture, in order to avoid contamination from commonly use media additives. In these regards, it is very important to use dialyzed Fetal Bovin Serum (dFBS) that is specifically depleted from low molecular weight components (< 10 kDa), that could interfere with metabolomics analysis[77]. Drugs and reagents in general added to the culturing media for specific treatments and/or conditions, will also have an impact on the metabolomics, and this must be taken in consideration when analyzing the data.

1.3.3.2 Tracing Methods

As previously mentioned, fluxes are not material entities and therefore cannot be measured in a mass spectrometer. However, they can be inferred thanks to the isotope tracers. In tracer-based metabolomics, there is a labeled precursor that, throughout several metabolic intermediates, will generate a specific distribution of ^{13}C [78,79]. The distribution of ^{13}C, together with the state-of-art knowledge in metabolic pathways, allows obtaining a clearer picture about precursor-product relationship and quantitative profiling. It is maybe more intuitive to think about a name, for example Maximillian. There are several forms, or abbreviations, for this name, we could for instance call our friend Maxim, Max, Maxy or Maxxy. All these four "products" have in common the root Max, which allows us to track back the original precursor, Maximillian. Some of them simply lost part of the name (Maxim and Max), while Maxy and Maxxy underwent an actual modification that is repeated twice in the latter nickname; nevertheless, we are still able to go back to the common precursor.

One-carbon metabolism		
Applications	**tracer**	**Metabolites readouts**
De novo synthesis of serine	[U-^{13}C]glucose	Serine M+3
Source of folate 1C units	[3-^{13}C]serine [U-^{13}C]glycine [U-^{13}C]sarcosine [U-^{13}C]formate	dTTP M+1 ATP M+1, M+2, M+3, M+4 Formyl-methionine M+1 Formate M+1
Location of serine catabolism to make cytosolic 1C units	[2,3,3-^{2} H]serine	dTTP M+1, M+2
Methylation through SAM	[Methyl-^{13}C,^{2} H_3] methionine	Methylated lysine (free or on histones)

Table 2: Tracers used for one-carbon metabolism pathway

To unravel fluxes information obtained from the tracer data, highly complex computational models, like CBM, are needed, however, intuitive interpretation can also offer reliable biological insights [77]. *Table 2* shows the tracers used for the study of one-carbon metabolism[77].

1.3.4 Genomics

The study object of genomics is the genome, which is the genomic material of an organism, more simply called DNA. For many years, genes have been the main focus of study when looking at the DNA, however non-coding DNA gained particular importance especially after discovering regulatory and epigenetic mechanisms [80,81] . A great advance in this field was provided by the introduction of genome-modifying tools, such as Transcription Activator-Like Effector Nucleases (TALEN), Zinc Fingers and more lately Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)[82] . The possibility of changing the DNA sequence in a specific way, created new approaches for the study of a specific disease. It allowed a deep investigation of specific risk polymorphisms, and opened new frontiers towards target therapies [83]. We can consider genomics as a science answering different questions about structure, function, evolution, regulation, mapping, editing and preservation of the genomes. In the era of high-throughput/large-scale studies, genomic screens are playing a major role in the discovery of new pathways and drug targets [84,85] . The possibility of combining genome wide studies with knock outs (KOs) and knock downs (KDs) on a screening level, opened new opportunities for unraveling the mechanisms at the basis of disease. In complex and heterogeneous diseases, such as cancer, it is very likely that several pathways are contributing to the malignancy, rather than just few single genes [86]. Genomic screenings allow to get an overview of the misregulated pathways, offering a better understanding for therapeutic possibilities [87–89] . Conducing KO or KD screens, gives more precise insights about gene functions and regulations. A common genomic approach aimed at finding therapeutic targets, is the so called synthetic lethality screen [30] . With synthetic lethality it is intended cell death provoked by a combination of deficiencies in the expression of two or more genes, while lack of expression of

only one of these genes cannot cause cell death. In this thesis, for instance, we want to find a gene that will be synthetic lethal to FPGS (found downregulated in the resistant cells) and cause cell death of the antifolate-resistant cells. Knock out screens will then help finding candidate genes that, together with some gene downregulations typical of the disease, can lead to cell death. Preferring KO over KD screens depends on the scientific question and on the model used. Some genes are vital to the cell [90]and therefore their KO can cause cell death, impeding the study of the associated phenotype. In these cases, a KD is preferred over a KO. Typical KD methods include siRNA and shRNA, which are small RNA sequences that cause transcript degradation upon its recognition by their complementary sequence [91].

1.3.5 Readout And Analysis Of Big Data

The massive amount of data obtained with omic sciences, has the downside of its interpretation. In fact, complex computational models and software are required to be able to analyze the obtained data. These models and software must take in consideration the biological aspects of the experiments in order to provide a fair output. Researchers are trying to provide freeware user-friendly platforms to allow big data analysis. In the proteomic context, MaxQuant [92] , Perseus [93] and David [94] are very popular softwares to process mass spectrometry data and perform consequent pathway analysis. Ingneutiy Protein Analysis (IPA)[95], offered by Qiagen, claims to be the most comprehensive pathway analysis software, with a very accurate up-to-date annotation. In metabolomics MAVEN [96] is also an open source platform for data analysis. When it comes to genomic screens, several labs provided in-house platforms, i.e. E-CRISP from Boutros lab in DKFZ [97] , and there are several on-line tools such as Bowtie[98] , used for sequences alignment. However, a problem remains about how to interpret the data and which pathways, proteins, genes, metabolites, are really significant and relevant to the study. Also, how to exclude top results that might simply be experimental biases. To this purpose, researches created annotation files containing possible lists of contaminants (biases), and a selection of parameters to obtain more or less strict results

[53,99,100] . Statistical parameters, such as p-value, are very common to filter out non-significant data or eliminate possible experimental mistakes. The enormous amount of data, requires the use of programming languages that enable the manipulation and filtering of the data at the source code level. In fact, a visual interface of all the data, allowing also filtering and modifications, would require a very powerful computer and too much memory for the data storage. MaxQuant, which offers a simple visual interface of the data, takes over 24 h to provide the data output when analyzing one single mass spectrometry run.

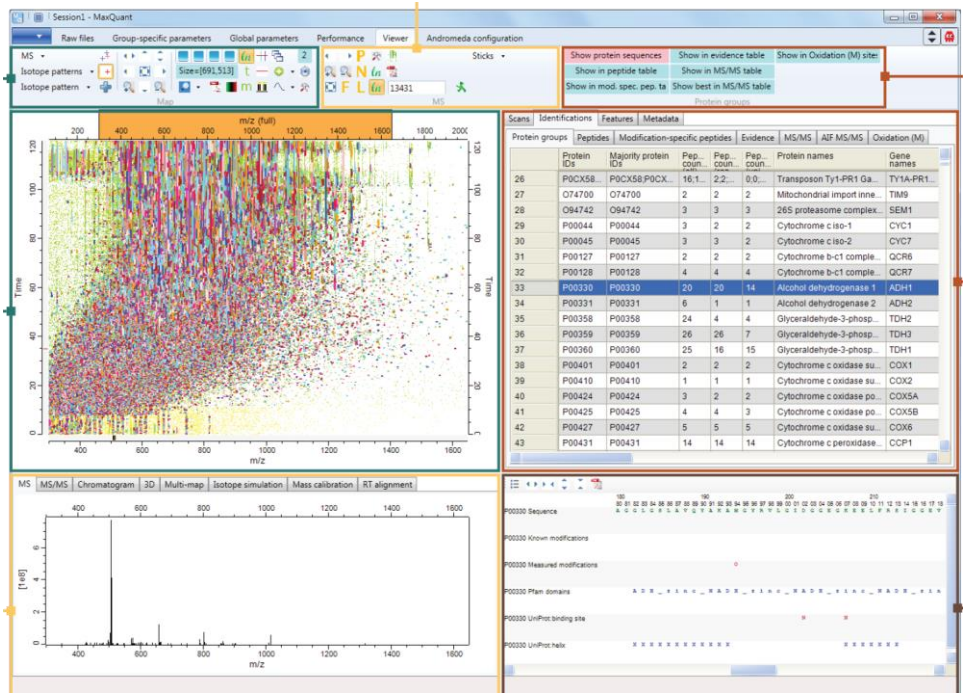


Figure 11: Example of mass spectrometry data processing with MaxQuant.

1.3.5.1 Candidate Selection

Once the big data are analyzed, the result will be a list with all the resulting hits, accompanied by significance (p-values). At this point there are two main ways for candidate selection. The most intuitive will be to simply consider the 5-10 top hits and proceed with their validation. Most of the times among the top results there are some totally unexpected candidates that might lead to the discovery of new mechanisms or pathways involved in the object of study. A

more thought-through approach consists in considering the actual biological context and stratify the results in modules, each of them specific for a determined pathway for instance. A great example is offered by the paper *"Dementia risk gens engage gene networks poised to tune the immune response toward chronic inflammatory sates"*[101]. With this last approach, the candidates selected for further validation studies, might not be on top of the list and their p-values might be slightly significant, however they might have high biological relevance in the specific context. In our specific case, we decided to stratify genomic screening results with a particular focus on enzymes and proteins involved in 1 C metabolism, and pick potential candidates that are interconnected with this pathway; while for the proteomic screenings we decided to consider candidates in common to all the three different approaches.

1.4 CRISPR Technology

The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) system was discovered and characterized in the years 1990-2010. It started being used for the actual modification of human cell lines in 2013 and since then it became more and more diffused among researchers in different areas of biology [102,103]. The discovery of CRISPR opened a whole new world for the genomic field. Genome editing became a relatively easy and fast procedure when compared to previous existing techniques, such as TALEN and Zinc Fingers (*Figure 12*) [82,104]. In particular, the huge improvement is especially seen in the generation of genetically modified mouse models. CRISPR technology in fact allows to bypass the whole procedure that involves the generation of animal chimeras and the selection of the wanted phenotype[105–107] . Previously, creating a mouse model from scratch would require years of work, nowadays CRISPR allows to obtain the model needed in a couple of months (Jackson technologies) [108] .

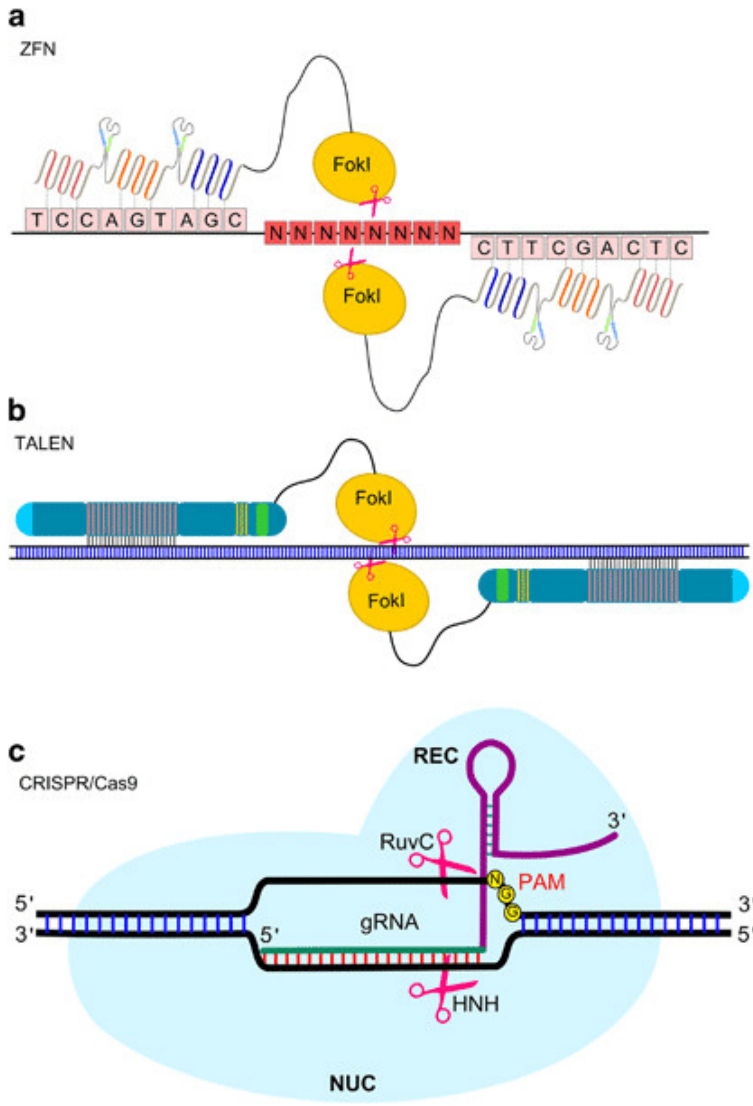


Figure 12: Schematic comparison of Zinc Finger Nucleases (a), TALEN (b) and CRISPR (c) technologies. Eid et al., 2016.

1.4.1 CRISPR Basics

CRISPR system is in reality the immune system used by bacteria to protect themselves from viral infections [109] . The short palindromic repeats are derived from fragments of viral DNA that previously infected the bacteria, and are then used to detect and destroy phages during new infections. These short sequences are part of the guide RNA (gRNA). An essential ally of CRISPR is Cas9 (CRISPR associated protein 9), which is an enzyme with endonuclease activity. Cas9 is guided by the gRNA to the specific strands of DNA that are complementary to

the gRNA sequence. Guides RNA are actually made of two parts: a crispr RNA (crRNA) that is the actual 20 nucleotide sequence complementary to the target DNA, and a tracr RNA, which constitute the binding scaffold for the Cas9. After DNA recognition, Cas9 exploits its activity and perform a double strand cut three nucleotides upstream of the PAM (protospacer adjacent motifs) site [110]. The PAM site is a short sequence of three nucleotides in the form NGG, where N can be any of the four DNA bases. The DNA is then repaired by the fastest and less-energy consuming system available, the Non Homologous End Joint (NHEJ)[111] . This pathway results in indel mutations which destroy the functionality of the gene. However, there can be cases of silent mutations or in-frame mutations that still result in the production of the protein, it is important therefore to establish CRISPR efficiency via DNA sequencing and western blot, and eventually proceed to the generation of single clones. Alternatively to the NHEJ mechanism, the cells can use the Homologous Recombination (HR) pathway [111,112], which requires a DNA strand to use as a template to repair the DNA damage (*Figure 13*).

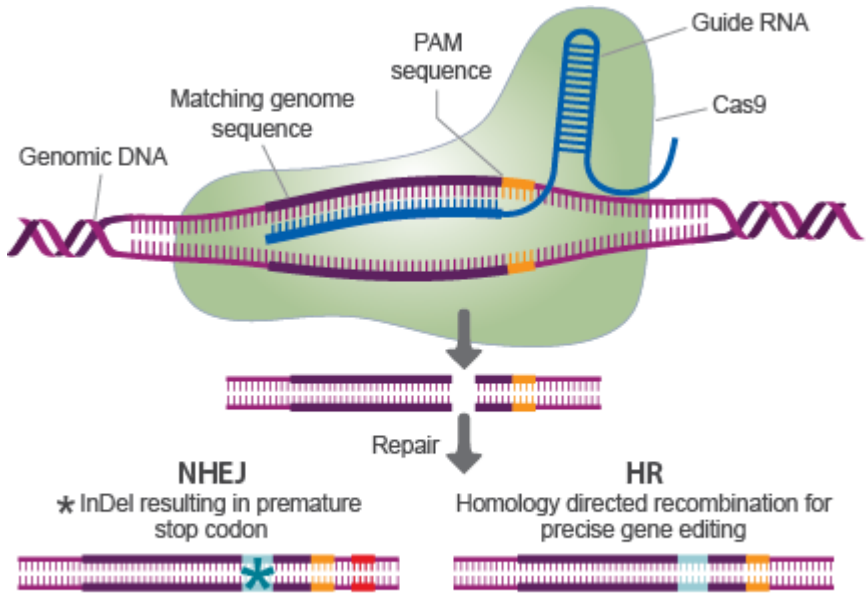


Figure 13: CRISPR technology and DNA repair. Transomic technologies.

This pathway is mostly used to introduce specific point mutations or sequences. For this purpose, a template DNA is provided to the cell together with some enhancers of the HR repair system or inhibitors of NHEJ that is usually the preferred pathway [113]. A common used enhancer is RS-1 which is stimulatory of the human homologous recombination protein RAD51. CRISPR-Cas9 is the most common CRISPR system used for genome editing, however Cpf1 nuclease is also gaining attention [114]. Cpf1 does not require tracrRNA but just crRNA, making the gRNA shorter when compared to Cas9. One of the biggest problems of the CRISPR system is the huge size of the plasmids, and in these terms Cpf1 can partially overcome this problem. Another positive aspect is that the cleavage occours 23 bp downstream from the PAM site, preventing disruption of the site itself that can therefore be used for sequent rounds of cutting.

1.4.2 CRISPR Screenings And Synthetic Lethality

CRISPR technology got extremely well combined with the high throughput screens, allowing the relatively easy planning and performing of genomic screenings to study pathways and genes involved in specific diseases, treatments or conditions. The introduction of CRISPRi (intereference) and CRISPRa (activation) also offered a possibility for studying certain genes that might be lethal in a knock out contest, further expanding screening opportunities. CRISPRi and CRISPRa use deactivated Cas9 (dCas9) that cannot perform a double strand cut, but that can still target the specific region[115,116] . The dCas9 is coupled to different transcriptional activation (CRISPRa) or repression (CRISPRi) domains that will upregulate or downregulate the gene expression. A common use of CRISPR screens, is to find those genes that act as synthetic lethal to an already pre-existant alteration. In particular, in a cancer contest, most cells that acquire resistance towards a specific drug, might show down- or up-regulation of a specific protein/enzyme/pathway. This specific alteration will not cause cell death, however when in combination with another up- or down- regulation of a specific gene, it will result lethal[117] (*Figure 14*) .

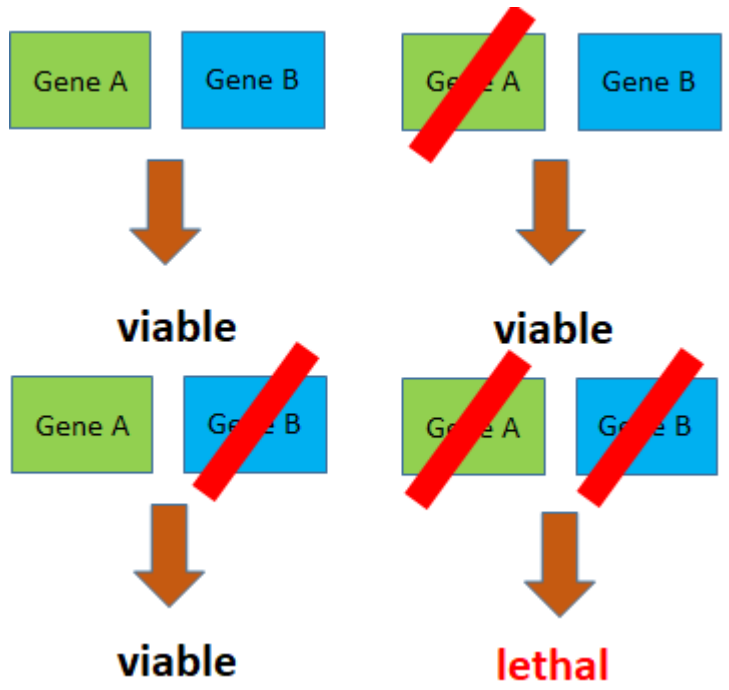


Figure 14: Schematic representation of synthetic lethality concept.

Being able to KO specific genes and looking at the phenotype, will provide essential information about pathways involved in the resistance mechanism and possible synthetic lethal targets.

2. MATERIALS AND METHODS

2.1 METHODS

In order to simplify the reading, the methods used in this thesis are divided in three sections, each of them corresponding to the respective results section.

2.1.1 Section I: The Relationship Between Folate Levels And Cytosolic Versus Mitochondrial One-Carbon Metabolism

2.1.1.1 Cell Culture

All cells were obtained from ATCC while CEM-7A cells were the gift of Prof. Yehuda G. Assaraf (Technion). None of the cell lines used here are listed in the database of commonly misidentified cell lines maintained by ICLAC and NCBI Biosample. All cell lines were confirmed to be mycoplasma free by EZ-PCR Mycoplasma Test Kit (Biological Industries, 20-700-20). Cell lines were grown in RPMI 1640 (Biological Industries) supplemented with 10% (v/v) heat-inactivated dialyzed fetal bovine serum (Sigma), 2 μM or 200 nM folic acid, 100 U ml−1 penicillin, and 100 μg ml−1 streptomycin in a 5% CO2 incubator at 37 °C.

2.1.1.2 Generation Of Knockout Cell Lines Using CRISPR–Cas9

Heterogeneous knockout cell populations were generated using lentiCRISPR v2 (Addgene catalogue number 52961). sgRNA against SHMT1 (*Table 4*) was designed based on a previous study[118]. For sgRNA cloning, lentiCRISPR v2 vector was digested with BsmBI (NEB-R0580S) and ligated with the sgRNA using Quick Ligation™ Kit (NEB-M2200L). From the ligation, 5 μL of the reaction was transformed into 50 μL of Stbl3 bacteria. The bacteria were incubated for 30 min on ice, heat shocked at 42 °C for 45 seconds, quickly placed on ice for 1-2 min, and then grown at 37 °C for 1 hr with 150 μl LB media. After being plated on ampicillin (100 μg/ml) plates for 16 h, bacterial colonies were picked and grown in 2 ml LB media plus ampicillin (100 μg/ml) for 24 h. Plasmid DNA was

extracted using NucleoSpin Plasmid EasyPure kit (MAN-740727.250). The DNA was then sent for sequencing using the U6 primer (sequence: 5' CATATTTGCATATACGATACAAGGCTC 3') to check the correct insertion of the sgRNA in the lentiCRISPR.v2 vector. To generate active lentivirus, 3.6 × 10^5 293T cells were first seeded in 100 mm plates and transfected the following day using a 4:2:1 ratio of lentiCRISPR v2: psPAX2 (Addgene catalogue number 12260):pMD2.G (Addgene catalogue number 12259) using PEI. Media were changed the next day. 72 hr post transfection, media were collected and passed through a 0.45 μm sterile filter. For cell transduction, 1 × 10^6 Jurkat cells per well were seeded in 96-well U-bottom microtiter plates and immediately pelleted by centrifugation (500 g, 4 °C, 3 min). After removing the supernatant, 100 μL of viral solution was added to each well to transduce the recipient cells in a final polybrene concentration of 10 μg/ml. After 7-9 h, cells were transferred to 6-well plates with 2 mL fresh media per well. For A549 cells, 2.5 × 10^5 cells were seeded in a 6-well plate. After 24 h, 100 μL of viral solution was added to the cells and polybrene was added to a final concentration of 10 μg ml^−1. Media were changed after 24 h. Transduced cells were selected using 2 μg ml^−1 puromycin for 72 hr and subjected to single cell cloning by limiting dilution in 96 well plates.

2.1.1.3 Generation Of HCT116 Cells With RFC Knockdown

Lentivirus production and viral transduction of HCT116 cells were performed as described above. shRNAs targeting SLC19A1 (RFC1) had sequences of CCGGCGACGGTGTTCAGAATGTGAACTCGAGTTCACATTCTGAACACCGTCGTTTTTG (TRCN0000043129) or CCGGCCAGTTATACTCCGTGTACTTCTCGAGAAGTACACGGAGTATAACTGGTTTTTG (TRCN0000043128). For control, pLKO.1 scrambled control vector (Sigma, Mission) was used.

2.1.1.4 Quantitative PCR

Cell lines were processed using Trizol RNA extraction (Thermo, 15596026). cDNA was reverse transcribed using the qPCRBIO cDNA Synthesis Kit (PCR Biosystems, PB30.11-10) with no reverse transcriptase samples serving as a negative control. Gene expression was quantified using the 2x qPCRBIO SyGreen Blue Mix (PCR Biosystems, PB20.15) and normalized to RSP11 and TPT1 (see Supplementary Table 2 for primer sequences). The CFX96 Touch™ Real-Time PCR Detection System (Biorad) was used for all reactions.

2.1.1.5 Immunoblotting

For protein extracts, cells were lysed in ice-cold Triton lysis buffer (40 mM HEPES, pH 7.4, 120 mM NaCl, 1 mM EDTA, 1% TritonX-100, 10 mM sodium pyrophosphate, 10 mM glycerol 2-phosphate, 50 mM NaF, 0.5 mM sodium orthovanadate, 1µM Microcystin-LR, 0.2 mM PMSF and protease inhibitor cocktail). Lysates were clarified by centrifugation (16,000 x g for 20 min at 4 °C) and protein concentrations were determined with Bradford assay (Biorad). Protein lysates were normalized for each experiment and equal amounts of protein were loaded into each lane of one or more polyacrylamide gels. Proteins were separated by SDS-PAGE, transferred to nitrocellulose membranes and subjected to immunoblotting with the indicated antibodies. Infrared western blot assay was performed using Odyssey Fc Imaging System (LI-COR Biosciences).

2.1.1.6 Proliferation Assay

2 × 10^4 cells per well were seeded in 24-well plates. For the 1-week growth assay, half of the media was replaced on day 3. Cell number was measured via a Multisizer Coulter Counter (Beckman Coulter).

2.1.1.7 Flow Cytometry

1 × 10^6 cells were washed twice and resuspended in 300 µL PBS. To the cell suspension, 700 µL of pre-chilled (-20 °C) 70% EtOH was added dropwise while vortexing. The samples were incubated on ice for 30 min, washed twice,

resuspended in 300 µL PBS, and propidium iodide (PI) was added at a final concentration of 20 µg ml^−1. Cells were analyzed with the BD LSR-II Analyzer (BD) using 488 nm blue laser and 575/26 filter. An unstained fixed sample was used to set the proper parameters. Cell cycle analysis was performed using the ModFit LT 5.0 software (Verity Software House).

2.1.1.8 RNA-Seq Profiling.

RNA was isolated using Trizol and mRNA libraries were prepared using the TruSeq RNA Library Preparation kit v2 (Illumina). cDNA libraries were sequenced on an Illumina HiSeq2500 to obtain >50-bp single-end sequence reads. Reads were aligned to the GRCh37 human reference genome using TopHat (2.1.0), and gene counts were obtained using HTSeq-count (0.6.1). All downstream analyses were performed in R using the Bioconductor framework. Library preparation and sequencing procedures were performed at the Technion Genome Center (Haifa, Israel).

2.1.1.9 Isotope Tracing.

Isotope tracing was performed by feeding exponentially growing cells with [2,3,3-^{2}H]-serine. Cells were fed with labeled substrates for 24 hr and metabolism was quenched immediately by adding −80 °C 50:30:20 (v/v/v) methanol:acetonitrile:water. Metabolite samples were stored at −80 °C until analysis. LC-MS was used to measure the mass-isotopomer distribution of metabolites. Measured mass-isotopomer distributions were corrected for natural abundance[119].

2.1.1.10 LC-MS Analysis.

Chromatographic separation was achieved on a SeQuant ZIC-pHILIC column (2.1 × 150 mm, 5 µm, EMD Millipore). Flow rate was set to 0.2 ml min−1, column compartment was set to 30 °C, and autosampler tray maintained 4 °C. Mobile phase A consisted of 20 mM ammonium carbonate and 0.01% (v/v) ammonium hydroxide. Mobile Phase B was 100% acetonitrile. The mobile phase linear gradient (%B) was as follows: 0 min 80%, 15 min 20%, 15.1 min 80%, 23 min

80%. A mobile phase was introduced to Thermo Q Exactive mass spectrometer with an electrospray ionization source working in polarity switching mode. Ionization source parameters were following: sheath gas 40, auxiliary gas 10, spray voltage −3.25 kV or +4.25 kV, capillary temperature 325 °C, S-lens RF level 50, auxiliary gas temperature 50 °C. Metabolites were analyzed in the range 70–1000 m/z. Positions of metabolites in the chromatogram were identified by corresponding pure chemical standards. Data were analyzed with MAVEN[120]. Absolute metabolite pool sizes were quantified using isotope-ratio with chemical standards[121].

2.1.1.11 Folate Measurements.

Intracellular folate and THF levels were measured as described previously[122].

2.1.1.12 NMR Detection Of Formate.

Suspension cells (packed cell volume = 10 µL) under exponential growth were collected and resuspended in T-75 flasks with 15 mL of fresh medium. After 24 h, 12 mL of spent medium was carefully collected post centrifugation (800 g, 5 min), and immediately mixed with 38 mL of −80 °C 50:30 (v/v) methanol:acetonitrile. After a brief vortex, samples were centrifuged (15,000 g, 30 min, 4 °C) and 12 mL of supernatant was dried by lyophilization. Dried samples were resuspended in 600 µL D2O. Samples were run on a Bruker Avance II 400 MHz NMR Spectrometer, and 512 scans were collected for each sample. Data was analyzed using TopSpin 1.3 software. Formate showed a clear peak at 8.53 ppm. Absolute quantification of formate was achieved using the standard curve method (fresh media with 0.2, 1.0, 2.0, 5.0, and 10.0 mM formate standard were processed as above to generate the standard curve).

2.1.1.13 In Vivo Xenograft Experiment.

All animal experiments were approved by the Animal Care Committee of the Technion (Haifa, Israel). For tumor growth studies, WT or SHMT1 knockout cells (1 × 10^6 cells in 200 µL 50% Matrigel) were injected in the rear flank of NOD/SCID mice and tumor growth was inspected over time using two caliper

measurements (volume = ½ [L × W2]). For tumor metabolomic studies, NOD/SCID mice were bilaterally injected on the rear flank with WT (right) and SHMT1 mutant cells (left) (1 × 10^6 cells in 100 µL PBS). Mice were sacrificed when control tumors had achieved an average size of ~100 mg (~5 mm) to avoid the development of tumor necrosis. Tumors were removed and immediately frozen in liquid nitrogen for LC-MS analysis. Isolated tumors were weighed, then 20 mg tissue was disrupted and lysed using TissueLyser II (Qiagen) and −80 °C 50:30:20 (v/v/v) methanol:acetonitrile:water. Metabolite samples were stored at −80 °C until analysis.

2.1.1.14 Metabolic Flux Analysis (MFA).

Extending upon standard MFA methodologies, our method further accounts for simulated compartment specific labeling patterns by de-convoluting whole-cell level metabolite isotopic labeling measurements into cytosolic and mitochondrial-specific simulated labeling patterns (Eq. 1).

$$\min_{v,\alpha} \sum_{i=1}^{N} \left(\frac{X_i^{WC-EXP} - \left(\alpha_i * X_i^{CY-COMP}(v) + (1-\alpha_i) * X_i^{MT-COMP}(v)\right)}{S_{X_i^{exp}}} \right)^2 \quad (1)$$

s.t.

$Sv = 0$	*Stoichiometric mass balance*
$v_{lb} \leq v \leq v_{ub}$	*Lower and upper bound on net fluxes*
$0 \leq \alpha \leq 1$	*Cytosolic relative pool size of metabolite*

where v and α_i denote the simulated fluxes and cytosolic relative pool size of metabolite i respectively. Simulated metabolite mass-isotopomer distributions of cytosolic and mitochondrial metabolites are denoted by $X_i^{CY-COMP}$ and $X_i^{MT-COMP}(v)$ and are uniquely determined by the forward and backward fluxes. We computed these mass-isotopomer distributions given all forward/backward fluxes via the Elementary Metabolite Unit (EMU approach). Measured whole cell level metabolite mass-isotopomer distribution and its standard deviation are denoted by X_i^{WC-EXP} and $S_{X_i^{exp}}$. The objective function minimizes the variance-weighted sum of squared residual of the differences between the measured mass-isotopomer distribution of metabolite i measured

on a whole-cell level and a convolution of the simulated mitochondrial and cytosolic mass-isotopomer distributions, considering the relative pool size of metabolite α_i in each compartment. The Stoichiometric matrix is denoted by S.

2.1.2 Section II: A New Effective Process For Virus Production And Cell Transduction Of Difficult Cell Lines

2.1.2.1 Cell Culture

CCRF-CEM, Jurkat and NCI-H1299 cells were cultured in RPMI medium supplemented with 10% FBS and 1% penicillin-streptomycin. Cell lines Suit-2 and Capan-1 were cultured in IMDM medium, 10% FBS and 1% penicillin-streptomycin. HEK-293T were cultured in IMDM medium, 10% FBS. All cells lines were kept at 37°C, 95% humidity and 5% CO_2. Suspension cells were kept at 10^6 cells/ml and cultured in suspension flasks. Adherent cells were kept at 70-80% confluency and cultured in standard cell culture flasks while expanding the cell lines.

2.1.2.2 Transduction Analysis

For FACS analysis, 10^6 cells were counted for each sample. Cells were spun down at 500 x g at 4°C for 3 min and washed twice with 500 µl PBS. Subsequently, the cells were resuspended in FACS buffer (PBS supplemented with 4% FBS) in a total volume of 300 µl and analysed on an LSRFortessa flow cytometry machine. Un-transduced cells were used as negative control (unstained) to set the right voltage for the mCherry detecting channel. The cell population was selected plotting forward scatter (FCS) versus side scatter (SSC); doublets were excluded from the analysis by plotting FCS width versus FCS area. The percentage of cells positive for mCherry fluorescence reflected the transduction efficiency. For microscopy, cells were analysed using a ZEISS Cell Observer microscope. Data was acquired by overlaying the brightfield image with that of the mCherry detection filter. Since the wavelength of mCherry is located in the far-red, we used the burner (a xenon lamp) to increase

the power of the signal in order to avoid underestimation of the transduction efficiency. Un-transfected cells were used to determine background.

2.1.2.3 Detailed Protocol For An Optimal Cell Transduction With Lentiviruses.

2.1.2.3.1 Virus Production

HEK-293T cells were cultured in IMDM growth medium supplemented with 10% fetal bovine serum. Cells were at 80% confluency before seeding; passage number was equal or lower than 12. Cells were seeded in culture dishes with 150 mm diameter at a cell density of 0.6 x 10^5 cells/cm^2 in 15 ml of medium. After 24 h, cells were transfected as follows:

DNA solution

- 17 µg plasmids of lentiviral sgRNA library consisting of over 260,000 molecules in vector sgLenti containing mCherry
- 17 µg psPAX2 (lentiviral packaging plasmid)
- 5.7 µg pMD2.G (VSV-G envelope expressing plasmid)
- Addition of sterile H_2O to a total volume of 60 µl
- Add 140 µl P3000 reagent
- Add 2.1 ml Opti-MEM medium

Lipofectamine mix

- 160 µl lipofectamine LTX
- 40 µl lipofectamine 3000
- 2.1 ml Opti-MEM medium

The DNA solution was added to the lipofectamine mix and incubated for 30 min at room temperature. The solution was then added dropwise to the HEK-293T cells with the 15 ml growth medium still in place. The cell culture dish was gently shaken in order to distribute the lipofectamine-DNA complex evenly. After 24 h cell growth, the medium was replaced with fresh IMDM, 10% FBS, 1% penicillin-streptomycin solution. Supernatant containing viruses was collected 72 h post transfection, loaded onto a 20 ml syringe and filtered through a 0.45 µm filter. Finally, the virus was concentrated tenfold via

centrifugation using Vivaspin 10.000 MWCO columns. The columns were centrifuged at 1000 x g for 20-30 min, the time necessary to reduce the volume to 1 ml.

2.1.2.3.2 Cell Transduction

For cell transduction, 25,000 CCRF-CEM cells per well were seeded in 96-well U-bottom microtiter plates. The cells were immediately pelleted by centrifugation at 500 x g and 4°C for 3 min; the supernatant was discarded by careful pipetting. A volume of 12.5 µl of the concentrated viral solution (equivalent to 125 µl of non-concentrated supernatant) was added to the cells and the volume of each well was adjusted to 50 µl with RPMI medium supplemented with 10% FBS without resuspending the cells. Polybrene was added to a final concentration of 10 µg/ml. Cells were incubated at 37°C, 95% humidity, 5% CO_2 for 7-9 h. Afterwards, cells were transferred to 48-well plates; the medium was replaced with 1 ml fresh RPMI, 10% FBS. Transduction efficiency was established via flow cytometry (FACS) four days after viral infection. For representative analyses with the highly complex CRISPR-library of some 260,000 sgRNAs, 10 million rather than 25,000 cells were used per microtiter well and 200 µl of the concentrated virus solution and polybrene were added. After a 7-9 h incubation, cells were transferred to a 1 l flask for cell growth. In total, 600 million cells were used, equivalent to 60 microtiter plate wells.

2.1.2.4 Alternative Protocols For Virus Production.

The other protocols used for producing lentiviruses in HEK293T cells in comparison were as follows:

2.1.2.4.1 HEPES-Buffered Saline (HBS) Protocol

HEK293T cells were seeded at a density of $0.6x10^5$ cells/cm^2 in 6 cm^2 culture dishes in 4 ml IMDM growth medium supplemented with 10% FBS and grown for 24 h.

DNA solution

- 6.42 µg plasmids of lentiviral sgRNA library
- 6.42 µg psPAX2
- 2.14 µg pMD2.G
- Add sterile H_2O to a total volume of 438 µl.
- Add 62 µl 2 M $CaCl_2$.

Dropwise, 500 µl of 2x HBS buffer (2x HBS: 50 mM HEPES, 280 mM NaCl, 1.5 mM Na_2HPO_4, pH 7.0; sterilised by filtration) were added to the DNA solution while mixing it by blowing air through the solution with a Pasteur pipette at the bottom of the tube. The resulting 1 ml transfection solution was added dropwise to the HEK293T cells while shaking the cell culture dish gently to distribute the solution evenly. After 10 h, the growth medium was replaced with 5 ml of fresh IMDM, 10% FBS, 1% penicillin-streptomycin solution. Supernatant containing viruses was collected 72 h post transfection, loaded onto a 20 ml syringe and filtered through a 0.45 µm filter.

2.1.2.4.2 Polyethyenimine (PEI) Protocol

HEK293T cells were seeded at a density of $0.6x10^5$ cells/cm^2 in 10 cm^2 culture dishes in 10 ml IMDM growth medium supplemented with 10% FBS and grown for 24 h.

DNA solution

- 8 µg plasmids of lentiviral sgRNA library
- 4 µg psPAX2
- 4 µg pMD2.G
- Add Opti-MEM medium to a total volume of 250 µl.

For a PEI:DNA ratio of 3:1, 48 µl of PEI solution (1 µg/µl) were added to 202 µl Opti-MEM. The DNA solution was added to the PEI solution; mixed well, spun briefly to collect all the liquid and incubated at room temperature for 20 min.

Then, the transfection solution was added dropwise to the HEK293T cells while shaking the cell culture dish gently to distribute the solution evenly. After 24 h, the growth medium was replaced with 10 ml of fresh IMDM, 10% FBS, 1% penicillin-streptomycin solution. Supernatant containing viruses was collected 72 h post transfection, loaded onto a 20 ml syringe and filtered through a 0.45 µm filter.

All other transfections were performed as suggested by the respective reagent manufacturers using the same cells and plasmids as above:

2.1.2.4.3 Lipofectamine 3000 Protocol

https://www.thermofisher.com/de/de/home/brands/product-brand/lipofectamine/lipofectamine-3000.html

2.1.2.4.4 Lipofectamine LTX Protocol

https://www.thermofisher.com/de/de/home/brands/product-brand/lipofectamine/lipofectamine-ltx-reagent.html

2.1.2.4.5 Jetprime Protocol

https://www.polyplus-transfection.com/products/jetprime/#files-btn

2.1.2.4.6 Transit-LT1 Protocol

https://www.mirusbio.com/products/transfection/transit-lt1-transfection-reagent

2.1.3 Section III: Unraveling FPGS-dependent antifolate resistance: proteomic and synthetic lethality CRISPR screening

2.1.3.1 Antibody microarray

Cell lysate. Cellular proteome was extracted using lysis buffer as reported previously[123]. The lysis buffer was constituted of 50 mM Bicine buffer (pH 8.5) containing 20% glycerol, 1.0 mM magnesium chloride, 5.0 mM EDTA, 1.0 mM

phenylmethanesulfonyl fluoride, 1.0 IU/ml benzonase, Halt Protease and Phosphatase Inhibitor Cocktail, 0.5% Nonidet P-40 substitute, 1.0% Cholic acid, 0.25% n-dodecyl-β-maltoside and 0.5% amidosulfobetaine14. Briefly, cells were washed with cold PBS, 100 ul of lysis buffer per 1*10^6 cells were added each sample and incubated on ice for 30 min. The cells were subsequently pipetted up and down with a syringe through a 23G needle followed by centrifugation at 18,000 g for 20 min at 4°C. Afterward, the supernatant was transferred to a fresh microcentrifuge tube and total protein concentration was determined using Bicinchoninic Acid (BCA) protein assay reagent kit according to manufacturer's instructions.

Sample labeling. For the labeling, samples were adjusted with 0.1M Bicine buffer to the same concentrations. The samples were labelled with the fluorescent NHS-ester dye DY-649P1 (Cy-5) in a molar ratio of 7.5 (dye/protein) in dark eppendorf tubes and were shaken at 4°C for 2 h. Afterwards, the unlabeled fluorescent dye was quenched by adding 10% glycine with a constant shaking for 20 min. For the reference background, cellular proteomes were pooled and labeled with DY-549P1 (Cy-3) as described previously[124].

Microarray incubation, scanning and image processing. Antibody microarrays were processed as previously described[70]. Before the incubation with the labeled samples, the arrays were equilibrated at room temperature for 30 min, followed by washing twice with TBST. The arrays were then blocked with 5 ml of 10% non-fat dry milk, prepared in TBST, in quadriPERM chambers and left on a shaker at room temperature for 3 h. The arrays were subsequently washed with TBST and incubated overnight with 50 µg of the labelled samples and pool reference in 1% milk at 4°C with constant shaking. Afterwards, the arrays were washed four times with TBST and once with distilled water followed by drying in a ventilated oven at 25°C. Consequently, arrays were scanned using a Tecan power scanner at constant laser power and the resulting images were analyzed with GenePix Pro 6.0 software.

2.1.3.2 SILAC analysis

CCRF-CEM, MTA-C3 and MTA-R15 cell lines were seeded in triplicates in 6 well plates at a concentration of 2*10^5 cells per well. Cells were grown in Heavy labeled or Light labeled media. Heavy labeled media was RPMI-lys 8- Arg 10- kit from SILANTES, presenting labeled lysine and arginine, while Light labeled media had the exact same composition but unlabeled lysine and arginine. After 7 days, cell lysates were prepared using 100 uL of RIPA buffer (additioned with protease inhibitor cocktail, Roche and DTT 1% w/v) per 1*10^6 cells. Cell lysate was resuspended several times through a pipette to mince particles and obtain an omogenous solution. One microliter of Benzonase (Merck, 25 units/ ul) was added per 100Ul LYSATE. Samples were incubated at 4°C on an orbital shaker at 300 rpm for 1 h. Subsequently samples were centrifuged at 15000 g for 30 min at 4°C to separate insoluble material. Supernatant was transferred to a new 1.5 ml tube and stored at -20 °C. Samples were sent for analysis to the mass spectrometry core facility of DKFZ institute.

2.1.3.3 LFQ analysis

CCRF-CEM, MTA-C3 and MTA-R15 cell lines were seeded in triplicates in 6 well plates at a concentration of 2*10^5 cells per well. Cells were grown in RPMI low folate media supplemented with 10% dFBS, and 1% penicillin streptomicyn. After 7 days, cell lysates were prepared as described in SILAC analysis and samples were sent for LFQ analysis to the Radboud institute for molecular life sciences in Nijmegen.

2.1.3.4 CCRF-CEM Cas9 and MTA-C3 Cas9 cell lines

CCRF-CEM Cas9 and MTA-C3 Cas9 cell lines were obtained by transducing the respective parental cell lines CCRF-CEM and MTA-C3, with lentiCas9-Blast viruses. The new virus production and transduction protocol described in this thesis was used for this purpose. After establishing the right working concentration of blasticidin for each cell line, blasticidn was added to the cells one day after transduction and kept in culture for 72 h. After blasticidin selection, the transduced cells were diluted to 4*10^5 cells/mL. A volume of 10

µl of this solution was added to 990 µl complete RPMI medium. A volume of 12.5 µl of the new solution was added to 10 ml of RPMI media, 20% FBS, 1% penicillin-streptomycin solution. At this point the solution has a final concentration of 5 cells/ml. A volume of 100 µl of 5 cells/ml solution was transferred to each well of a 96 well-plate together with 100 µl of RPMI media, 20% FBS, 1% penicillin-streptomycin solution, to reach a final volume of 200 µl. The plate was incubated at 37°C, 95% humidity, 5% CO_2 till single clones started growing and were therefore transferred to 24-well plates for further expansion and testing of the clones. Cas9 functionality was tested by knocking out the surface cell receptor CD44 (*Table 4*) via CRISPR technology, using sgLenti as backbone and inserting the sgRNA using AarI cloning sites. FACS analysis were performed to quantify remaining CD44 signal and therefore Cas9 functionality.

2.1.3.5 FACS analysis

A total number of 1 × 10^6 cells was washed twice and resuspended in 100 µl PBS. The samples were incubated in the dark on ice with anti CD44 antibody, dilution 1:50, for 1 h. Afterwards samples were washed twice with PBS, centrifuged at 300 g for 3 min at 4 °C and resuspended in 300 ul PBS. Propidium Iodide (PI) was added to the cell suspension at a final concentration of 1 ug/ml. Cells were analyzed with the Fortessa Analyzer (BD). An unstained and untransduced sample was used to set the proper parameters.

2.1.3.6 CRISPR screening setup and sample preparation

CCRF-CEM and MTA-C3 cell lines were used for the CRISPR screening. CCRF-CEM Cas9 and MTA-C3 Cas9 single clones were obtained transducing the original cell lines with viruses generated from lentiCas9-Blast plasmid (Addgene #52962). Cas9 expressing cell lines were transduced with sgLenti viruses at an initial coverage of 250 X. For both transductions, the new transduction protocol described in this thesis was used. Cells were grown in 1 L roller bottles (Corning CLS431644) for 20 days, corresponding to 14 doublings (cell doubling time: 34 h). DNA was extracted as follow. Cells corresponding to 1000 X coverage were washed twice with PBS and pellet was

resuspended in 20 ml of buffer P1 (Qiagen), RNase A was added at a final concentration of 100 ug / ml, and SDS at a final concentration of 0.5 %. Samples were incubated at 37°C for 30 min. Proeinase K was added at a final concentration of 100 ug / ml and samples were incubated at 55°C for 30 min. Afterwards samples were passed three times through a 18G needle and a 22G needle. First and second round PCRs (primers sequences in *Table 3*) were done using NEB Next ultra II, following manifacturer's instructions. Library samples were measured with Qbit, diluted to 10 nM and sent for sequencing with a coverage of 1000 X using Illumina HiSeq 2000 V4 platform, 50 bp single reads.

2.1.3.7 Dead-Live ficoll isolation

A total number of 50*10^6 cells were resuspended in 25 ml media and aliquoted in a 50 ml falcon tube. 12,5 ml Ficoll Histopaque-1077 were carefully added to the media by gently pipetting the ficoll from the bottom of the falcon tube. The sample was then centrifuged at 700 g for 20 min at 4°C. The inner layer contains the living cells.

2.1.3.8 Proliferation assay

A total number of 2 × 10^4 cells per well were seeded in 2 ml in 24-well plates. For the 1-week growth assay, half of the media was replaced on day 3. Cell number was measured via a Multisizer Coulter Counter (Beckman Coulter).

2.1.3.9 AHCY, AMD1, OIP5, SAT2, BAX1 and PTPN6 KO.

Candidates knockouts were performed inserting the respective sgRNAs (*Table 4*) in sgLenti vector using AarI cloning sites. Virus were produced following the guidelines of the new protocol described in this thesis.

2.1.3.10 Isotope tracing, LC-MS analysis and folates measurements were performed as described in the first section of methods.

2.2 MATERIALS

2.2.1 Cells

- BxPC3 (ATCC CRL-1687); adherent culturing.
- Capan-1 (ATCC HTP-79); adherent culturing.
- CCRF-CEM (ATCC CCL-119); suspension culturing.
- HEK-293T (ATCC CRL-3216); adherent culturing.
- Jurkat (ATCC TIP-152); suspension culturing.
- NCI-H1299 (ATC CRL-5803); adherent culturing.
- Suit-2 (JCRB1094); adherent culturing.
- A549 (ATCC CCL-185)
- HCT116 (ATCC CCL-247)
- THP1 (ATCC TIB-202)
- HL-60 (ATCC CCL-240)
- MTA-C3 and MTA-R15 [50]

2.2.2 Cell Culture Medias And Solutions

- DMSO cell culture grade (AppliChem GmbH)
- Dulbecco's Modified Eagle Medium (DMEM) (Gibco, Life Technologies)
- Fetal Bovine Serum (FBS), heat-inactivated (Gibco, Life Technologies)
- Dialyzed Fetal Bovine Serum (Gibco, Life Technologies)
- RPMI Lysine (8) Arginine (10) Kit (Silantes)
- Ficoll Histopaque-1077 (Sigma-Aldrich)
- Iscove's Modified Dulbecco's Medium (IMDM) (Gibco, Life Technologies)
- L-Glutamine 200mM (Gibco, Life Technologies)
- Penicillin 1,000u/ml -Streptomycin 100µg /ml (Gibco, Life Technologies)
- Phosphate buffered saline (DPBS)+ MgCl2 (Gibco, Life Technologies)
- Phosphate buffered saline (PBS) (Gibco, Life Technologies)
- RPMI 1640 with L-glutamine and phenol red (Gibco, Life Technologies)
- Trypsin-EDTA (0.05%) with phenol red (Gibco, Life Technologies)
- RPMI 1640, Custom formulation, no folic acid (Biological Industries)

2.2.3 Antibodies

- anti-SHMT1 (Cell signaling, D3B3J, 1:250)
- anti-SHMT2 (Sigma, HPA020549, 1:250)
- anti-MTHFD1 (Santa Cruz Biotechnology, sc-271412, 1:1000)
- anti-MTHFD1L (Proteintech, 16113-1-AP, 1:1,000)
- anti-MTHFD2 (GeneTex, GTX104990, 1:1,000)
- anti-RFC (Santa Cruz Biotechnology, sc-390948, 1:100)
- anti-FPGS (Abcam, ab184564, 1:1000)
- anti-GAPDH (EMD Millipore, CB10001, 1:10,000).
- IRDye 680RD goat anti-rabbit (Licor, 926–68071, 1:15,000)
- IRDye 800CW goat anti-mouse (Licor, 926–32210, 1:15,000)
- Anti-CD44 conjugated PE-Cy7 (Biolegend, BJ18, 1:50)

2.2.4 Reagents

- Oligonucleotides were obtained from Biomers
- HEPES (Sigma-Aldrich)
- 5-methyltetrahydrofolic acid (Cayman Chemical)
- (6S)-tetrahydrofolic acid (Cayman Chemical)
- Lipofectamine 3000 with P3000 reagent (Thermo Fisher)
- Lipofectamine LTX (Thermo Fisher)
- Opti-MEM reduced serum medium (Thermo Fisher)
- PBS (Life technologies)
- Penicillin-streptomycin (Life Technologies)
- pMD2.G (Addgene)
- Polyethyenimine (PEI) solution (Polysciences,)
- Polybrene (Merck)
- psPAX2 (Addgene)
- [2,3,3-^{2}H]-serine (Cambridge Isotope Laboratories)
- sgRNA library of over 260,000 molecules[125] in sgLenti vector (Addgene)
- Folic acid (Sigma)
- sodium formate (Sigma)
- Sodium chloride (NaCl) (Sigma-Aldrich)

- Sodium phosphate dibasic (Na_2HPO_4) (Sigma-Aldrich)
- Blasticidin (Thermo Fisher Scientific)
- Puromycin (Thermo Fisher Scientific)
- Quick ligase (NEB)
- Benzonase (Merk Millipore)
- Bicine (Biomol)
- NEB Next Ultra II (NEB)
- Buffer P1 (qiagen)
- Proteinase K (Thermo Fisher Scientific)
- RNase A (Thermo Fisher Scientific)
- Nitrocellulose membrane (Protran 0.45) (GE Healthcare)
- Nuclease free water (Ambion)
- RIPA buffer (Thermo Fisher Scientific)
- Tween 20 (Sigma)
- Buffers were prepared according to Arigo Biolaboratories recipes (https://www.arigobio.com/)

2.2.5 Plastic ware

- 48-well suspension microtiter plate (Greiner-Bio-One)
- 96-well microtiter plate, U-bottom (Greiner-Bio-One)
- Cell culture dish, 150 mm (Corning)
- Cell culture standard flasks (Greiner Bio One)
- Cell culture suspension flask (Greiner Bio One)
- Filter, 0.45 µm (Merck)
- Syringe, 20 ml (BD)
- Vivaspin 20 10.000 MWCO columns (Sartorius)
- Falcon tubes (15 ml, 50 ml) (Greiner Bio-One)
- Non-treated roller bottle (Corning)
- Serological pipettes (5 ml, 10 ml, 25 ml) (Greiner Bio-One)

2.2.6 Equipment

- Centrifuge: Eppendorf centrifuge 5810 R
- Flow cytometer equipped with yellow-green laser: LSRFortessa (BD Biosciences)
- Fluorescent microscope: Zeiss Cell Observer SD
- FACS Fortessa (BD)
- LC-MS (Thermo Scientific)
- Tecan infinite M200 plate reader (Tecan)
- Trans-Blot turbo transfer system (Biorad)
- Odissey (Licor)
- Vi-CELL XR cell viability analyser (Beckman Coulter)
- Beckman Coulter Z2 (Beckman Coulter)
- MP 230 ph-meter (Mettler Toledo)
- Nanodrop ND-1000 (Thermo Fisher Scientific)
- Cryotubes 1 ml (Thermo Fisher Scientific)
- Analytical balance scale 434-33 (Kern)
- Safe lock tubes 1.5 ml and 2 ml (Eppendorf)
- Water bath (Grant Instruments)

Primer	Sequence 5' – 3'
SHMT1 fw	TCTCTGATGCAGCTGTGTCA
SHMT1 rev	GCAGTGTTCAAATGGGGAGG
SHMT2 fw	ACTCAGACTGGGGAAGCAAA
SHMT2 rev	AGGTCCTGAGTGTGAAAGGG
U6 fw	CATATTTGCATATACGATACAAGGCTC
Sequencing primer sgLenti	GAGACTATAAGTATCCCTTGGAGAACCACCTTGTTGG
1st PCR screening fw	GGCTTGGATTTCTATAACTTCGTATAGCA
1st PCR screening rev	CGGGGACTGTGGGCGATGTG
2nd PCR screening fw	AATGATACGGCGACCACCGAGATCCACAAAAGGAAACTCACCCTAAC
2nd PCR screening rev	CAAGCAGAAGACGGCATACGAGAT-(N)6-GTGACTGGAGTTCAGACGTG

Table 3: Sequences of primers used in this thesis.

Guide RNA	Sequence 5' – 3'
SHMT1	GAACGGGGCGTATCTCATGG
SHMT2	GAGAAGGACAGGCAGTGTCG
AHCY	ACCGCTCCCGCATACGCATC
SAT2	TCTATGTGATGCCGGAATAT
BAX1	TTTCTGACGGCAACTTCAAC
PTPN6	TAAGACCTACATCGCCAGCC
CD44	TCGCTACAGCATCTCTCGGA

Table 4: Sequences of sgRNA used in this thesis.

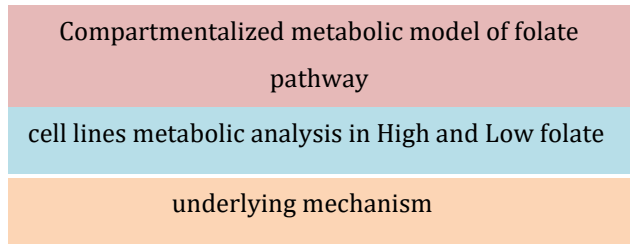
Index (N6)	sequence
#1	ATCCTC
#2	ATCACG
#3	ACTTGA
#4	TAGCTT
#5	GGCTAC

Table 5: Sequences of indexes used for Illumina screening library sequencing.

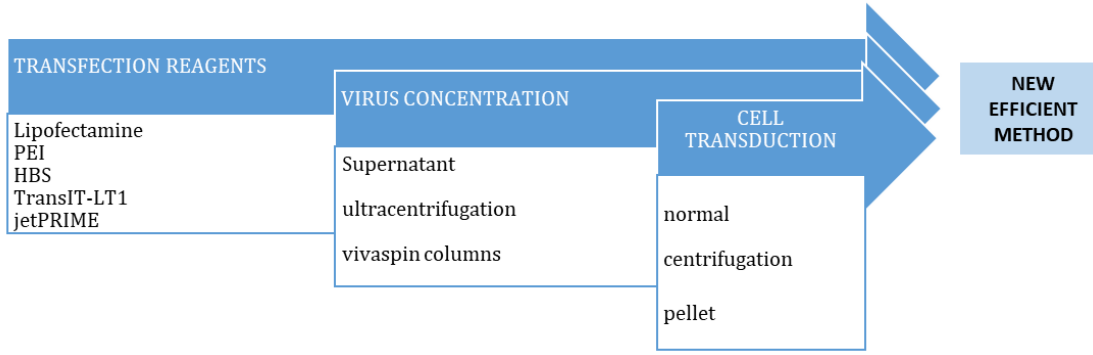
3. RESULTS

This thesis work focuses on FPGS-dependent antifolate resistance, and in particular wants to give an answer about proteomic and/or metabolic pathways involved in the mechanisms of resistance. The ultimate goal is to find a possible target to restore drug sensitivity in the resistant cells. To this aim, we performed a synthetic lethality CRISPR screening and combined these data with the knowledge obtained from different proteomic screenings. However, in order to obtain reliable data, we first decided to better characterize our study model under a metabolic profile. It is quite intuitive that studying antifolate resistance means focusing mainly on one carbon metabolism, which sees the folate pathway as its core. It is therefore important to first unravel the folate pathway and how folate levels affect cell metabolism, before moving to antifolate resistance. The following results can overall be divided in three sections:

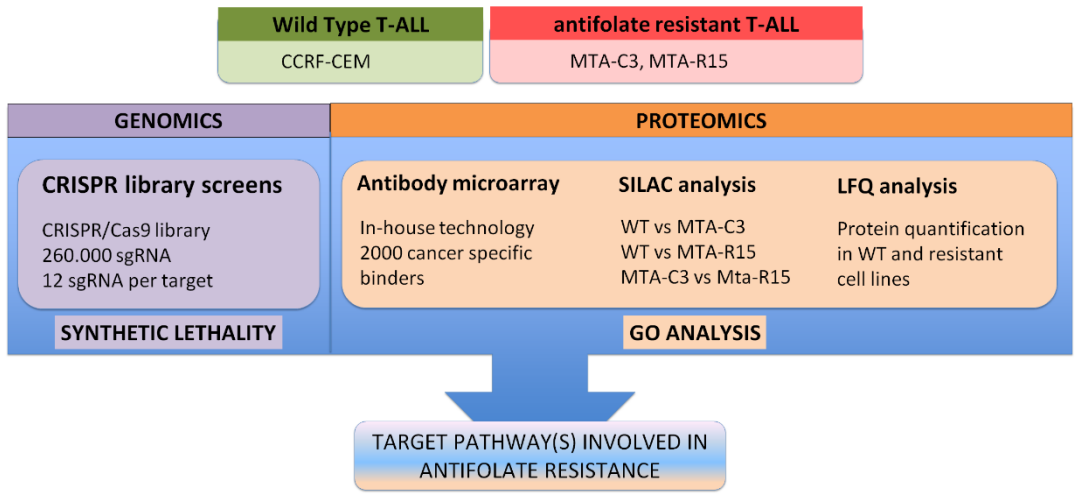
1- The relationship between folate levels and cytosolic versus mitochondrial one-carbon metabolism



2- A new effective process for virus production and cell transduction of difficult cell lines



3- Unraveling FPGS-dependent antifolate resistance: proteomic and synthetic lethality CRISPR screening



3.1 THE RELATIONSHIP BETWEEN FOLATE LEVELS AND CYTOSOLIC VERSUS MITOCHONDRIAL ONE-CARBON METABOLISM

The results of this section are part of a manuscript entitled "Tumor dependence on cytosolic one-carbon metabolism is determined by the cellular capacity to retain folates" which is currently under review and is part of the work conduced in Israel in collaboration with Shlomi lab. It is quoted in this section wherever it seemed useful, including the majority of figures and legends.

3.1.1 The Cytosolic Pathway Is The Major Net Contributor Of 1C Units For Pyrimidine Biosynthesis Under Physiological Folate Conditions

Our understanding of 1C metabolic flux in cancer mostly relies on experiments performed in standard tissue culture conditions, in which the total folate concentration is an order of magnitude higher than that in human serum: 150-450 nM in serum[3,4] versus 2.2 μM and 9 μM in RPMI and DMEM, respectively. To examine the relative contribution of cytosolic versus mitochondrial folate pathway to pyrimidine biosynthesis under physiological folate levels, we performed stable isotope tracing across a panel of human cancer cell lines. Cells were fed [2,3,3-^{2}H]-serine and the incorporation of deuterium labeling in synthesized thymidine triphosphate (dTTP) was monitored. Thymidine with two deuterium atoms (M+2) is synthesized via the cytosolic folate cycle through SHMT1, and that with one deuterium (M+1) is synthesized within mitochondria through SHMT2; mitochondrial serine-derived methylene-THF is oxidized to formate, transported to the cytosol, incorporated into THF, and eventually reduced back to methylene-THF, synthesizing dTTP (*Figure 15a*). We found that lowering the total folate concentration in culture media to a physiological level (from 2 μM to 200 nM folate) led to a ~10-fold increase on average in the relative contribution of the cytosolic pathway across different cell lines (dTTP M+2/M+1; *Figure 15b*). Under physiological folate, the cytosolic pathway became the predominant contributor for 1C units in tumors of various origins (dTTP M+2/M+1 > 1; *Figure 15b*), including T-cell acute lymphoblastic leukemia, glioblastoma, and

non-small-cell lung carcinoma. Notably, media folate levels had no effect on cell growth rate, suggesting that relying primarily on cytosolic 1C flux under physiological folate conditions is sufficient to meet anabolic demands (*Extended Data Figure 1a*). The increased reliance on SHMT1 for producing 1C units under physiological folate media was associated with a drop in whole-cell serine hydroxymethyl-transferase flux, suggesting a marked decrease in SHMT2 flux: A drop in whole-cell serine hydroxymethyl-transferase flux is evident by a ~5-fold drop in serine consumption from culture media in CCRF-CEM cells grown in physiological versus in high folate conditions (p-value < 0.001; Fig. 1c); and consistent drop in the secretion of the two products, glycine and formate. While glycine was excreted under high folate media, substantial glycine uptake flux was observed under physiological folates (~0.8 mM/h; ~50% of serine uptake). Increased cellular reliance on extracellular glycine versus *de novo* production from serine was further evident by a significant drop in glycine M+1when feeding [2,3,3-^{2}H]-serine under physiological versus high folate media (p-value < 0.001; *Figure 15d*). Formate secretion rate showed a significant ~7-fold drop under physiological folate (p-value < 0.001; *Figure 15c*). A decrease in mitochondrial 1C flux under physiological folate condition was further evident by a drop in the contribution of serine catabolism to the reduction of mitochondrial NADP+ via methylenetetrahydrofolate dehydrogenase 2 (MTHFD2). Feeding [2,3,3-^{2}H]-serine under physiological folate significantly dropped the fractional deuterium labelling of proline (p-value < 0.001 *Figure 15e*). Notably, we found a similar major drop in mitochondrial derived 1C flux in several cell lines grown under physiological folate conditions (*Figure 15e* and *Extended Data Figure 1b*). These results suggest that while 'overflow metabolism' in mitochondrial serine catabolism do appear in some cultured cell lines under physiological folate and *in vivo*[126,127], it is not ubiquitous under physiological conditions. To quantitatively determine the net contribution of the cytosolic versus mitochondrial 1C flux to pyrimidine biosynthesis, we utilized Metabolic Flux Analysis (MFA; Methods); overcoming a potential bias in the interpretation of the cytosolic versus mitochondrial 1C flux based on [2,3,3-^{2}H]-serine tracing, due to isotope exchange effects (enzymes close to chemical

equilibrium simultaneously catalyzing oxidative and reductive flux). To enable inference of mitochondrial and cytosolic 1C fluxes, we employed a variant of MFA, searching for the most likely mitochondrial and cytosolic fluxes such that whole-cell level measured isotopic labeling of serine and glycine match a convolution of the simulated labeling of these metabolites in the two compartments. The result further supports the cytosolic pathway being the major net contributor of 1C units for pyrimidine biosynthesis under physiological folates (*Figure 15f*).

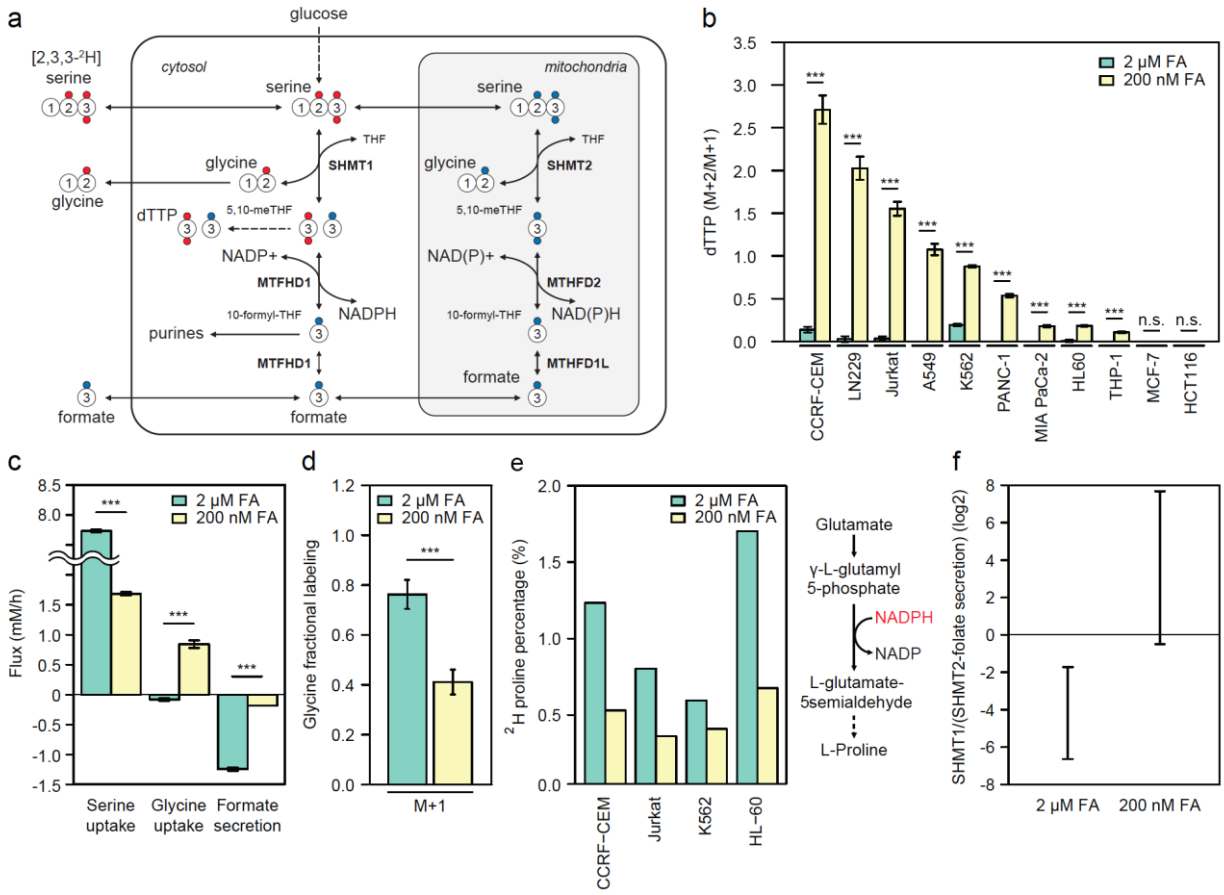


Figure 15: **The cytosolic folate cycle is the prime source of 1C units in a variety of cancer cells under physiological folate levels.** (a) A compartmentalized view of the folate cycle and [2,3,3-^{2}H]-serine tracer for inferring cytosolic versus mitochondrial 1C flux into pyrimidines. (b) Measured dTTP M+2/M+1, representing the cytosolic over the mitochondrial 1C flux, when feeding various cell lines with [2,3,3-^{2}H]-serine in 2 μM and 200 nM media folate levels. (c) The measured uptake/secretion rate of serine, glycine, and formate in CCRF-CEM in 2 μM and 200 nM media folate levels. (d) The fractional labeling of intracellular glycine when feeding CCRF-CEM with [2,3,3-^{2}H]-serine in 2 μM and 200 nM media folate levels. (e) The isotopic labeling of intracellular proline when feeding leukemia cells with [2,3,3-^{2}H]-serine in 2 μM and 200 nM media folate levels. Proline biosynthesis involves oxidation of mitochondrial NADPH, hence its labeling represents the contribution of serine catabolism to NADPH production. (f) The ratio between the cytosolic and mitochondrial 1C metabolic flux towards purine and pyrimidine biosynthesis, inferred based on computational Metabolic Flux Analysis (MFA). n.s. not significant. ***P < 0.001 by two-sample t-test. Data are mean ± SD, n = 3 independent biological replicates.

3.1.2 Downregulation Of Mitochondrial 1C Flux Under Physiological Folate Is Due To The Depletion Of Intracellular Reduced Folates

Exploring how cells regulate cytosolic versus mitochondrial 1C flux in response to folate availability, we found that media folate level had no effect on the concentration of key 1C metabolic enzymes in Jurkat cells (*Figure 16a*); or an effect on the global mRNA expression pattern (*Extended Data Fig. 2a-b*) or 1C enzyme levels (*Figure 16a*). Metabolomics analysis using liquid-chromatography mass-spectrometry (LC-MS), revealed a significant >50-fold increase in the intracellular concentrations of the purine biosynthesis intermediates 5′-phosphoribosyl-5-aminoimidazole-4-carboxamide (AICAR) and 5′-phosphoribosyl-glycinamide (GAR) in cells under physiological folate (*Figure 16b*, *Extended Data Figure 2c*); indicating a potential drop in intracellular folate pools[128]. Indeed, the intracellular concentrations of folate and its reduced form (THF) in Jurkat and CCRF-CEM showed a significant 80-90% drop while switching to primarily cytosolic 1C flux under physiological folate conditions (*Figure 16c-d*). In comparison, MCF-7 and HCT116 that primarily rely on mitochondrial 1C flux under physiological folate conditions showed significantly higher capacity to retain intracellular folates (*Figure 16c-d*). These results suggest substrate level downregulation of mitochondrial 1C flux in CCRF-CEM and Jurkat under physiological folate due to the depletion of intracellular reduced folates. In support of metabolic regulation of mitochondrial 1C flux, we found that switching Jurkat cells from growth on physiological to high folate media resulted in rapid induction of mitochondrial 1C flux (dTTP M+2/M+1 dropped ~16% in 1 hr after switching from high to low folate media; p-value < 0.05; *Figure 16e*).

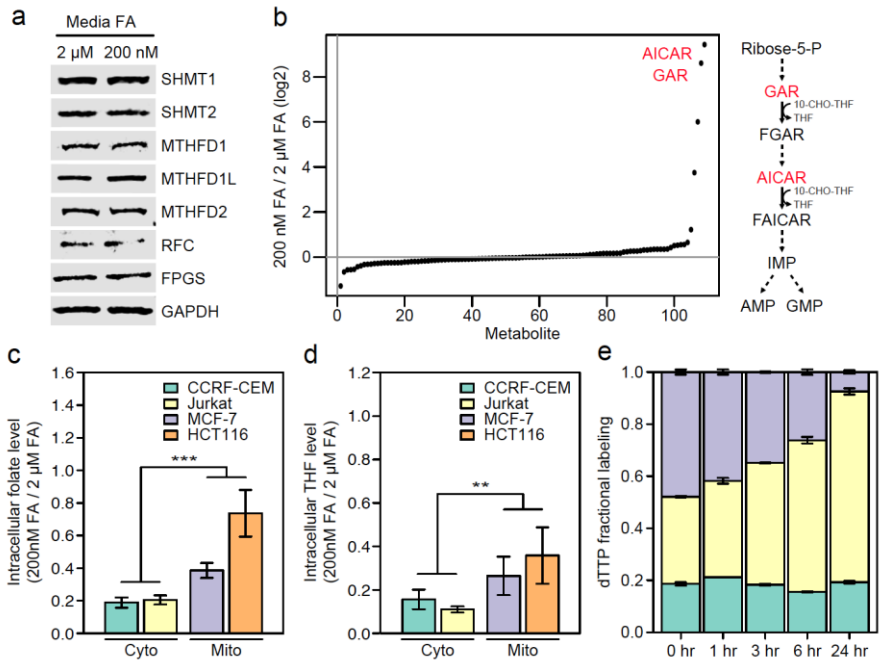


Figure 16: **Intracellular folate levels determine the cytosolic versus mitochondrial 1C flux.** (a) Western blot of 1C enzymes in Jurkat cells under 2 μM and 200 nM media folate. (b) Ratio between intracellular metabolite concentrations in Jurkat cells grown in 200 nM and in 2 μM folate media. (c-d) Intracellular folate (c) and THF levels (d) in 2 μM and 200 nM media folate. CCRF-CEM and Jurkat cells are more dependent on the cytosolic folate pathway (Cyto) while MCF-7 and HCT116 cells are more dependent on the mitochondrial pathway (Mito). (e) The mass-isotopomer distribution of dTTP in Jurkat cells fed with [2,3,3-^{2}H]-serine. 200 nM folate media was switched to media with 2 μM folate and dTTP labeling was measured at different time points. **P < 0.01 and ***P < 0.001 by two-sample t-test. Data are mean ± SD, n = 3 independent biological replicates.

3.1.3 RFC Expression Determines The Intracellular Folate Concentration, Which Modulates The Cytosolic Versus Mitochondrial 1C Metabolism

To explore the mechanism underlying cell line specific reliance on the cytosolic versus mitochondrial 1C flux, we compared the expression levels of 1C genes across the studied cell lines. Analyzing transcriptomics data from the Cell Line Encyclopedia Collection (CCLE), we found that the expression level of SLC19A1, which encodes the reduced folate carrier (RFC), is significantly lower in cell lines that rely on the cytosolic 1C pathway (*Figure 17a-b*). Consistently, we found that SLC19A1 expression (*Extended Figure 3a*) levels are significantly

lower in these cells. RFC is an anion antiporter that utilizes the phosphate gradient to achieve uphill folate transport into cells and is ubiquitously expressed across tissues[129]. Considering that RFC has a *Km* of 200-400 μM that far exceeds the serum folate level (150-450 nM)[3,4], variability in RFC level between cancer cells affects the rate of folate intake. Note that, the expression levels of other folate transporters such as proton-coupled folate transporter (PCFT) and folate receptors (Frα, Frβ, and Frγ) were not significantly correlated with the cytosolic 1C flux (*Figure 17a*). To test whether RFC level modulates the balance between cytosolic and mitochondrial 1C flux, RFC overexpressing CEM-7A cells[130] and their parental CEM cells were cultured with [2,3,3-^{2}H]-serine and dTTP labeling was monitored (*Figure 17d*). We found that RFC overexpression induced the mitochondrial 1C flux under physiological folate levels, as dTTP M+2/M+1 ratio remained low in CEM-7A cells regardless of the media folate levels (*Figure 17d*). Note that this was associated with an increase in the intracellular levels of reduced folates to support biosynthesis, as evident by a significant drop in AICAR and GAR levels in the RFC overexpressing cells (*Extended Data Figure 3b*). Analogously, knockdown of RFC in HCT116 cells, which exclusively utilize the mitochondrial 1C cycle in both high and physiological folates (*Figure 15a and 17e*), significantly increased the dTTP M+2/M+1 ratio in physiological folate media (p-value < 0.01; *Figure 17e*). In addition, AICAR and GAR levels in HCT116 cells significantly increased upon RFC knockdown (*Extended Data Figure 3c*). Taken together, our results suggest that RFC expression determines the intracellular folate concentration, which modulates the cytosolic versus mitochondrial 1C metabolism.

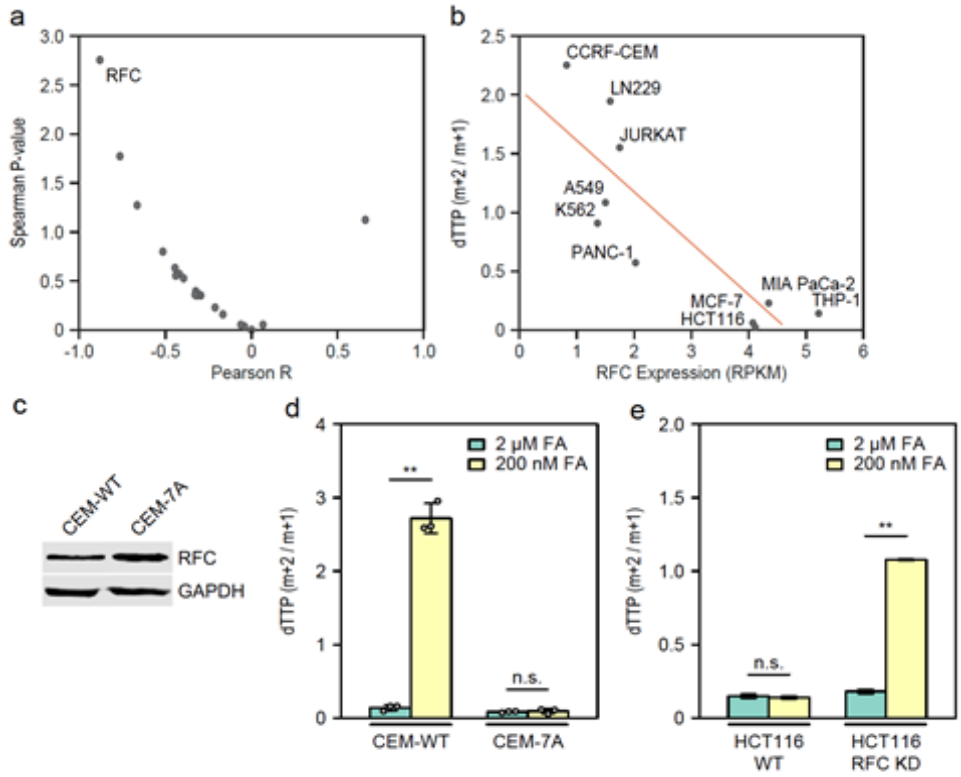


Figure 17: **Cytosolic 1C flux is induced in cancer cells with high expression of the Reduced Folate Carrier (RFC).** (a) The Spearman correlation (x-axis) and p-value (y-axis) between the expression of 1C metabolic genes (based on data from the Cell Line Encyclopedia Collection; CCLE) and the activity of the cytosolic 1C pathway under physiological folate across cell lines (as in Fig. 15b). (b) SLC19A1 expression level from CCLE (x-axis) versus the cytosolic over mitochondrial 1C flux ratio (y-axis) in different cell lines. (c) Western blot of RFC in CCRF-CEM (CEM-WT) and a subline over-expressing RFC (CEM-7A). (d) Cytosolic over mitochondrial 1C flux ratio in CEM-WT and CEM-7A in high and physiological folate media. (e) Cytosolic over mitochondrial 1C flux ratio in HCT116 wt and HCT116 RFC KD in high and physiological folate media. n.s. not significant. **P < 0.01 by two-sample t-test. Data are mean ± SD, n = 3 independent biological replicates.

3.1.4 SHMT1 Acts Synthetic Lethal To RFC, Negatively Affecting Tumor Growth *In Vivo*

Considering the induced contribution of cytosolic 1C flux to pyrimidine biosynthesis in cancer cells with low RFC expression, we sought to test whether SHMT1 knockout in these cells will hinder tumor growth. We found that SHMT1-deficient Jurkat cells (*Figure 18a*), which resort to strictly mitochondrial 1C flux (*Extended Data Figure 4a*), showed a significant ~65% drop in growth rate under physiological folate levels (p-value < 0.001); a significantly larger decrease in growth rate than that observed in high folate conditions (p-value < 0.001; *Figure 18b*). A similar drop in growth due to SHMT1 KO was observed in A549 under physiological folate conditions (*Extended Data Figure 4b*). Feeding SHMT1 KO cells with formate partially rescued the impaired growth, suggesting that mitochondrial folate pathway cannot fully compensate for the loss of cytosolic 1C flux under physiological folate condition (*Figure18b*). The drop in cell proliferation in the SHMT1 KO cells was associated with a significant ~27% increase in the fraction of cells in S phase (p-value < 0.01; *Figure 18c*) and significant increase in the dUMP to dTTP ratio (p-value < 0.001; *Figure 18d*, *Extended Data Figure 4c*), suggesting insufficient thymidylate syntheses (TS) flux to support dTTP production for DNA replication. To evaluate tumor growth *in vivo*, SHMT1 WT and KO Jurkat cells were implanted onto the hind flanks of immunocompromised NOD/SCID mice. As expected, SHMT1 KO tumors grew significantly slower, resulting in a significant 10-fold smaller tumor volume after 7 weeks (*Figure 18e-f*).

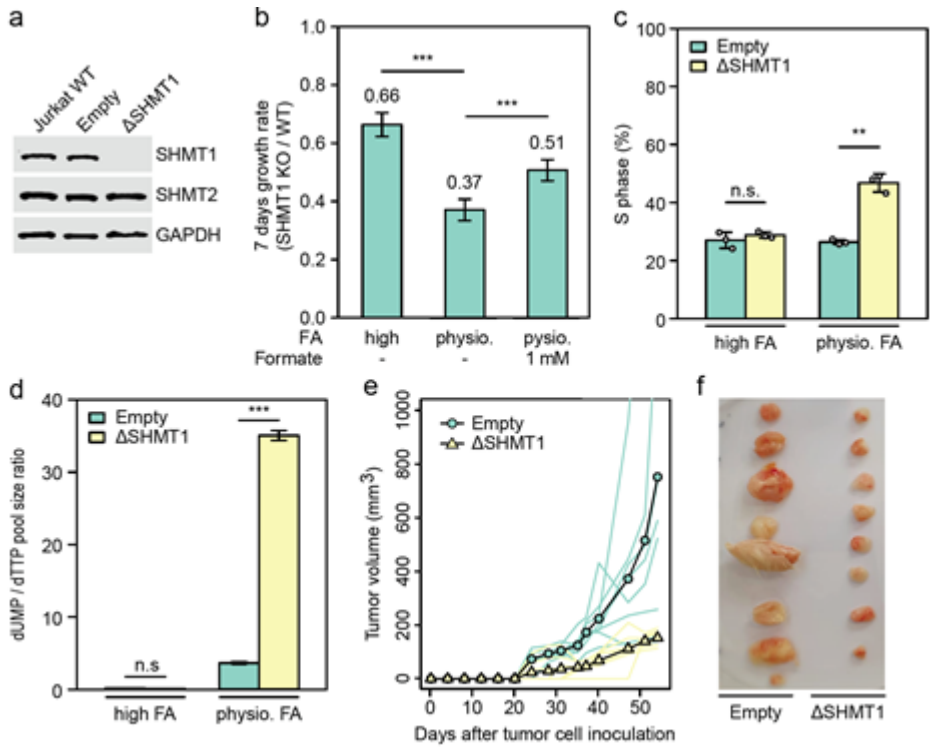


Figure 18: **SHMT1 is essential for growth of tumors with low RFC expression in vitro and in vivo.** (a) Western blot validation of CRISPR-Cas9 knockout of SHMT1. (b) In vitro growth assay for WT and SHMT1 KO Jurkat cells in high folate media, physiological folate media, and physiological folate with formate media (y-axis showing the relative growth rate in KO versus in WT cells). (c) Fraction of WT and SHMT1 KO cells in S phase when grown in high and physiological folate media. (d) Ratio between intracellular dUMP and dTTP in Jurkat WT and SHMT1 KO cells grown in high and physiological folate media. (e) Growth of SHMT1 WT and KO Jurkat xenograft (n = 8). (f) Tumors 8 weeks post injection of SHMT1 WT (left) and KO (right) Jurkat cells. n.s. not significant. **P < 0.01 and ***P < 0.001 by two-sample t-test. Data are mean ± SD, n = 3 independent biological replicates.

3.2 A NEW EFFECTIVE PROCESS FOR VIRUS PRODUCTION AND CELL TRANSDUCTION OF DIFFICULT CELL LINES

The results of this section are part of a manuscript entitled "Process for an efficient lentiviral cell transduction" which is currently under review. It is quoted in this section wherever it seemed useful, including the majority of figures and legends.

In order to perform a CRISPR/Cas9 screening in CCRFCEM and MTA-C3 cell lines, it was very important to establish an efficient transduction method. In fact, T-ALL cells are quite known as difficult to manipulate and the first transduction trials show very low efficiency (not more than 5%). These very low numbers could potentially select a cell subpopulation more prone to be infected, causing therefore an experimental bias. For this reason, we first needed to establish a new effective protocol for virus production and cell transduction. We initially focused on CCRFCEM cell line, and then expanded it to other cell lines. We specifically focused on the optimization of three aspects: (i) identification of the best performing transfection reagent for lentivirus production; (ii) increasing lentivirus concentration prior to transduction without compromising the viral particles and thus their functionality; and (iii) achieving a higher recipient cell concentration without compromising cell viability (*Figure 19*). In the analysis, we used a lentiviral sgRNA library consisting of over 260,000 molecules[125] . By virtue of the mCherry fluorescence marker gene in the sgLenti vector, cells could be analyzed by flow cytometry in order to determine quantitatively the percentage of transduced cells.

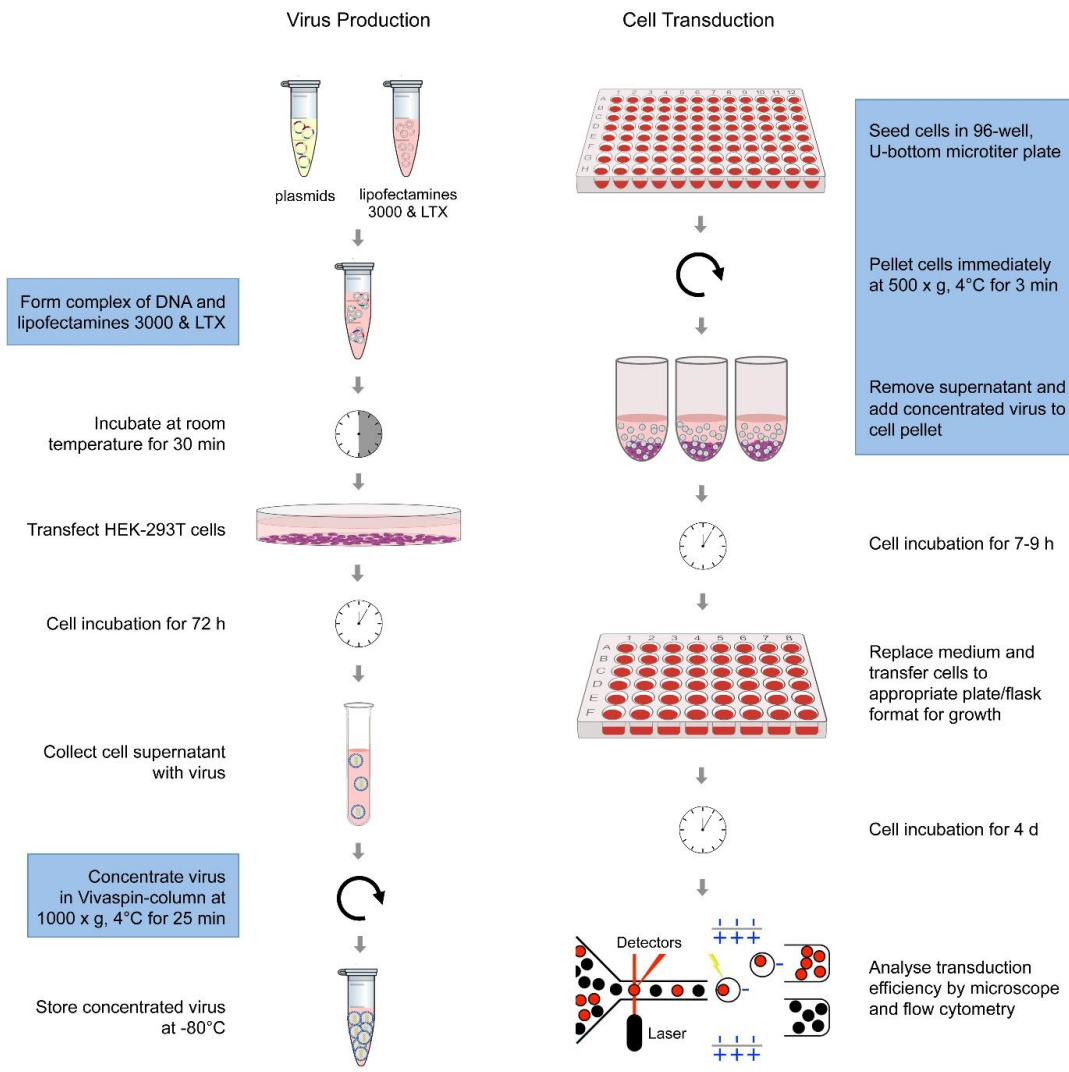


Figure 19: **Schematic overview of the various steps of the optimised transduction process.** The steps in blue-labelled frames are different to standard protocols and jointly increase transduction efficiency substantially.

3.2.1 A Combination Of Lipofectamine 3000 And Lipofectamine LTX Strongly Improves Transduction Efficiency

First, we compared the performance of transfection reagents commonly used for virus production in HEK-293T cells: polyethyenimine (PEI), HEPES-buffered saline (HBS), lipofectamine 3000 and lipofectamine LTX, alone and in combination, as well as jetPRIME and TransIT-LT1. For all transfection reagents, the protocols recommended by their manufacturers were used, unless otherwise stated in the Materials and Methods section. Transduction yields

obtained with virus supernatants from the respective HEK-293T cultures varied widely (*Figure 20*). In cell suspension, best results were obtained with PEI and jetPRIME with transduction efficiencies of about 3% and 7%, respectively. However, mixing lipofectamine 3000 and lipofectamine LTX for a combined use actually led to a substantial rise in transduction efficiency to about 14% (*Figure 20A*).

3.2.2 Cell Concentration During Viral Infection Affects Transduction Efficiency

In order to improve yields further, we expected that a higher concentration of CCRF-CEM cells during transduction might boost the process. Cells were spun down in a 96-well, U-bottom microtiter plate at 500 x g and 4°C for 3 min. The supernatant was removed and the lentivirus then added to the cell pellet. This led to transduction efficiencies in an overall range as without cell pelleting *(Figure 2B*). However, the order of performance changed considerably in comparison to the results with cell suspension. Viruses produced with PEI and jetPRIME exhibited much lower yields, for example. As before, however, viruses produced with a mixture of lipofectamines 3000 and LTX yielded the best results by a wide margin.

3.2.3 Gentle Viral Concentration Using Vivaspin Protein Colums Positively Affects Transduction Efficiency

Since the lipofectamine 3000 and LTX mixture had consistently exhibited much superior performance in comparison to all other transfection reagents, we focused our further efforts on the optimization of yields with viruses produced with this compound mix. We contemplated that concentrating the virus could improve transduction further. In one approach, the virus supernatant produced with HEK-293T cells – as always filtered through pores of 0.45 µm for the removal of cell debris and other contaminants – was centrifuged at 200.000 x g at 4°C for 2 h and the pellet resuspended in 200 µl PBS. Alternatively to ultracentrifugation, the virus solution was loaded onto an ultrafiltration column (Vivaspin, 10.000 MWCO), followed by centrifugation at

1000 x g at 4°C for typically 25 min until the remaining volume was about 1 ml. In all measurements, an equivalent number of viruses was used for transduction. To our surprise, ultracentrifugation decreased transduction efficiency (*Figure 20C, ultracent*), while virus concentration by spin-column improved transduction markedly *(Figure 20C, column*). Differences were such pronounced that they were directly visible looking at relatively few cells microscopically (e.g., *Figure 20D, E)*. Interestingly, adding column-concentrated virus to pelleted cells further enhanced transduction efficiency significantly to almost 30% (*Figure 20C*). For confirmation that the higher virus concentration was responsible for this improvement, irrespective of the reagent used for transfection during virus production, also viruses produced with PEI were concentrated by ultrafiltration column. Again, this resulted in a substantially higher transduction efficiency (*Extended Data Figure 5*), which nevertheless was far lower with 7-8% compared to the nearly 30% obtained with the mixture of lipofectamines 3000 and LTX.

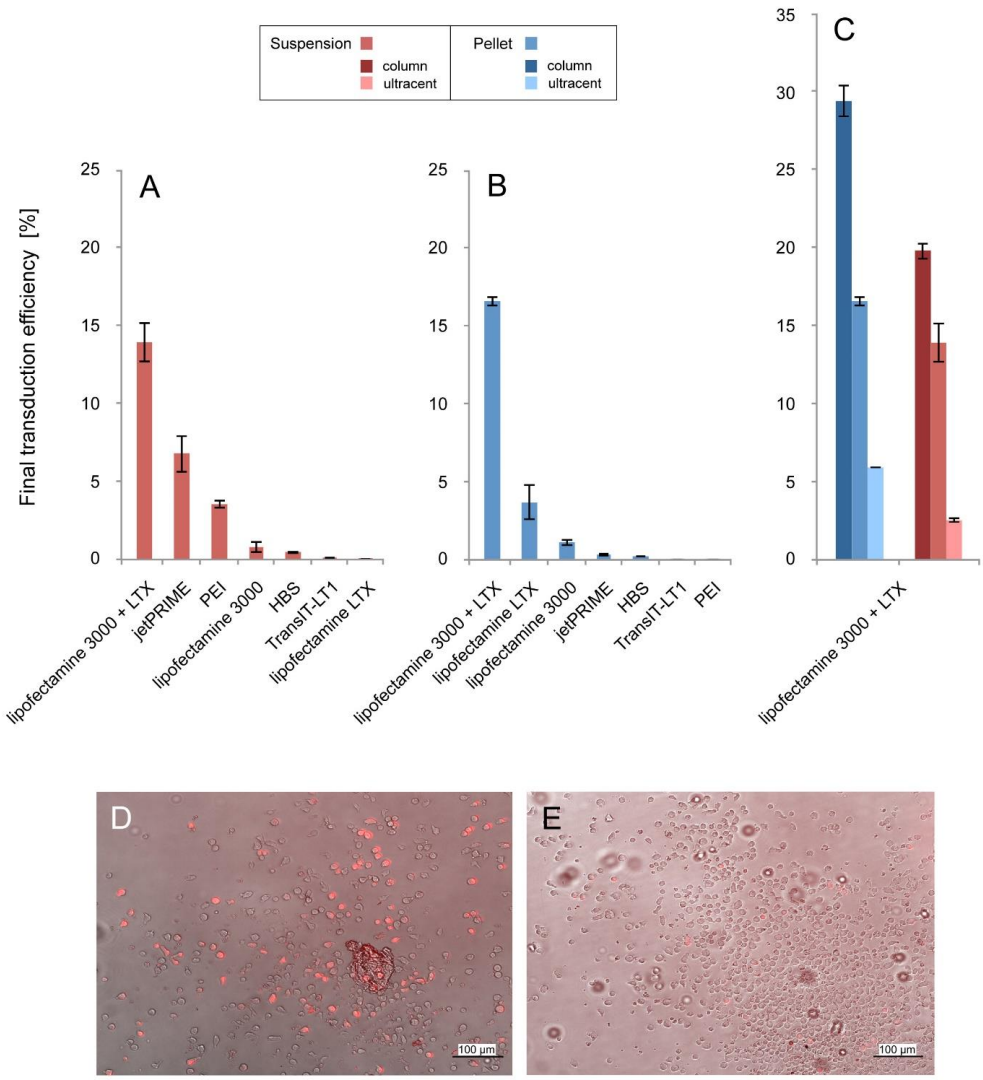


Figure 20: **Comparison of transduction efficiencies with CCRF-CEM cells**. Viruses were produced by transfecting HEK-293T cells with virus constructs using the reagents indicated in panels A-C. The virus-containing cell supernatant was then used for transduction of CCRF-CEM cells. This was done by directly adding the virus to the cell suspension (A; red) or by gently pelleting the cells before adding the virus (B; blue). In panel C, results are shown that were obtained by using a mixture of lipofectamines 3000 and LTX for virus production. Prior to transduction of CCRF-CEM cells, the virus was either used directly as supernatant as in panels A and B (red = cell suspension; blue = pelleted cells) or concentrated via ultracentrifugation (ultracent; light red or blue) or Vivaspin column (column; dark red or blue). Panels D and E: Microscopic images that were acquired four days after transduction; red signals indicate successfully transduced cells. Results are shown of a transduction of CCRF-CEM cells with the optimised protocol (D) and with viruses that had been concentrated by ultracentrifugation rather than spin-column (E).

3.2.4 The Newly Established Protocol Increases Virus Transduction Efficiency In Different Cell Lines

After getting such strikingly improved results with our overall protocol on CCRF-CEM cells (T-ALL, suspension culturing), we studied its effect on five other cancer cell lines: Jurkat (another T-ALL cell line often used in leukaemia studies and grown in suspension culture), BxPC3 (pancreatic ductal adenocarcinoma, primary tumour, adherent culturing), Capan-1 (pancreatic ductal adenocarcinoma, liver metastasis, adherent culturing), Suit-2 (pancreatic ductal adenocarcinoma, liver metastasis, adherent culturing) and NCI-H1299 (lung cancer, lymph node metastasis, adherent culturing). The lipofectamine 3000 and LTX transfection for virus production was combined with concentrating the resulting lentivirus supernatant by means of an ultrafiltration spin-column as well as gentle pelleting of the recipient cells prior to transduction; yields were compared to results obtained with the most commonly applied protocol, using PEI[131] . The new protocol achieved almost 100% transduction efficiency with NCI-H1299 and around 90% with BxPC3 and Suit-2, up from about 50% and 30% with PEI, respectively (*Figure 21*). Transduction of Capan-1, known to be a cell line that is difficult to transduce, most likely due to its high level of mucin expression[132,133], yielded 28% transduction efficiency with our method while 17% were obtained in the control experiment. For Jurkat cells, the difference was about five-fold with 50% versus 9%.

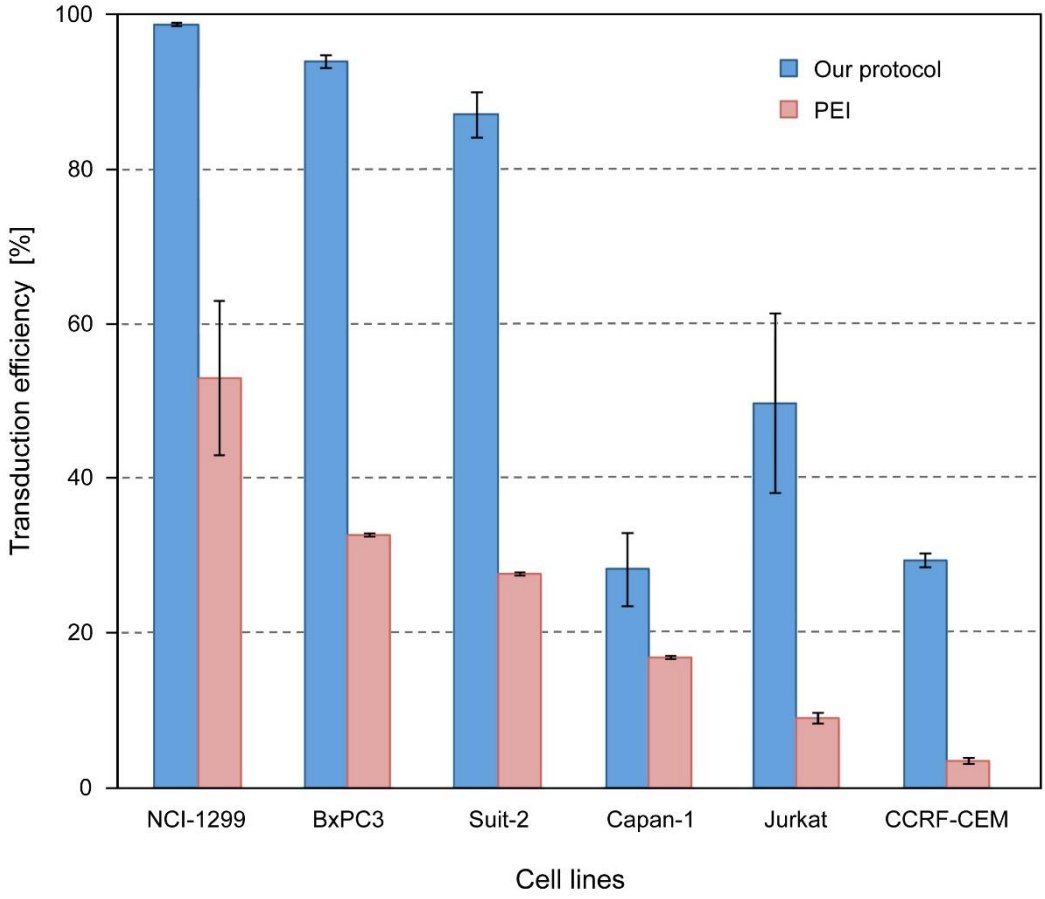


Figure 21: **Comparison of transduction efficiencies in different cell lines.** Blue columns represent the results obtained with the method of mixing lipofectamine 3000 and lipofectamine LTX to transfect HEK-293T cells, concentrating the resulting virus supernatant through Vivaspin 10.000 MW columns, and gently pelleting the recipient cells prior to cell transduction. In comparison, red columns stand for results obtained with a commonly used PEI protocol without concentrating viruses and cells.

3.3 UNRAVELING FPGS-DEPENDENT ANTIFOLATE RESISTANCE: PROTEOMIC AND SYNTHETIC LETHALITY CRISPR SCREENING

Once established an efficient method for cell transduction and unraveled the mechanisms underlying the usage of folate pathway by CCRFCEM cells, we moved on to investigating antifolate resistance. The cell lines models used to assess this question were CCRFCEM (wild type T-ALL leukemia), MTA-C3 and MTA-R15 (antifolate resistant cells) obtained by treating CCRFCEM cells with different concentration of Pemetrexed [50]. Our models of antifolate resistance cell lines were previously characterized for FPGS profiling. Both of them showed afunctional FPGS (loss of activity of over 90%). However, MTA-C3 seemed to be a more defined model compared to MTA-R15, in fact while MTA-C3 presents a specific point mutation in FPGS gene (G1088A) possibly responsible for its inactivation, MTA-R15 shows a wide corollarium of alterations. For this reason, we decided to mainly focus on CCRF-CEM versus MTA-C3. In order to obtain a better picture of the pathways altered in resistant cells, we performed proteomic screenings and antibody microarray analysis. We then merged the obtained knowledge with the candidates resulting from a CRISPR/Cas9 screening, with the aim of finding synthetic lethal genes to FPGS, which results afunctional in MTA-C3 and MTA-R15 cell lines [50,51].

3.3.1 Proteomic Screenings

We performed SILAC, LFQ and antibody array analysis to investigate the protein profile of wt CCRFCEM vs the two resistant cell lines. The data were analyzed using Ingenuity Pathway Analysis (IPA), GenClip, DAVID and GOrilla for Gene Enrichment analysis, focusing on CCRFCEM and MTA-C3 cell lines.

3.3.1.1 SILAC

CCRF-CEM, MTA-C3 and MTA-R15 cell lines were grown in either heavy or light labeled media, provided by SILANTES. An initial incorporation check was performed according to the book "protein and peptide analysis by LC-MS. Experimental strategies" by Letzel T. 2011, to establish optimal growing conditions for the cells. SILAC analysis were performed comparing the two resistant cell lines with the wt CCRFCEM (MTAC3/CCRFCEM and MTAR15/CCRFCEM) to find proteins involved in mechanisms of resistance, and comparing the two resistant cell lines (MTAR15/MTAC3) to find proteins explaining possible differences between the cell lines. PCA analysis show different clustering of the samples.

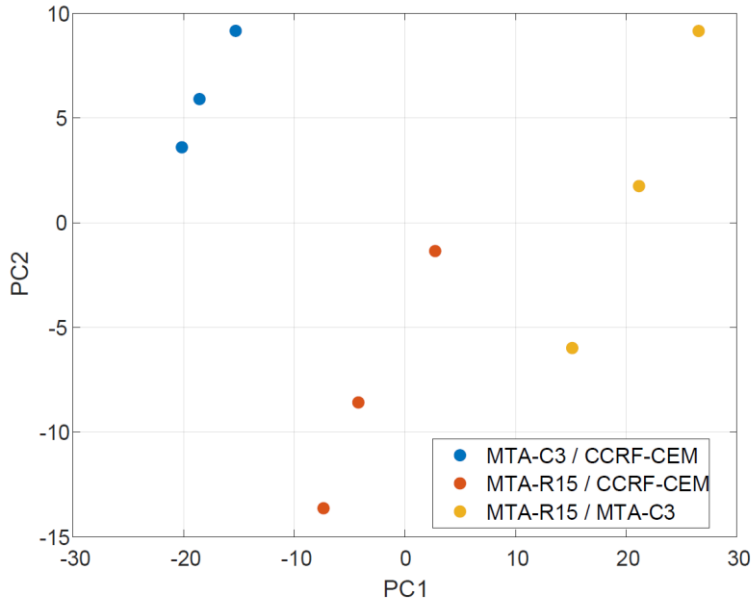


Figure 22: PCA analysis of SILAC results show different clustering for the specimen analysed.

3.3.1.2 LFQ

LFQ analysis were performed for CCRFCEM, MTA-C3 and MTA-R15. The principal component analysis shows separate clustering for the different cell lines.

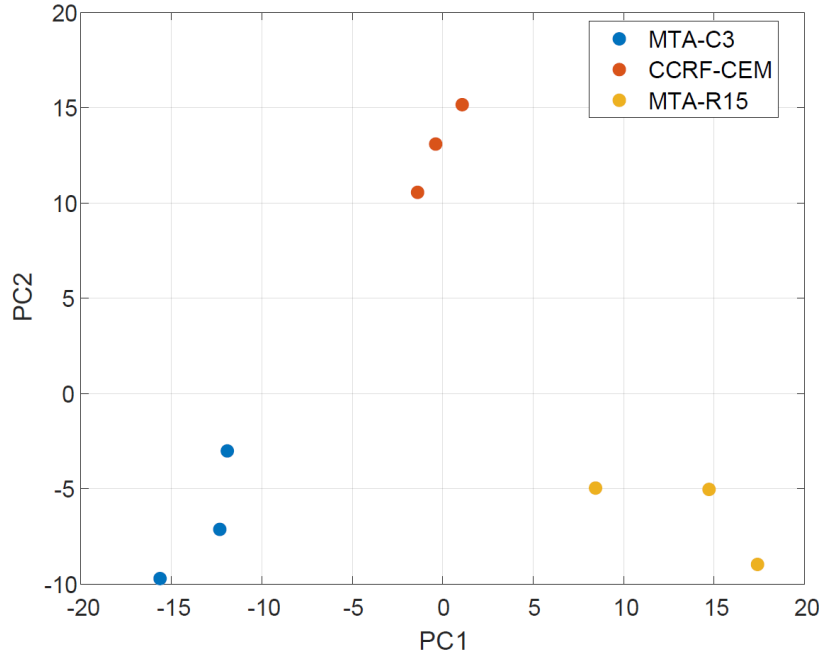


Figure 23: PCA analysis of LFQ results. The three cell lines analyzed show different clustering.

3.3.1.3 Antibody Microarray

Antibody microarrays were performed for CCRFCEM, MTA-C3 and MTA-R15 cell lines. The cells were lysed and labeled with Cy5. A reference sample constituted of a mixture of pancreatic cancer cell line lysates, was labeled with Cy3 and used as normalization. Each microarray slide had 2000 antibodies cancer-protein related, spotted in multiple copy to reduce bias and variability. Figure 8 shows an example of the antibody microarray slides. IPA analysis of MTA-C3 vs CCRF-CEM profile, show an increase of pathways involved in cell growth, survival, movement (Figure 9). This is in accordance with the malignancy and resistance of MTA-C3 cell line.

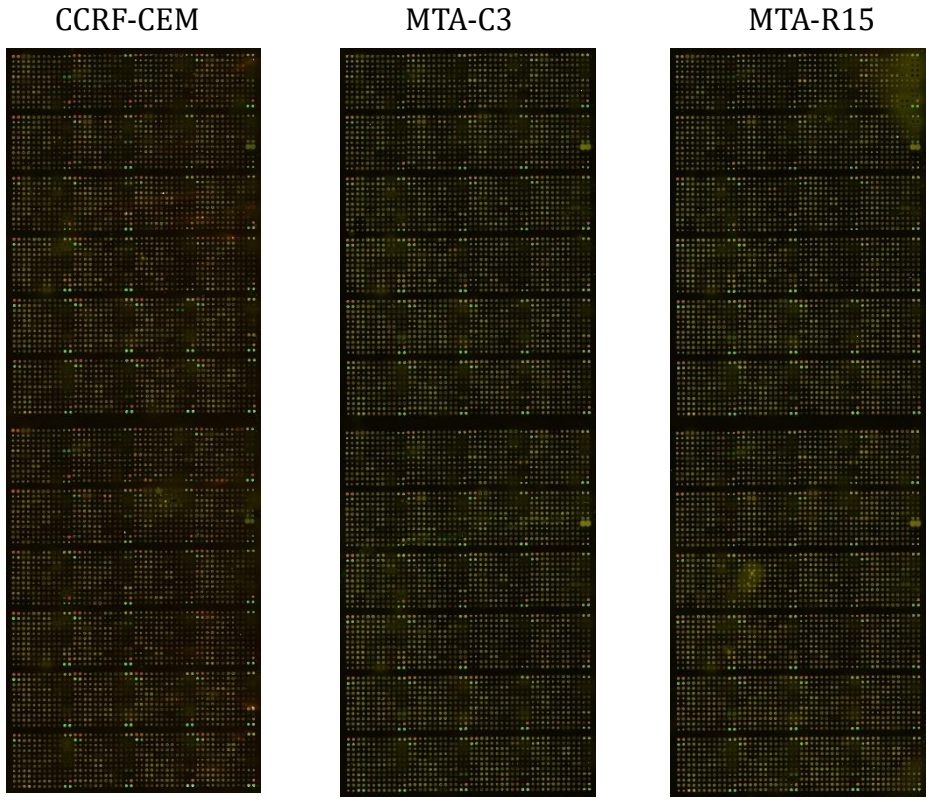


Figure 24: An example of antibody microarrays. The slides were analyzed with genpix software to estimate protein up- and downregulation.

Categories	Diseases or Functions Annotat...	p-value	Predicted Activation State	Activation z-score
Cell Death and Survival	Necrosis	1.88E-128		1.041
Cell Death and Survival	Apoptosis	1.29E-120		-1.045
Cell Death and Survival	Cell survival	1.01E-109	Increased	7.951
Cell Death and Survival	Cell viability	6.93E-109	Increased	7.521
Cell Death and Survival	Cell death of tumor cell lines	9.62E-109		0.296
Cellular Movement	Migration of cells	3.79E-103	Increased	6.111
Cellular Movement	Cell movement	1.91E-101	Increased	6.219
Cell Death and Survival	Apoptosis of tumor cell lines	4.47E-101		0.395
Cellular Development,Cellular Gro...	Cell proliferation of tumor cell lines	9.02E-99	Increased	3.327
Cell-To-Cell Signaling and Interact...	Activation of cells	3.04E-95	Increased	5.884
Cellular Movement,Immune Cell T...	Leukocyte migration	4.08E-95	Increased	4.801
Cellular Movement	Cell movement of blood cells	4.58E-95	Increased	4.925
Cancer,Organismal Injury and Abn...	Advanced malignant tumor	2.74E-93		1.734
Tissue Morphology	Quantity of cells	3.86E-92	Increased	3.294
Cellular Development,Cellular Gro...	Proliferation of blood cells	5.87E-92	Increased	3.172

Figure 25: IPA analysis of antibody microarrays results regarding MTA-C3 vs CCRFCEM protein expression. The results show an increase in cell viability, migration, movement and activation. All these functions correlate with tumor aggressiveness and resistance.

3.3.1.4 Gene Enrichment And Pathway Analysis

The three protein screenings were merged together to obtain a more comprehensive analysis of the pathways underlying the differences between wild type CCRFCEM and resistant MTA-C3. Gene Enrichment analysis show how the upregulated genes found in the resistant cell line versus wild type CCRFCEM, have a strong connection with the mitochondria. The same result is confirmed by both DAVID and GenClip (Figure 12, 13). This outcome matches also preliminary metabolic results, that showed a preference in the usage of 1C-mitochondrial pathway for the resistant cell lines versus CCRFCEM. Downregulated genes cluster mainly with DNA replication and repair and cell cycle arrest (Figure 14). These cellular functions result in fact highly altered in cancer contest. GOrilla analysis on genes resulting downregulated in MTA-C3 show mainly alteration regarding T cell pathways , confirming the robustness of the screenings, while upregulated genes again cluster into small molecule metabolic processes and the mitochondria (data not shown), as predicted also from DAVID and GenClip.

Annotation Cluster 1	Enrichment Score: 14.58			Count	P_Value	Benjamini
UP_KEYWORDS	Mitochondrion	RT		44	2.0E-20	4.7E-18
GOTERM_CC_DIRECT	mitochondrion	RT		44	5.2E-16	1.0E-13
UP_KEYWORDS	Transit peptide	RT		27	1.0E-14	7.9E-13
UP_SEQ_FEATURE	transit peptide:Mitochondrion	RT		25	9.9E-14	5.8E-11
GOTERM_CC_DIRECT	mitochondrial matrix	RT		20	1.3E-11	1.2E-9

Figure 26: DAVID gene enrichment analysis associated with upregulated genes in MTA-C3 vs CCRFCEM

Keyword	Hit	Total	P-Value	Q-Value
cluster1 Enrichment score : 9.23				
MITOCHONDRIA	17	520	1.952e-12	1.207e-09
OXIDATIVE	14	481	7.923e-09	1.633e-06
MITOCHONDRIAL MEMBRANE	19	847	1.323e-08	2.454e-06
METABOLIC	17	693	2.532e-08	4.27e-06
CATALYTIC ACTIVITY	22	1201	4.77e-07	7.373e-05
HEAT SHOCK PROTEIN	20	1052	6.47e-07	9.232e-05
PLASMA MEMBRANE	49	3954	1.596e-06	0.0002

Figure 27: GenClip gene cluster associated with upregulated genes in MTA-C3 vs CCRFCEM

Keyword	Hit	Total	P-Value	Q-Value
CELL DEATH	60	4184	4.278e-10	1.3e-07
cluster1 Enrichment score : 7.56				
S PHASE	29	1239	1.733e-12	1.844e-09
DNA REPLICATION	20	973	2.981e-07	2.643e-05
DNA REPAIR	29	1246	2.241e-12	1.59e-09
CELL CYCLE ARREST	25	1218	8.594e-09	1.663e-06
DOUBLE STRAND BREAK	14	540	5.055e-07	3.984e-05
KINASE ACTIVITY	26	1693	2.371e-05	0.0015
PROTEIN PROTEIN INTERACTION	26	1794	9.38e-05	0.005
CELL GROWTH	55	4205	2.294e-07	2.219e-05
ATPASE ACTIVITY	14	535	4.073e-07	3.467e-05

Figure 28: GenClip gene cluster associated with downregulated genes in MTA-C3 vs CCRFCEM

3.3.2 CRISPR Cas9 Screening

The CRISPR screening was conducted on CCRF-CEM vs MTA-C3 cell lines. We wanted to find genes possibly involved in antifolate resistance and acting synthetic lethal to FPGS, causing therefore MTA-C3 cell death, possibly also restoring sensitivity to Pemetrexed. We used a homemade library CRISPR/Cas9 constituted of 260.000 sgRNA (see Material and Methods). We first generated CCRF-CEM – Cas9 and MTA-C3 – Cas9 single clones, tested them for Cas9 activity, and then proceeded with the library screening.

3.3.2.1 Generation Of CCRFCEM - Cas9 And MTA-C3 - Cas9

CCRF-CEM and MTA-C3 cells were infected with lenti Cas9- Blast virus and single clones were selected and tested for Cas9 efficiency. To this purpose, we selected a highly expressed T-cell surface marker, CD44, and designed sgRNA against it. The library vector backbone (sgLenti, see material and methods) subsequently used for the screening, was used to clone and deliver CD44 sgRNA. Infected cells were selected for mCherry expression, coded by the library vector, and Cas9 efficiency was tested by quantifying the remaining CD44 signal on the cell surface. Figure 11 shows an example of Cas9 efficiency test based on CD44 expression.

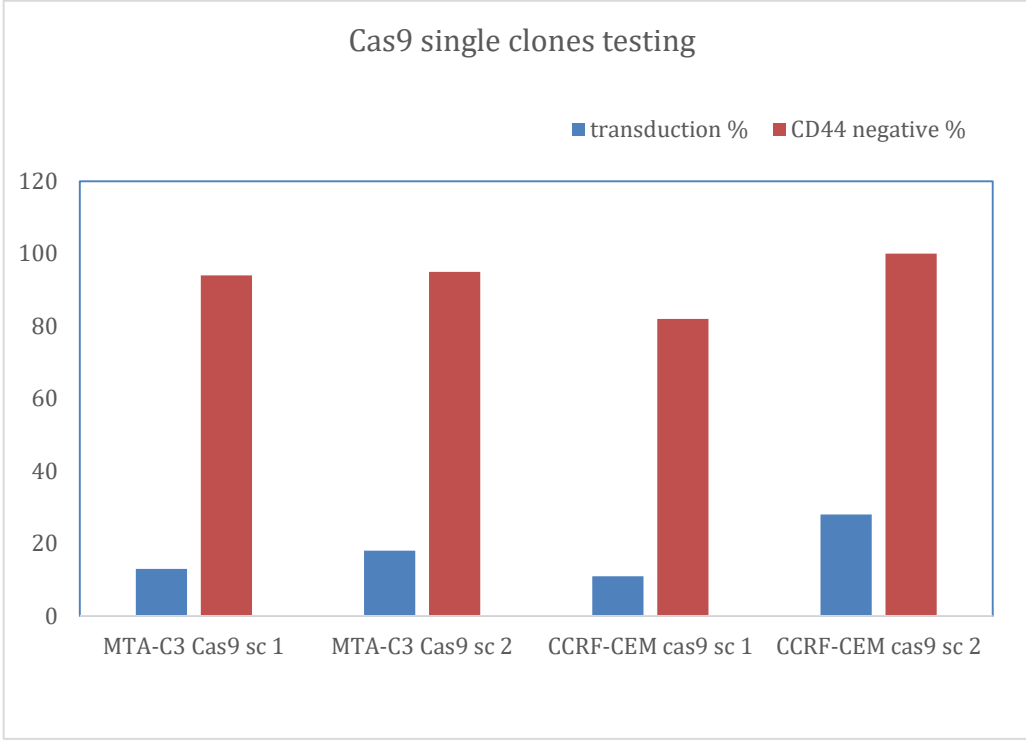


Figure 29: FACS results of CD44 ko in CCRF-CEM Cas9 and MTA-C3 Cas9 single clones.

3.3.2.2 Screening Calculations

The vector used for the screening was sgLenti CRISPR/Cas9 library vector, kindly provided from Michael Boettcher (see Material and Methods). The vector is constituted of 267109 different sgRNAs, targeting XXX genes, having therefore XXX sgRNA for each gene. The screening was conducted on CCRF-CEM Cas9 and MTA-C3 Cas9 cell lines. In order to maintain a better representation of the original cell lines, we decided to pull together five single clones with high Cas9 efficiency (>90%). This would better represent cell line heterogeneity and avoid biased results based on a specific single clone. Cells were transduced with an initial coverage of 250 X and grown for 20 days, corresponding to 14 doublings (cell doubling time: 34 h). Afterwards, dead cells were removed from the media using ficoll dead/live isolation protocol (see Material and Methods), DNA was extracted and library samples were prepared with a final concentration of 10 nM and sent for sequencing with a coverage of 1000 X using Illumina HiSeq 2000 V4 platform, 50 bp single reads.

3.3.2.3 Screening Results

Aim of the screening was to find synthetic lethal candidates to FPGS in MTA-C3 cell line. For this reason, CCRF-CEM candidates were used to normalize MTA-C3 data and specifically select genes that would cause cell death in MTA-C3 but not in CCRF-CEM.

The data were analysed and filtered according to four parameters:

1- sgRNAs which had less than 50 reads were removed from downstream analysis
2- Read count for each sgRNA was normalized based on sequencing depth of each sample. (Read count -> RPM (Reads per million))
3- Genes which had less than 3 sgRNAs were removed from downstream analysis.
4- CS (CRISPR Score) was calculated as: average(log2(final RPM/initial RPM))

The resulting candidates list, as expected, shows a Gaussian distribution of the CS (*Figure 30*). We mainly focused on candidates with CS < -1 (*Figure 31*). GO analysis on top list candidates using GOrilla platform, show clustering on DNA assembly and regulation level (*Figure 32*), including histone and chromatine remodeling, known to be highly affected by methylation processes.

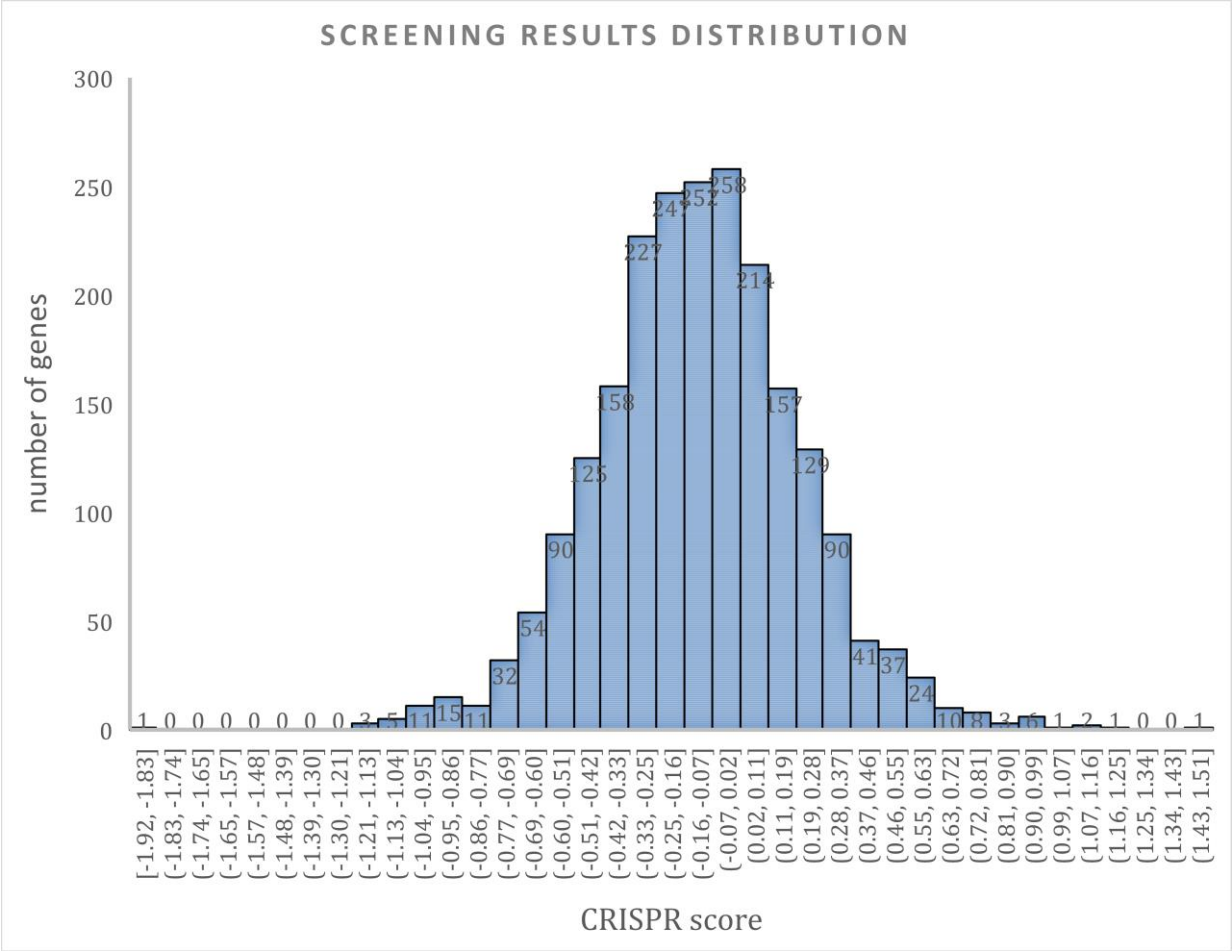


Figure 30: Graphical representation of the resulting screening candidates. High CS depict possible synthetic lethal genes to FPGS in MTA-C3 Cas9 cell line.

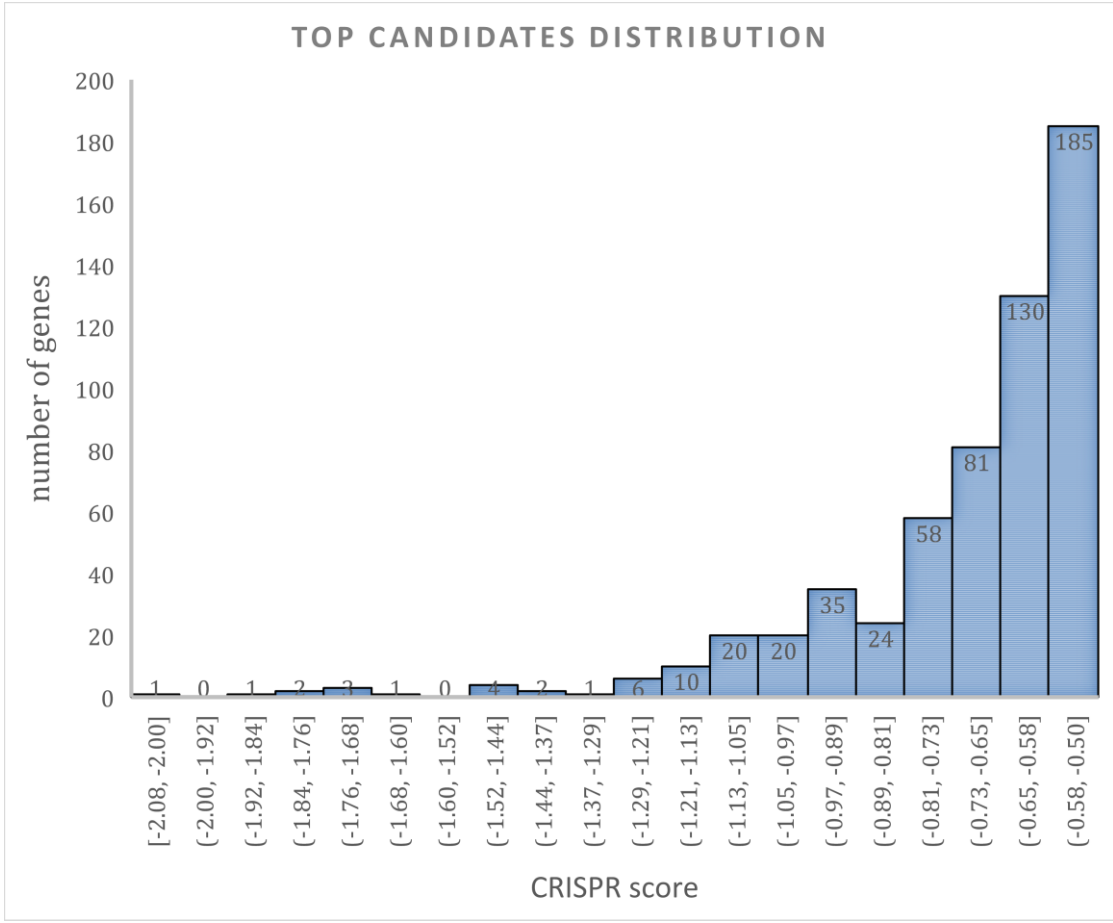


Figure 31: Graphical representation of relevant candidates (< - 0.5 CS).

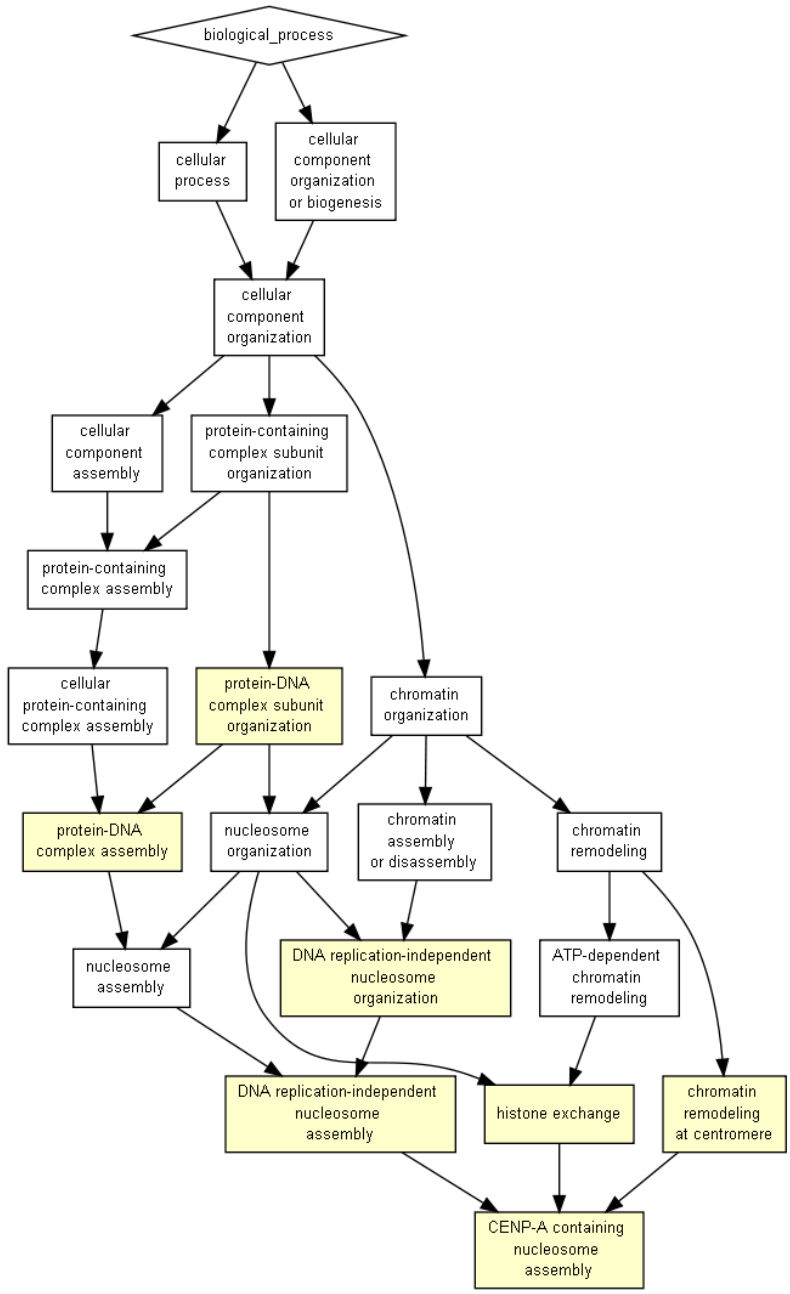


Figure 32: GOrilla analysis of < - 1 CS screening candidates showing biological processes associated with the TOP score genes.

3.3.3 Candidate Selection

Relevant candidates were selected combining two different approaches:

- top list candidates
- candidates involved in one-carbon metabolism

The three protein screenings, SILAC, LFQ, antibody microarray, were merged to obtain a more robust analysis readout. Common candidates were subsequently filtered for proteins that could have a connection with 1-C metabolism and folate pathway in general. Figure 15 shows a representation of top candidates possibly related with folate pathway. Annexin 1 (ANXA1) and dihydrofolate reductase (DHFR) result overexpressed in resistant cell line MTA-C3, in accordance with previous literature [134,135] that proposes a role for them in drug resistance. The candidates that we picked for further investigation were Tyrosine-protein phosphatase non-receptor type 6 (PTPN6) and Inhibitor of apoptosis-promoting 1 (BAX1) (*Figure 33, 34*).

PTPN6 is a member of the protein tyrosine phosphatase (PTP) family. It is therefore part of a class of signaling molecules that are involved in cell growth, differentiation, mitotic cycle, and oncogenic transformation. It is expressed primarily in hematopoietic cells and previous studies already depicted it as a possible tumor suppressor gene in different cancers[136] . It has been shown that its downregulation is mainly due to gene promoter methylation[137–139]. A polymorphism in MTHFR was found to be responsible for hypermethylation of PTPN6 in human sperm under high folic acid supplementation [140], suggesting possible connections between PTPN6 downregulation and folate pathway. In our screenings, PTPN6 resulted downregulated in resistance cells compared to wt CCRF-CEM, confirming its negative correlation with cancer aggressiveness. Despite many studies of PTPN6 in chronic myeloid leukemia (CML), no data are available regarding its function in T-ALL disease. For this reason, we decided to further validate a possible role of PTPN6 in T-ALL antifolate resistance.

BAX1 is a member of the B-cell lymphoma 2 (Bcl2) gene family. This protein functions as an apoptotic activator, and has opposite function compared to BCL2, that is in fact anti-apoptotic . Bcl-2 family is well known in cancer context, in fact altered levels of BCL2 and rearrangements of the gene are often detected in several cancers [141,142] . In particular, BCL2 binds to BAX1 to prevent BAX1 homodimer formation, that would allow membrane translocation with

consequent pro-apoptotic signaling [143,144]. BAX1 has been shown to be involved also in p-53 mediated apoptosis, in particular p-53 can directly activate BAX1 to start the apoptotic program [145] , but it can also upregulate its expression [146–148]. We decided to pick this candidate to investigate whether beside all its implications in different types of cancer, it could also have a direct role in antifolate resistance.

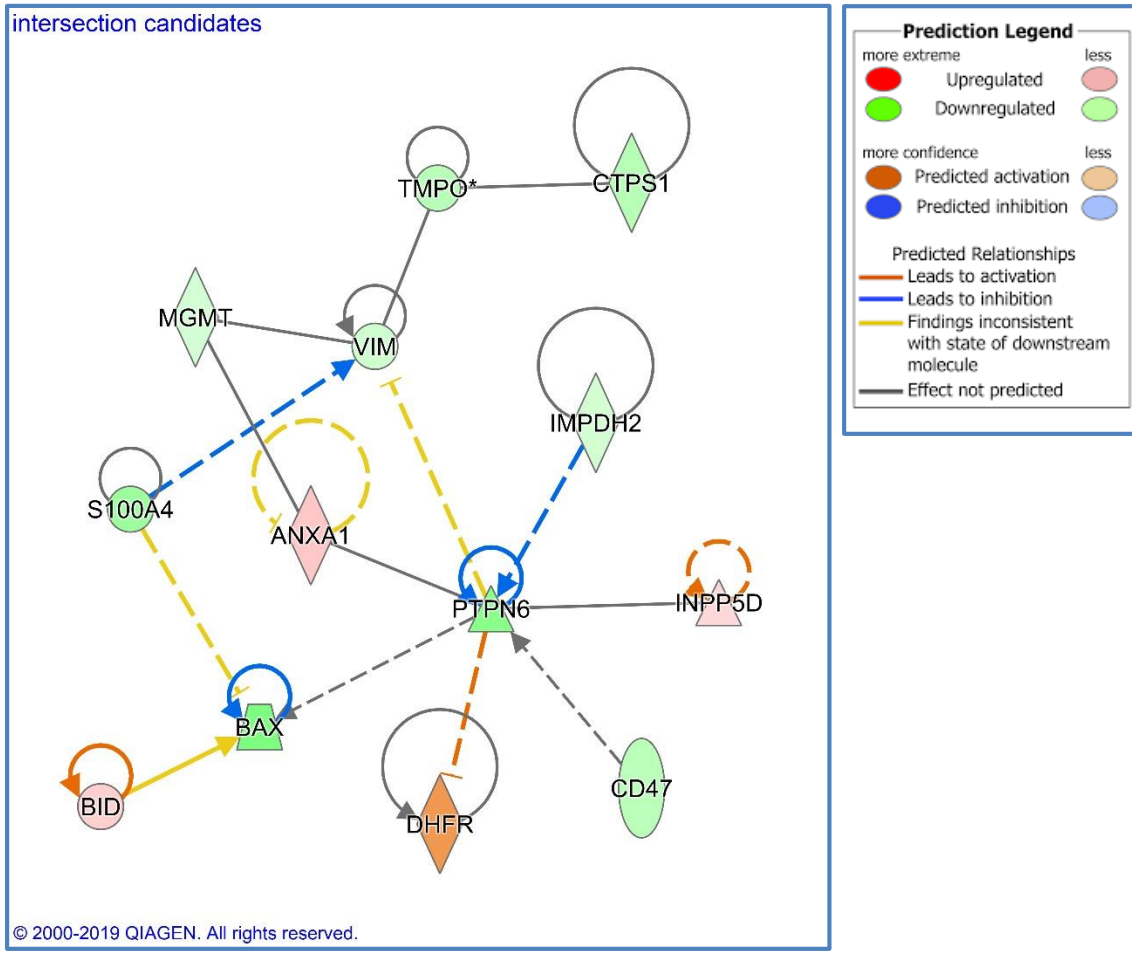


Figure 33: IPA analysis on common candidates resulting from the tree different proteomic screenings. The colors are representative of the ratio between MTA-C3 and CCRF-CEM cell lines. Green and blue colors show downregulation and inhibition of certain proteins and pathways in MTA-C3 compared to CCRF-CEM, while red and orange show upregulation and activation.

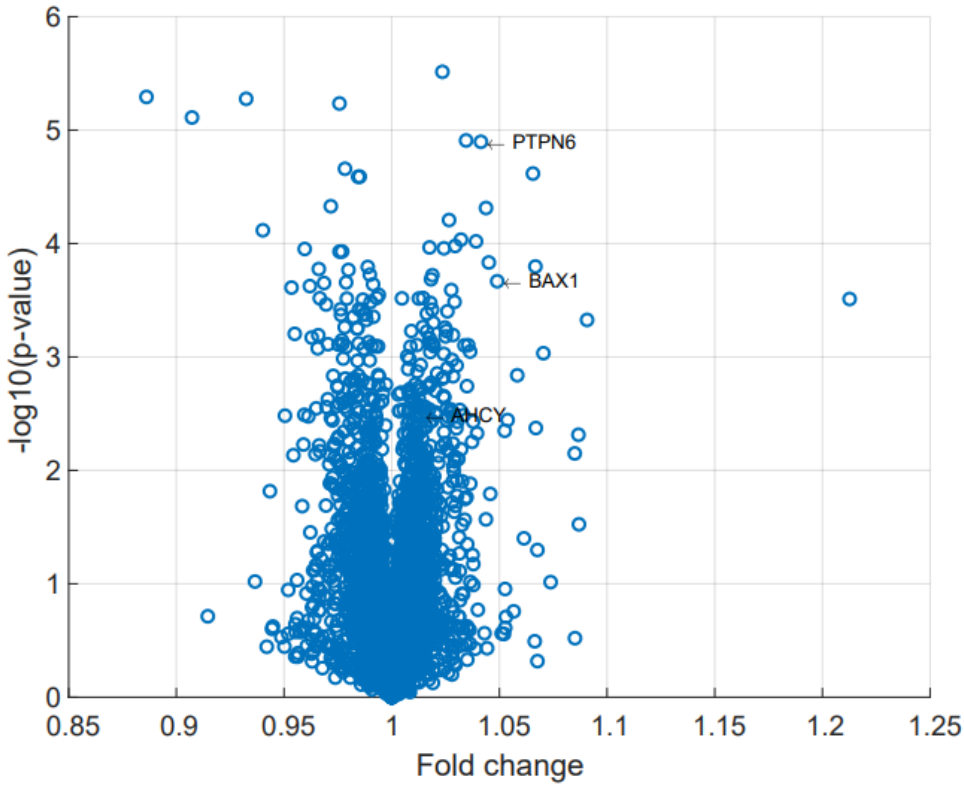


Figure 34: Volcano plot of LFQ analysis show significance (p-value) and relevance (fold change) of our proteomic candidates PTPN6 and BAX1. AHCY resulted a common candidate between the proteomic and the genomic screenings.

The CRISPR/Cas9 screening was also analyzed with the same approach. We selected four candidates: AHCY, AMD1, SAT2 and OIP5. Opa Interacting Protein 5 (OIP5) resulted top candidate of the screening, and previous studies already showed upregulation of this gene in several cancers, also connected with aggressiveness and drug resistance[149,150]. OIP5 protein localizes to centromeres where it is essential for the recruitment of CENP-A (Histone H3-like nucleosomal protein), that is specifically found in centromeric nucleosomes [151]. When H3 is replaced by CENP-A, the nucleosome structure presents protruding DNA and results less compact and ordered. The other three candidates were mainly chosen for their implication with the 1C-metabolism pathway. S-adenosylhomocysteine hydrolase (AHCY) was a candidate in common between genomic and proteomic screenings, and it plays a side role in 1 C metabolism

[152,153]. AHCY is an enzyme that catalyzes the conversion of S-adenosylhomocysteine into adenosine and homocysteine (Hcy), therefore essential in the methionine cycle. In particular, AHCY has a key role in what is called the "methylation index", defined by the ratio of SAM/SAH [154]. Addition of methyl groups (methylation) is important in many cellular processes, including DNA transcription and processing of neurotransmitters [155–157]. AHCY is connected to 1 C metabolism through MTHFR, in fact Hcy is converted to methionine by MTHFR, essential to maintain intracellular SAM concentrations [158] .

Adenosylmethionine Decarboxylase 1 (AMD1) is an enzyme involved in polyamine biosynthesis. As for AHCY, it is also involved in SAM/SAH ratio, in fact AMD1 decarboxylates SAM redirecting it into the biosynthesis of spermine and spermidine [34,159].

Spermidine/Spermine N1 acetyltransferase 2 (SSAT2) shares 45% identity with SSAT1, that is an acetyltransferase enzyme inducible by polyamines and polyamines analogues. Despite the analogy with SSAT1, SSAT2 is not a polyamine catabolic enzyme [160]. However, existing literature seems to not have completely unraveled or understood SSAT2 role, therefore we decided to further validate this possible candidate.

3.3.4 Validation Of AHCY, AMD1, OIP5, SAT2, BAX1 And PTPN6

In order to proceed with the validation, we performed KO of the different candidates in CCRF-CEM Cas9 and MTA-C3 Cas9 cell lines. We then proceeded with growth rate analysis and eventually moved to metabolomics profiling.

3.3.4.1 CRISPR KO Of The Different Candidates

The same library vector backbone (sgLenti) was used to deliver sgRNA to the cells. The CRISPR guides were designed according to sgLenti guides and Zhang database (*Table 4*, Materials and Methods). Viral transduction resulted

successful for AHCY, SAT2, BAX1 and PTPN6 and was validated via FACS by mCherry detection (*Figure 35*).

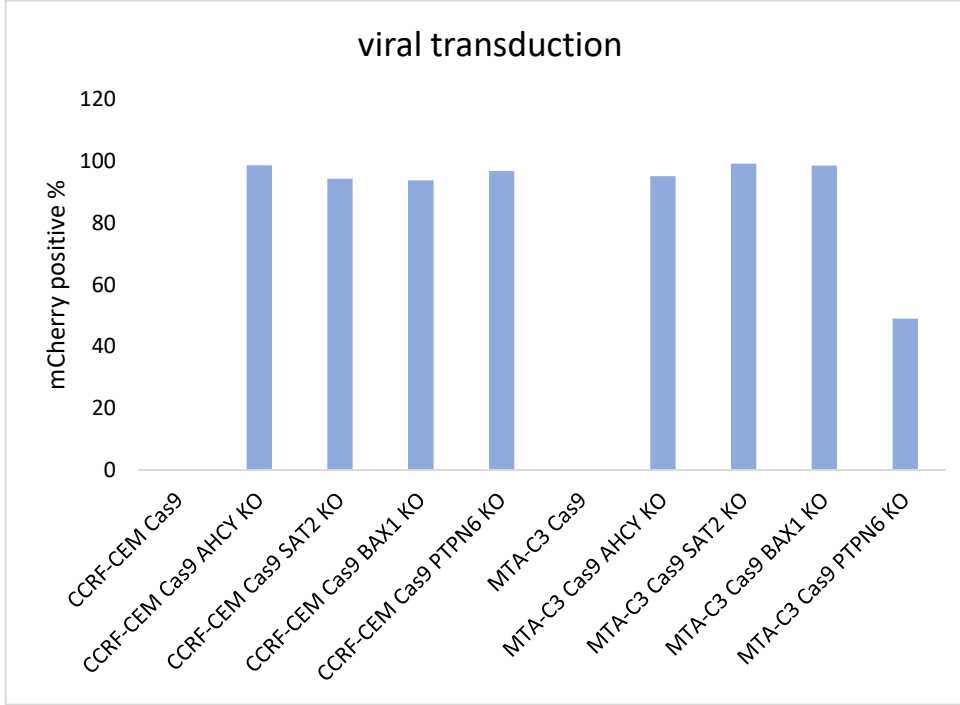


Figure 35: Representation of mCherry expression, and therefore transduction efficiency, in the different KO cell lines.

Successful CRISPR KO was validated via DNA sequencing (data not shown), and we then proceeded with growth assays and metabolomics analysis.

3.3.4.2 KO Cell Lines Show Different Growth Rate Compared To Transduction Control

Once established the KO models, we proceeded with growth rate assays using Culture Counter technology. Surprisingly, BAX1 and PTPN6 KO show a slight decrease in growth rate in both cell lines, though their downregulation was expected to boost cell proliferation. SAT2 KO shows a decrease in proliferation in both wild type and resistant cell line (*Figure 36*); interestingly, AHCY KO shows significant reduced growth rate in MTA-C3 cell line, but not in CCRF-CEM, making of AHCY a putative synthetic lethal

candidate to FPGS (*Figure 37*). We wanted to test whether KO would affect Pemetrexed sensitivity in the resistant cell lines, therefore we exposed the cells to 100 nM Pemetrexed, concentration that resulted lethal for wild type cells (*Figure 38*). None of the knockouts showed an effect on Pemetrexed sensitivity for the resistant cell lines (*Figure 39*).

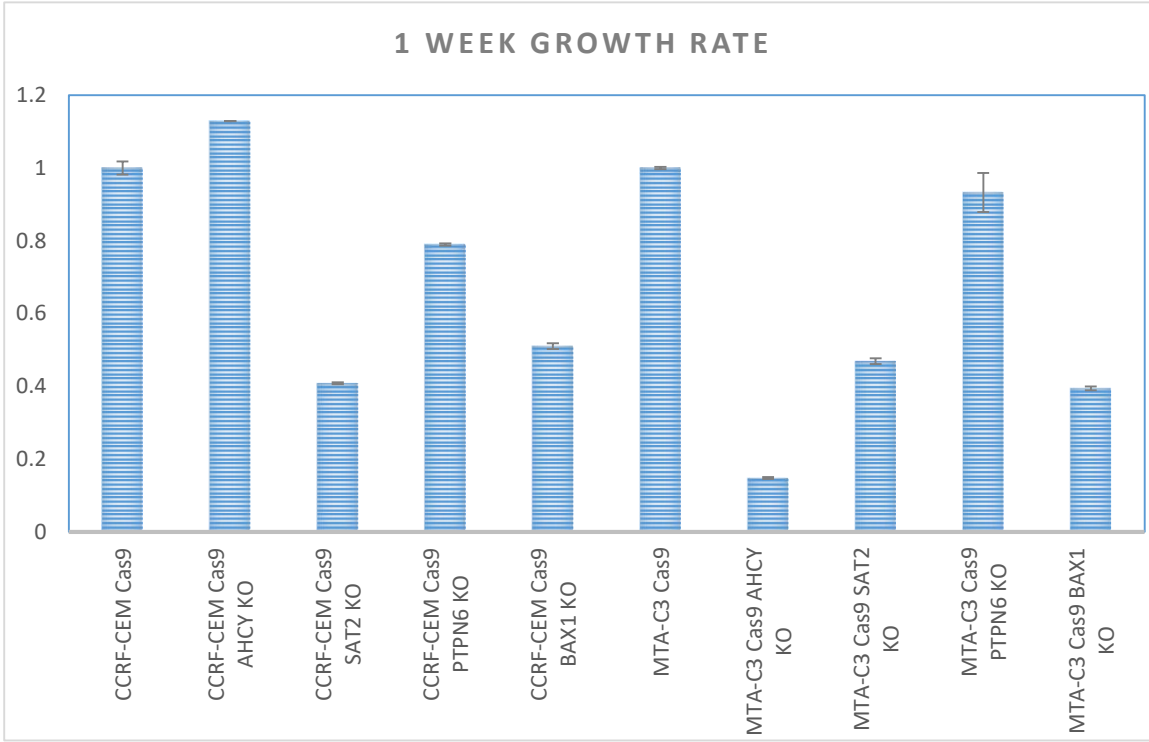


Figure 36: Colture Counter assay showing one week growth rate across the different knockouts and parental cell lines.

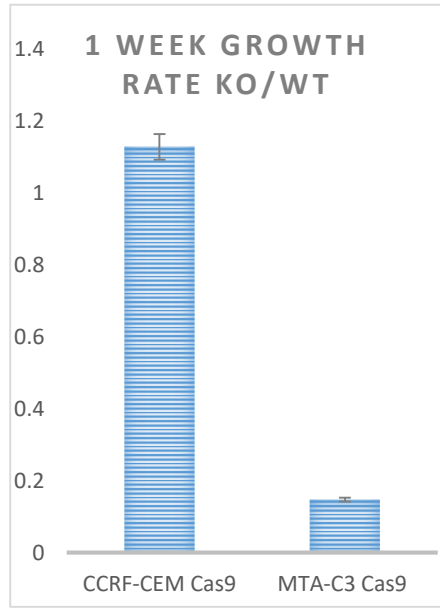


Figure 37: Growth ratio of knockout AHCY cell lines versus parental ones. AHCY KO has a strong effect on proliferation in resistant cells while it is not a sensible target in wild type CCRF-CEM

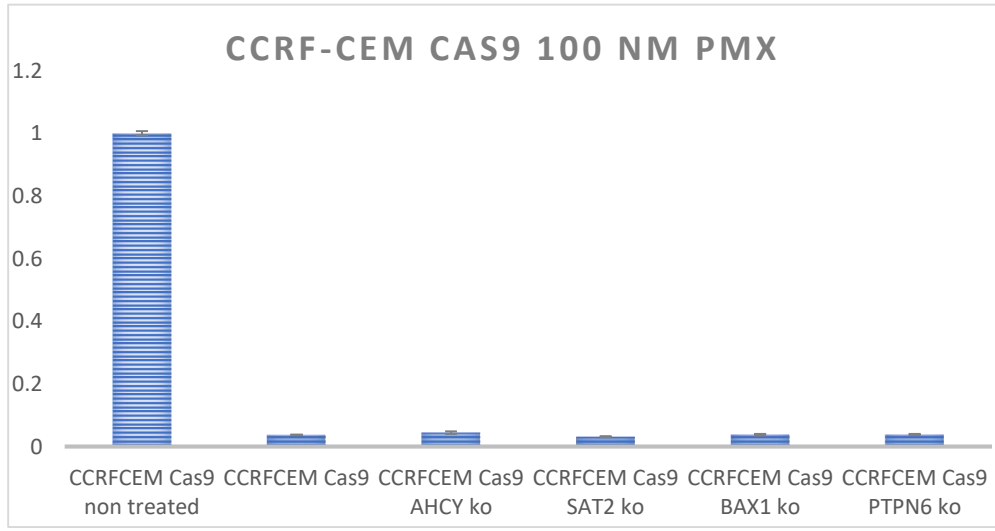


Figure 38: Effect of 100 nM Pemetrexed on one week growth rate in candidates knockout CCRF-CEM cell lines and parental.

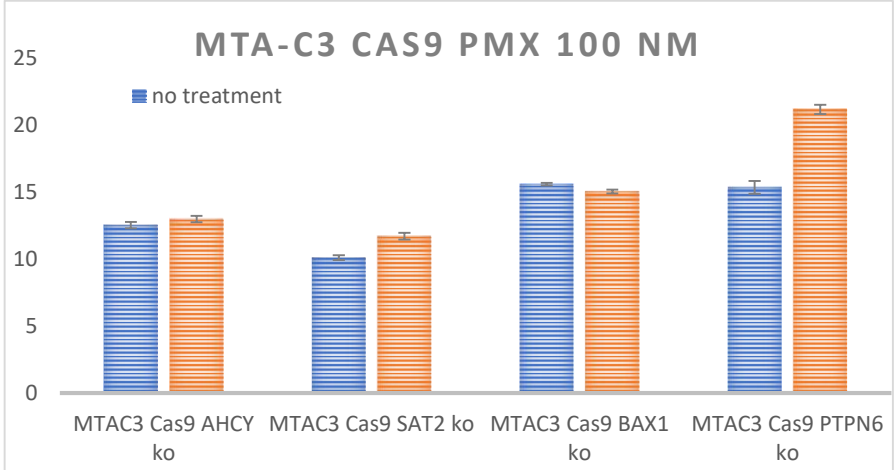


Figure 39: One week growth rate comparison between MTA-C3 knockout cell lines untreated (blue) or supplemented with 100 nM Pemetrexed (orange).

3.3.4.3 AHCY KO cell lines show no difference in mitochondrial versus cytosolic 1C metabolism compared to parental cell lines

The interesting results depicting AHCY as a possible targetable candidate in resistant cells, made us further investigate and characterize AHCY KO cell lines. AHCY is an enzyme catalizing the reaction that from S-adenosyl-homocysteine (SAH) leads to homocysteine (HCY) plus adenosine.

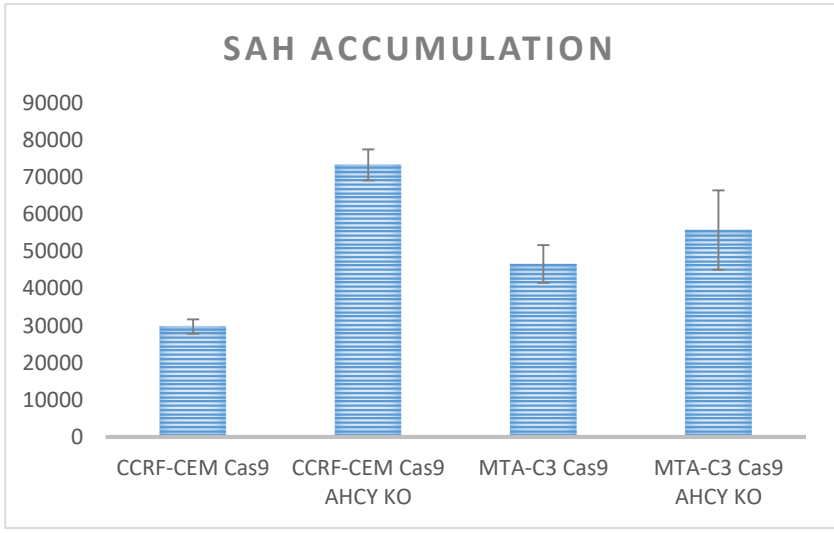


Figure 40: SAH accumulation in AHCY KO cell lines and parental ones

We expected therefore to have SAH accumulation in AHCY KO. Our metabolomics experiments on AHCY KO and parental cell lines, confirm higher levels of SAH in AHCY deficient cells (*Figure 40*).

Previous experiments performed in our lab showed how resistant cell line MTA-C3 presents a minor switch to the cytosolic pathway (M+2) under Low Folate (LF) conditions, when compared to CCRF-CEM.

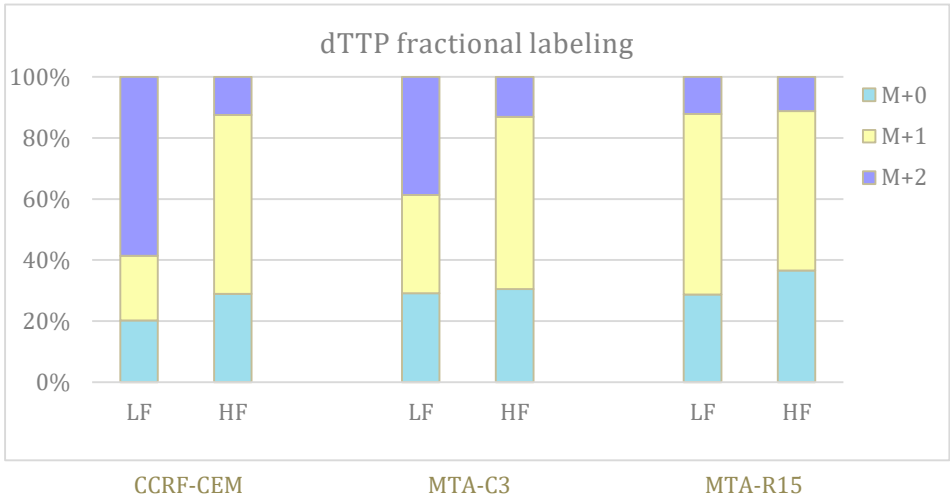


Figure 41: dTTP labeling in wild type and resistant cell lines. M+1 and M+2 are respectively representative of mitochondrial and cytosolic flux.

The resistant clone MTA-R15 does not show a switch to the cytosolic pathway in LF conditions, relying completely on the mitochondrial pathway (*Figure 41*). We then wanted to investigate whether cytosolic vs mitochondrial 1-C flux metabolism would be affected in AHCY KO cell lines (*Figure 42*). Results show that AHCY KO does not have an impact on cytosolic vs mitochondrial 1C flux metabolism.

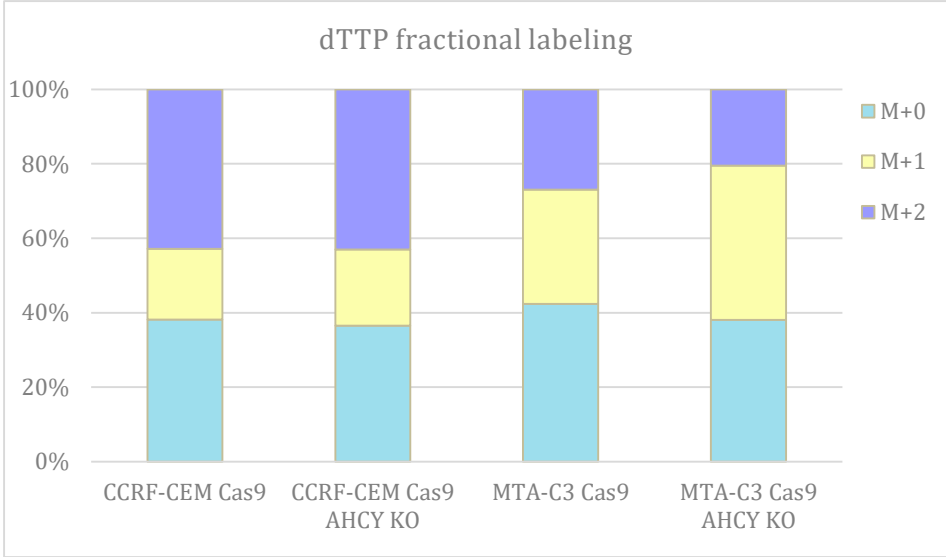


Figure 42: dTTP labeling in AHCY KO and parental cell lines. M+1 and M+2 are respectively representative of mitochondrial and cytosolic flux.

3.3.4.4 CCRF-CEM and MTA-C3 AHCY KO cell lines show different patterns of folate forms distribution

We then thought to investigate folate forms distributions across wild type and resistant AHCY KO cell lines. Levels of pABG and pAcABG confirm that the general amount of folates is not changing significantly between KO and parental cell lines. Interestingly CH-H2FA, CH3-H3FA and CHO-H3FA result accumulated in CCRF-CEM AHCY KO compared to the parental cell line, however the pattern is completely inverted when looking at MTA-C3 AHCY KO versus parental cell line (*Figures 43, 44*).

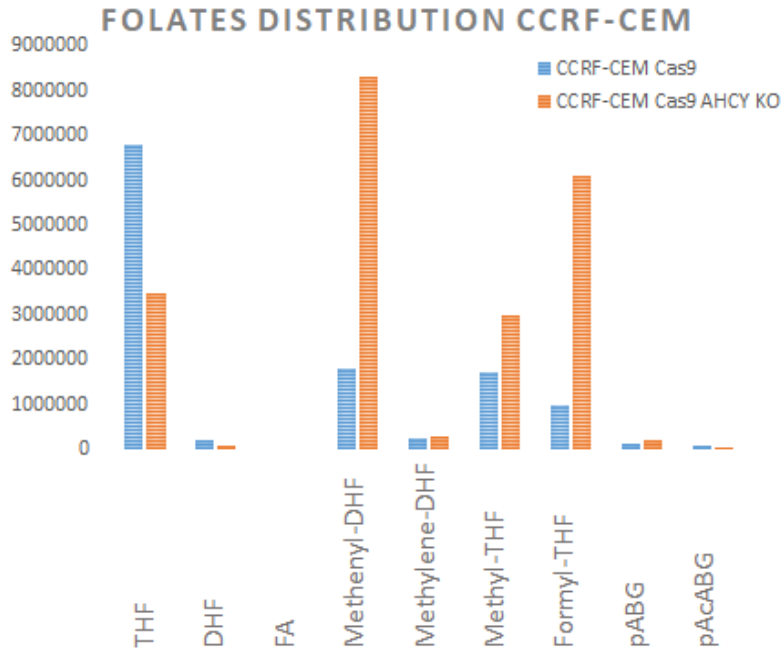


Figure 43: Folates distribution in CCRF-CEM Cas9 wild type and AHCY KO cell lines.

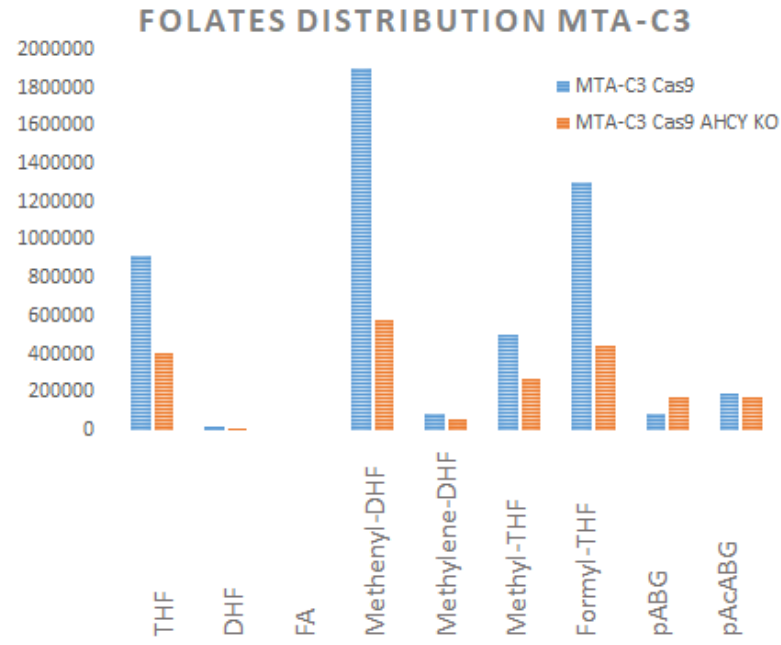


Figure 44: Folates distribution in MTA-C3 Cas9 resistant and AHCY KO cell lines.

We also checked for ability of polyglutamylation (FPGS) activity in the different cell lines. All the different folate forms show lack of polyglutamylation in the resistant cell lines, as expected, confirming an impaired FPGS activity. *Figures 45-48* show different polyglutamylation patterns for CH3-H3FA (Methyl-THF) and H4FA (THF).

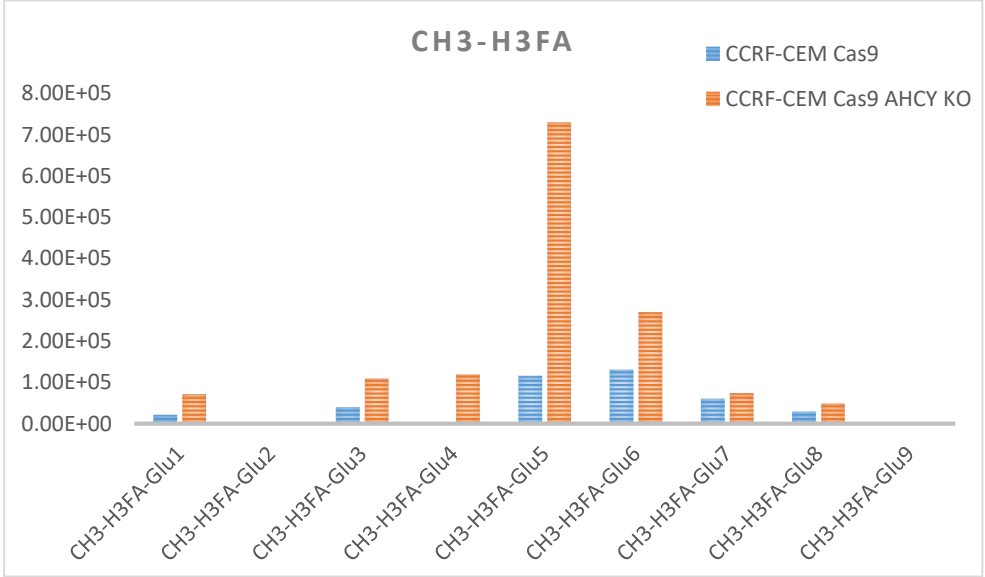


Figure 45: Polyglutamylation levels of methyl THF in AHCY KO and wild type CCRF-CEM Cas9 cell lines.

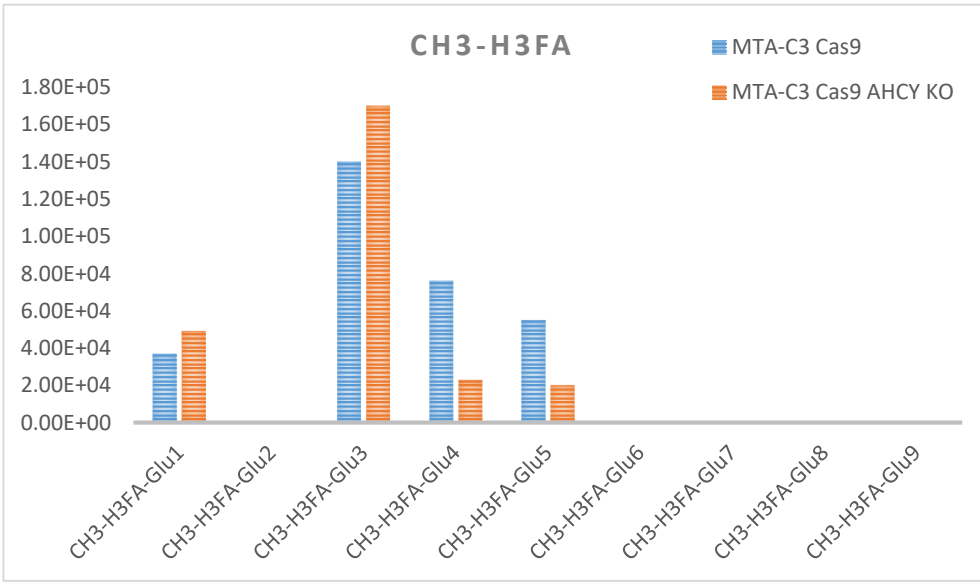


Figure 46: Polyglutamylation levels of methyl THF in AHCY KO and resistant MTA-C3 Cas9 cell lines.

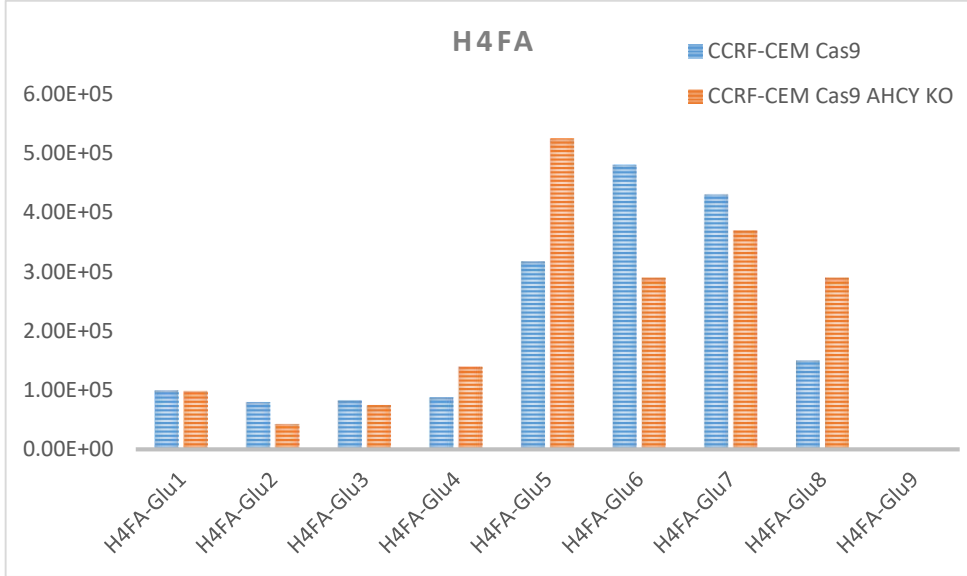


Figure 47: Polyglutamylation levels of tetrahydrofolate (THF) in AHCY KO and wild type CCRF-CEM Cas9 cell lines.

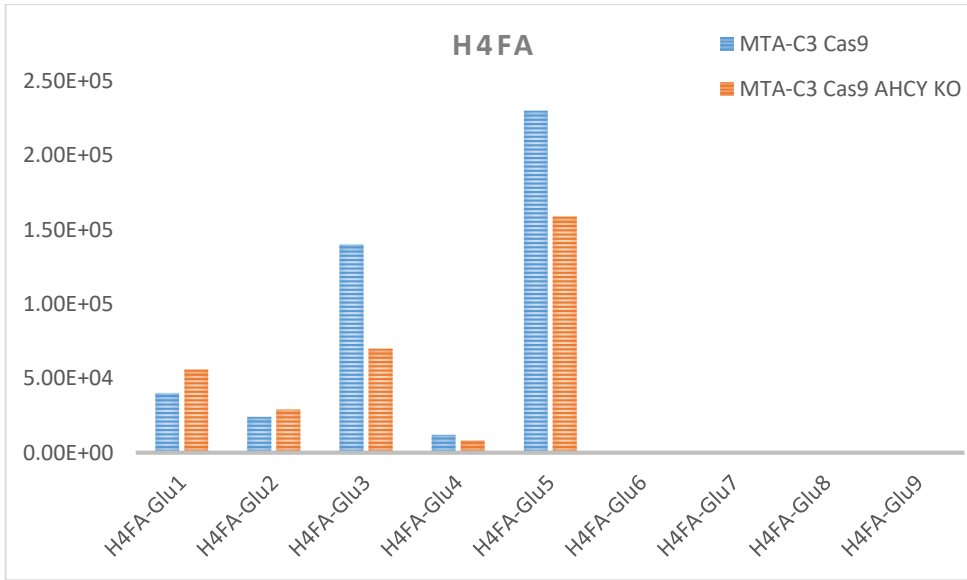
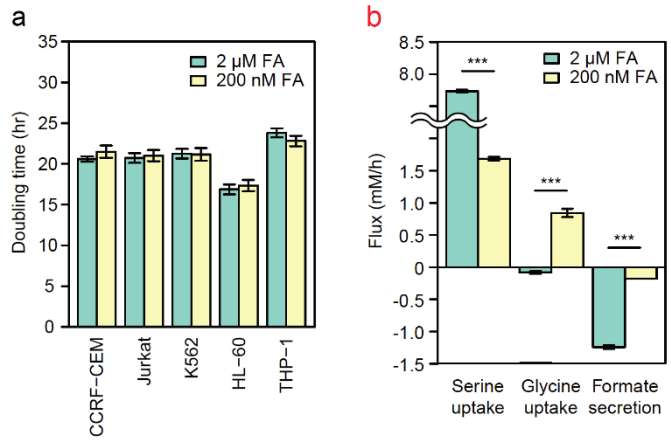
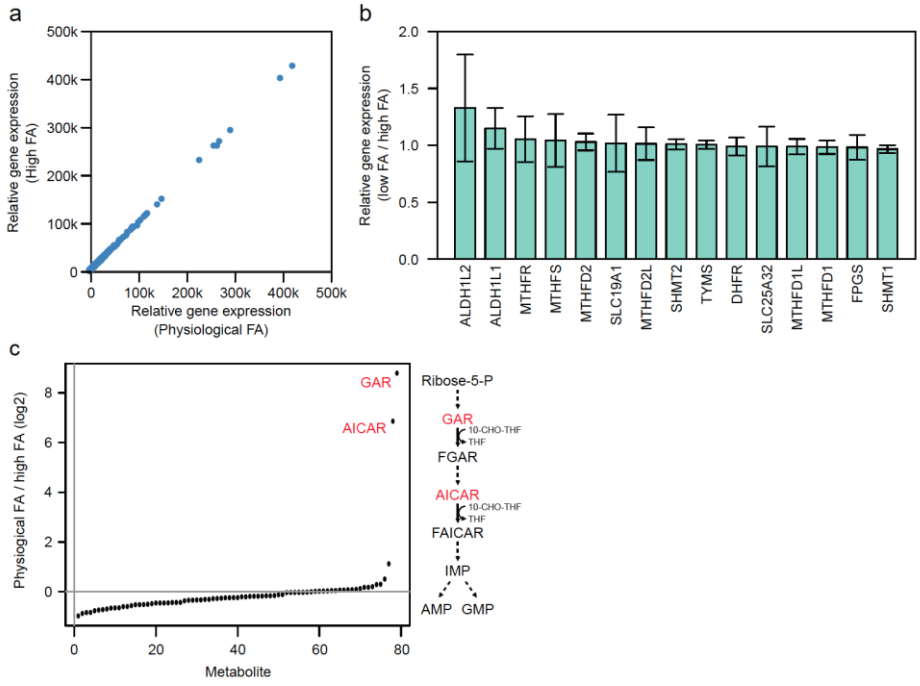


Figure 48: Polyglutamylation levels of tetrahydrofolate (THF) in AHCY KO and wild type CCRF-CEM Cas9 cell lines.

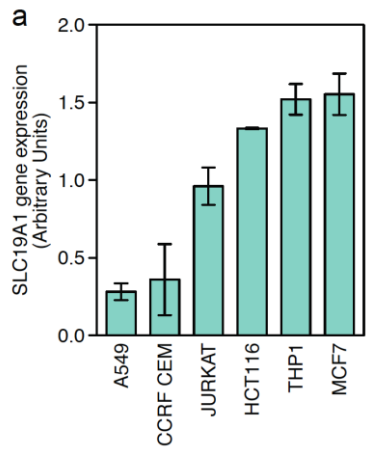
4. SUPPLEMENTARY RESULTS



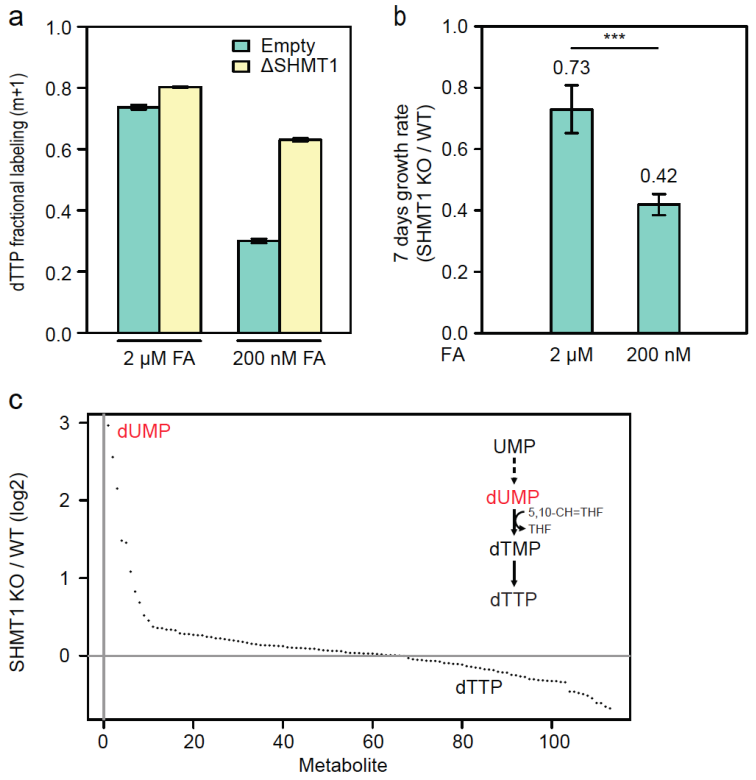
Extended data figure 1: (a) Cell doubling time of various cancer cells in high and physiological folate media. (b) The measured uptake/secretion rate of serine, glycine, and formate in Jurkat in 2 µM and 200 nM media folate levels. ***P < 0.001 by two-sample t-test. Data are mean ± SD, n = 3 independent biological replicates.



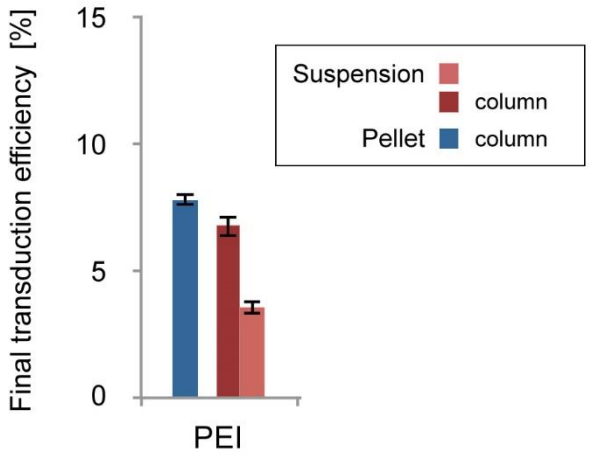
Extended data figure 2: (a) Global gene expression in CCRF-CEM cells under high and physiological folate measured via RNAseq. (b) 1C gene expression in CCRF-CEM cells under high and physiological folate. Data are mean ± SD, n = 3 independent biological replicates. (c) Ratio between intracellular metabolite concentrations in CCRF-CEM cells grown in 200 nM and in 2 µM folate media.



Extended data figure 3: (a) SLC19A1 gene expression in different cell lines. Data are mean ± SD, n = 3 independent biological replicates.



Extended data figure 4: (a) The fractional labeling of intracellular dTTP when feeding Jurkat WT and SHMT1 KO cells with [2,3,3-^{2}H]-serine in 2 μM and 200 nM media folate levels. (b) In vitro growth assay for WT and SHMT1 KO A549 cells in high and physiological folate media (y-axis showing the relative growth rate in KO versus in WT cells). (c) Ratio between intracellular metabolite concentrations in Jurkat WT and SHMT1 KO cells grown in high and physiological folate media. ***P < 0.001 by two-sample t-test. Data are mean ± SD, n = 3 independent biological replicates.



Extended data figure 5: Comparison of transduction efficiencies with CCRF-CEM cells. Viruses were produced by transfecting HEK-293T cells with virus constructs using the PEI protocol. The virus-containing cell supernatant was then used for transduction of CCRF-CEM cells. This was done by directly adding the virus to the cell suspension (red bar) or by adding it after virus concentration by means of ultrafiltration in a Vivaspin column (column; dark red bar). Additionally pelleting the cells gently prior to transduction improved the overall transduction efficiency slightly further (column; blue bar). The same number of viruses was used in all analyses.

5. DISCUSSION

Antifolates have been extensively used in the past six decades, however their efficacy has been limited due to the frequent emergence of drug resistance[37,38]. Antifolates act as folates mimics and therefore enter the cells through regular folates transporters, and get polyglutamated by folypolylglutamase synthase (FPGS)[12,41,47]. Polyglutamylation of folates and antifolates, results in their retention inside the cells, where antifolates can exert their cytotoxic activity. Mechanisms of resistance to antifolates have been shown to involve downregulation of folate transporters, alterations of their function, and/or impairment of FPGS [1,36,46] . Lack in polyglutamylation, would result in extrusion of the antifolates, therefore conferring resistance to the cells. Existing literature associated FPGS downregulation to metothrexate resistance, however it is still not clear whether the enzyme deficiency is a result of the antifolate treatment or rather a pre-existing condition of certain tumors[1]. Nevertheless, it remains clear how FPGS plays a role in antifolate resistance, and this calls for the need of investigating metabolic mechanisms with the aim of finding possible targetable pathways that would restore drug sensitivity or possibly affect proliferation of resistant cells. The folate pathway is at the center of one-carbon metabolism, an essential metabolic pathway that provides one-carbon units, necessary for the production of purines and thymidine, amino acid homeostasis and methylation processes. One-carbon metabolism occurs both in cytosolic and mitochondria and is regulated respectively by SHMT1 and SHMT2[21]. Cytoplasmic SHMT1 generates THF, this 1C donor will contribute to the biosynthesis of methylene and thymidylate. SHMT2 instead generates 1C units under the form of formate, that gets the exported to the cytoplasm where it will contribute to purine biosynthesis. In order to study, quantify, and differentiate between cytosolic and mitochondrial fluxes, mathematical models and different tracers are used.

In this thesis we tried to get a better understanding of the metabolic mechanisms underlying FPGS-antifolate resistance. To achieve this, we first focused on the one-carbon metabolism pathway in relation to folate concentrations, and therefore established proper conditions to conduct an unbiased CRISPR screening comparing wild type CCRF-CEM T-ALL leukemic cell line with MTA-C3, an antifolate resistant cell line presenting afunctional FPGS.

5.1 INTRACELLULAR FOLATES MODULATES THE MITOCHONDRIAL VERSUS CYTOSOLIC ONE-CARBON FLUX : A ROLE FOR RFC

Cell culturing conditions play a very important role, especially when focusing into cells metabolism. In fact, small molecules present in culturing media can easily interfere with metabolomics analysis. Moreover, substances concentrations can potentially have an effect on preferred metabolomic pathways used by the cells[161] . Being antifolates, and therefore folate, the focus of our research, we first directed our attention to the common basic medias (DMEM, RPMI) used in cell culture, and realized that folate concentrations are 10 folds higher compared to physiological ones. We therefore wanted to investigate folate concentrations effects on CCRF-CEM, since this was the cell line model chosen for our research. We showed how metabolic fluxes in cytosolic versus mitochondrial one-carbon metabolism change depending on folate concentration, in particular, CCRF-CEM cell line mainly uses the cytosolic pathway under physiological folate concentrations, while higher concentrations (as in basic cell culture medias) cause a switch to the mitochondrial pathway. This finding was crucial to set appropriate culturing conditions for our model cell line, pointing out the importance of growing our models in physiological folate levels. Next we wondered whether other cell lines also show a difference in the usage of one-carbon metabolism pathway depending on folate concentrations. Surprisingly, T-ALL cell lines CCRF-CEM and Jurkat showed a switch from mitochondrial to cytosolic pathway under physiological folate conditions, but THP-1 and HL-60 AML cell lines didn’t. Other cell lines analyzed also presented different patterns, with HCT116 and MCF-7 standing at one

extreme using mitochondrial pathway independently from the folate concentrations, and LN229 and A549 positioning at the other edge, showing cytosolic switch under low folate conditions. CCRF-CEM and Jurkat cells presented a significant increase of purine biosynthesis intermediates AICAR and GAR under physiological folate conditions, suggesting a potential drop in intracellular folates, which was confirmed from our analysis. Cells relying on mitochondrial 1C flux under physiological conditions, showed instead a higher capacity to retain intracellular folates. We therefore tried to understand whether the switch observed was related to some specific folate transporter, and CCLE analysis showed strong correlation between RFC expression and cytosolic over mitochondrial flux ratio. RFC is a transporter for intake of folates and reduced folates, and it plays a role in folate homeostasis. To investigate RFC and the possible correlation with intracellular folate concentrations, we picked HCT116 as study model since from CCLE analysis this cell line showed high expression of RFC and no cytosolic switch under low folate conditions. We generated RFC-shRNA expressing HCT116 cell lines and could prove that RFC levels are responsible for the intracellular folates concentration. Moreover, HCT116 RFC knockdown cell line, presents a switch to the cytosolic one-carbon pathway under low folate conditions, suggesting a role for RFC in the modulation of cytosolic versus mitochondrial one-carbon metabolism pathway usage.

5.2 SHMT1 ACTS SYNTHETIC LETHAL TO RFC AND SIGNIFICANTLY AFFECTS TUMOR GROWTH

Considering the contribution of cytosolic one-carbon metabolism (SHMT1) in cells with low RFC expression, we wanted to test whether SHMT1 KO would substantially affect growth and proliferation. We first investigated SHMT1 KO in vitro creating Jurkat SHMT1 KO cells. We could prove that under low folate conditions this cell line presents significative growth rate reduction, only partially rescued with the addition of formate, stating that the mitochondrial pathway cannot fully compensate for the inactivation of the cytosolic one. Next

we moved to *in-vivo* studies to check whether SHMT1 KO would affect tumor growth, and we saw a dramatic decrease in tumor formation in mice injected with Jurkat SHMT1 KO compared to Jurkat wild type. These results point out a role for SHMT1 being synthetic lethal to RFC, with important clinical consequences. In fact, one mechanism for antifolate resistance is lowering RFC levels, and in this context targeting SHMT1 could be a new therapeutic strategy.

5.3 A NEW PROCESS FOR VIRUS PRODUCTION AND TRANSDUCTION OF DIFFICULT CELL LINES

Once established the importance of folate concentrations and the right culturing conditions in studies focusing in changes in one-carbon metabolism, we moved to the CRISPR screening. The cell line model used in this study, CCRF-CEM, resulted very difficult to manipulate. Our first attempts of CRISPR library transduction showed very low efficiencies, accounting for less than 5%. Low transduction efficiencies can cause a selection bias, where just a specific subpopulation of the cell line gets infected, not being representative of the heterogeneity of the cell line. This could eventually translate in biased results. We therefore decided to improve the existing virus production and transduction methodologies establishing a new protocol. After several testing, we were able to find out that a mix of lipofectamines LTX and 3000 results in better infective titers. The reason might lay on a better viral packaging ability for the two lipofectamines together, due to higher transfection rates of HEK293T cells. However, this is just a speculation that eventually needs to be further investigated. A second key point of our protocol implies the usage of protein separation columns for virus concentration, instead of the common ultracentrifugation. The usage of column concentration instead of ultracentrifugation, might better preserve viral particles, in fact column concentration requires a much lower speed (1000 g) when compared to ultracentrifugation (200.000 g). High centrifugation speeds can be harmful for viral particles, especially in the case of big viruses, dramatically lowering the functional titer [162]. One last point that helped improving transduction efficiency,

consisted in gently pelleting the cells prior to virus administration. This allowed for the transduction procedure to occur in very low volumes, increasing the concentration of viral particles in the media surrounding the cells. In this situation, the close contact cell-virus resulted in a better transduction efficiency, when compared to the same amount of viral particles in a bigger (more diluted) volume. In conclusion our new method offers new possibilities for studies that require manipulation of difficult cell lines.

5.4 AHCY AS A PUTATIVE SYNTHETIC LETHAL CANDIDATE TO FPGS

Having a clearer picture of the dynamics underlying folate and one-carbon metabolism pathway, and provided the necessary tools needed to manipulate our study model, we proceeded with the CRISPR screening. We compared T-ALL wild type leukemia cell line CCRF-CEM with resistant MTA-C3. The resistant cell line was previously characterized for showing impaired FPGS activity, however this was only shown by feeding cells with metothrexate and checking its polyglutamylation. In our study we were able to directly look at endogenous folates and show the lack of polyglutamylation in the resistant cell line. As previously mentioned, FPGS was found downregulated in antifolate resistant tumors, in particular in leukemic children resulting resistant to methotrexate treatment, therefore with our screening we aimed at finding potential metabolic candidates synthetic lethal to FPGS. To obtain a broader knowledge and a better characterization of the model cell lines, we also performed several proteomic screenings. Gene Enrichment analysis conducted on proteomic data, showed how the upregulated genes found in the resistant cell line versus wild type CCRFCEM, have a strong connection with the mitochondria, in accordance with our metabolomics data that resistant cells tend to have a prediligence for mitochondria pathway over cytosolic. LFQ and SILAC analysis confirmed altered levels of Annexin V and dyhydrofolate reductase in resistant cells[134,135], as expected from literature. Merging LFQ, SILAC and antibody arrays, the latter being a technique established in our lab, we were able to pinpoint two candidates: PTPN6 and BAX1. PTPN6 is a member of protein tyrosine

phosphatase (PTP) family, and it regulates cellular processes such as cell growth, differentiation and oncogenic transformation. It is already well known in leukemic context and its downregulation is associated with aggressiveness and metastatic potential in chronic myeloid leukemia[136,138]. We wondered whether its downregulation in CCRF-CEM cell line would affect growth rate and Pemetrexed sensitivity. Our results show that neither proliferation nor drug sensitivity were altered, suggesting that PTPN6 downregulation alone is not sufficient to conferee an aggressive phenotype, but might instead act in combination with other alterations. Downregulation of BAX1, known to play a role in apoptosis regulation, surprisingly showed a decrease in growth rate in both wild type and resistant cell lines. This could be explained by the intricate regulation between BAX and Bcl2 members, and the accumulation of mutations and alterations characteristic of Bcl2 in cancer context. In general, BAX1 tends to proapoptotic (therefore, oncosuppressor) and Bcl2 results as an oncogenic gene, however specific mutations might completely switch the role of these two members[141,142]. We next analyzed the genomic screening results and tried to combine them with our proteomics and metabolomics knowledge. In particular, we decided to look for candidates potentially involved in one-carbon metabolism, and therefore associated with (anti)folates processing. We therefore decided to further validate SAT2 and AHCY, the latter being also a candidate resulting from the proteomic screenings, though not showing strong significance at the proteomic level. SAT2 function seems still controversial, in fact its analogy with SAT1 made it believed to be involved in polyamine catabolism, however Coleman et al.[160], speculated that it might instead have a function in thialysine acetylation. AHCY has been recently shown to play a role in DNA damage[155], and eventually cell proliferation. Our results show that SAT2 KO had an effect on growth rate in both wild type and resistant cell lines, excluding a possible relationship between SAT2 and FPGS downregulation. However, CCRF-CEM cell line resulted insensitive to AHCY KO while growth rate dropped dramatically in resistant cell line MTA-C3 AHCY KO. This finding made AHCY a possible synthetic lethal candidate, therefore we decided to further characterize parental and AHCY KO cell lines under a metabolomics profile. As

expected, AHCY KO cell lines show accumulation of S-adenosyl-homocysteine (SAH), proving the knockout efficiency. Previous dTTP labeling analysis in CCRF-CEM and MTA-C3 cell lines showed a preference for the resistant cell line to use mitochondrial 1C flux even under physiological folate conditions. AHCY showed no impact on the usage of cytosolic versus mitochondrial 1C flux in parental versus KO cell lines, suggesting no specific changes in folates retention. However, looking at specific intracellular folates distribution, we were able to detect a complete different pattern in wild type CCRF-CEM AHCY KO verus MTA-C3 AHCY KO. In particular, CH-H2FA (methenyl-DHF) and CHO-H3FA (Formyl-THF) result accumulated in CCRF-CEM AHCY KO compared to the parental cell line, but the readout is completely inverted in the resistant cell line. In fact, Methenyl-DHF and Formyl-THF rather decrease in MTA-C3 AHCY KO compared to the parental cell line. Total amount of intracellular folates results stable across the different cell lines, suggesting that the variations observed in the accumulation of the different folate forms, are rather due to some specific internal mechanism. In particular, the decrease in accumulation of Formyl-THF in MTA-C3 AHCY KO, might explain the reduction in proliferative rate, since it directly affects purine production, needed for DNA replication. However, it is unclear how AHCY KO in CCRF-CEM results instead in formyl-THF accumulation. Folates polyglutamylation patterns across AHCY KO and parental cell lines do not show significative differences, therefore a possible interconnection between AHCY and FPGS is still unclear and further investigations need to be done.

5.5 SUMMARY

In summary:

- we showed that folate levels in culture can affect one-carbon metabolism pathway usage. In particular, RFC plays a crucial role in intracellular folates concentrations that modulate cytoplasmic versus mitochondrial one-carbon flux usage.
- Moreover, SHMT1 results synthetic lethal to low RFC levels, in fact tumors derived from SHMT1 KO Jurkat cells are significantly smaller when compared to Jurkat wild type.
- We propose a new method for virus production and transduction, to increase transduction efficiency in difficult cell line models.
- We were able to show polyglutamylation impairment of endogenous folates in antifolate resistant cell line MTA-C3, previously characterized with afucntional FPGS
- We were also able to find a putative synthetic lethal candidate to FPGS, showing that AHCY KO in antifolate FPGS-dependent resistant cell line MTA-C3, causes a dramatic drop in growth rate. However, further investigations need to performed.

AKNOWLEDGMENTS

I would first like to thank my supervisor, Jorg Hoheisel, for accepting me in his group and giving me the chance to work on this project. But most of all, I want to thank him under a personal level, for being not only a scientific supervisor but a very understanding person, supporting me and pushing me during hard times, not letting me backing down. To these regards, I also want to thank my collaborator, Tomer Shlomi, for taking me in his group for the end of my PhD, and for offering me understanding and support during hard times. It is hard to make a complete list of all the people that stood by my side during this journey. Some people stayed till the end, some dropped off, some joined at different times, but all of them had their role and their importance in this phase of my life. PhD is much more than a research project: it is a whole life challenge, a crush into different cultures, a reset and rebuilt of our person both work and private wise. I want to thank all the B070 lab for laughs, pasta-clubs, parties, beers, and fun times we had together! All of you has a special place and memory in my heart. I want to especially thank Beiping for being such a close caring friend, even when I left for another Country (continent!), distance could not break this bound! I want to thank Yenan for our profound conversations and meditations, probably inspired by Patrizio that infected all of us with his yoga chilled mind. A special thank to our secretaries, Anke and Marie, that had such a hassle with all my travels, documents, papers, and almost had to learn Hebrew to communicate with the Israeli institute. Thank you to Gianmatteo, that was my little Friuli in Germany, and kept improving my own language. How to not mention all the reception amazing crew!!!!!! On top of all, my favorite, the gentle giant Kai! I gave him such headaches with all my life stories! And of course Uwe, Sabastian, Pia, Patrik, Giuseppe (Italians are everywhere!). They have been all kind and amazing, at any time of the day or of the night, no matter the weather. It is always nice to start a fresh working day with a smile at the entrance ☺ Thank you to Isa and to Stefy for being there for me during hard times, even if now our paths got divided, I won't forget the warmth they offered me when I

was looking for a repair. Thank you to Paula, for bringing some latino charm in gray cloudy days. Thank you to Chiara, Ko and Venu, that I met almost when I was about to leave, but that filled my heart with so much energy, and happiness, and strength! And how to forget Athena! The sparkle of sun in TP3! With her fresh red lipstick! Spreading happiness everywhere she goes. Now I should move to my Israeli life, and here I want to thank Won for pushing me through the project, and Elina for taking care of me as a sister. Thanks to Praveen for being always present, even at 8am on Shabbath at the bus stop, just to come with me to the hospital. Thanks to Matjo and Jiaqi, you guys give me hope that intelligent humble people exist, and that any topic can be an interesting topic to discuss. Thanks to Boris, that even if he keeps saying he's not social, he always found a way to show me his affection. Thanks to Jose and Niko, my Argentinian crew! thanks to the international school, especially Orna and Michal, that created a home for me in Israel. I never felt so attached to a place in such a short time. Thanks to Gil, for our breakfasts and our Hebrew-Italian classes, and to Gennady for our yard coffees. Thanks to Aviv, for becoming like an acquired brother in such a short time, and please…..no Egypt anymore!!!!!! Thanks to Mor for giving me all her unconditional love and pushing me to speak out my emotions. Thanks to Rad, Einel, Antonio, and everyone that everyday salutes me with a smile and gives me some positive energy…and a good coffee! But most of all, thanks to Hodaya! My crazy super different, and at the same time so similar to me, religious (for now, you know I always hope that one day we can go to the sauna together and then eat bacon fries) friend. Your energy, your temperament, your "never give up" attitude, and all the positivity that you are able to create around you even in difficult times, are such an example for all of us. I am so grateful that I could spend this time of my life with you, maybe I'm not a Jew but we for sure belong together. Getting back to old roots I want to thank my family, especially my dad for managing to not become crazy with a daughter like me. My brother and my grandma, for being always my secure pillars. I want to thank my "borc Gjavons" neighbors, that always save a spot for me on Christmas' Eve dinner. I want to thank my home friends, Raf, Jezz, Sere, for being so present in my life even with the distance. I want to thank Lee-Ann,

and all her multiple personalities! Nadia, my baby! I want to thank the Wild Types, for giving me the chance to play crazy wild music during my PhD! Thank you to Micha and Ralf, for chilling me down, and to Matteo, Valentina, Giannina, Anson, Helena. Thanks to Carlos, my greatest supervisor and friend at the same time! Thanks to Chiara, Casimiro (I still read all his best quotes and laugh at times), so many people that took part in this journey. But now, most of all, thank you to the three people that literally walked with me day by day, step by step, no matter how dark was the way and how impossible it felt. Every day, these three people, were there for me. In three different continents, but always so close to me. Thank you to Luca, Fra and Alena. You are the proof that the distance excuse is just blabbering. You are a gift in my life. Even though it was rather difficult times that pulled us close, I am sure the future has such great new adventures for us and a lot of happiness to share together!! And now, at the very end, thank you to Sabbath, my life companion, being like a sister and a mum at the same time, even though she is in fact a black cat. And thank you to myself, for not giving up.

You Rock!

BIBLIOGRAPHY

1. Wojtuszkiewicz, A. *et al.* Methotrexate resistance in relation to treatment outcome in childhood acute lymphoblastic leukemia. *J. Hematol. Oncol.* (2015). doi:10.1186/s13045-015-0158-9

2. Mahmood, L. The metabolic processes of folic acid and Vitamin B12 deficiency. *J. Heal. Res. Rev.* (2014). doi:10.4103/2394-2010.143318

3. Fazili, Z., Pfeiffer, C. M., Zhang, M., Jain, R. B. & Koontz, D. Influence of 5,10-methylenetetrahydrofolate reductase polymorphism on whole-blood folate concentrations measured by LC-MS/MS, microbiologic assay, and Bio-Rad radioassay. *Clin. Chem.* (2008). doi:10.1373/clinchem.2007.096545

4. Wishart, D. S. *et al.* HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Res.* (2018). doi:10.1093/nar/gkx1089

5. Bekaert, S. *et al.* Folate biofortification in food plants. *Trends in Plant Science* (2008). doi:10.1016/j.tplants.2007.11.001

6. Shea, B. *et al.* Folic acid and folinic acid for reducing side effects in patients receiving methotrexate for rheumatoid arthritis. *Cochrane Database of Systematic Reviews* (2013). doi:10.1002/14651858.CD000951.pub2

7. Greenberg, J. A., Bell, S. J., Guan, Y. & Yu, Y.-H. Folic Acid supplementation and pregnancy: more than just neural tube defect prevention. *Rev. Obstet. Gynecol.* (2011). doi:10.18370/2309-4117.2017.34.57-63

8. Yang, T. Y. *et al.* Effect of folic acid and vitamin B 12 on pemetrexed antifolate chemotherapy in nutrient lung cancer cells. *Biomed Res. Int.* (2013). doi:10.1155/2013/389046

9. Ducker, G. S. & Rabinowitz, J. D. One-Carbon Metabolism in Health and Disease. *Cell Metabolism* (2017). doi:10.1016/j.cmet.2016.08.009

10. Morgan, S. L., Oster, R. A., Lee, J. Y., Alarcón, G. S. & Baggott, J. E. The effect of folic acid and folinic acid supplements on purine metabolism in methotrexate-treated rheumatoid arthritis. *Arthritis Rheum.* (2004). doi:10.1002/art.20516

11. Safi, J., Joyeux, L. & Chalouhi, G. E. Periconceptional folate deficiency and implications in neural tube defects. *Journal of Pregnancy* (2012). doi:10.1155/2012/295083

12. Frigerio, B. *et al.* Folate receptors and transporters: Biological role and diagnostic/therapeutic targets in cancer and other diseases. *J. Exp. Clin. Cancer Res.* (2019). doi:10.1186/s13046-019-1123-1

13. Lawrence, S. A. *et al.* Mammalian mitochondrial and cytosolic folylpolyglutamate synthetase maintain the subcellular compartmentalization of folates. *J. Biol. Chem.* (2014). doi:10.1074/jbc.M114.593244

14. Tibbetts, A. S. & Appling, D. R. Compartmentalization of Mammalian Folate-Mediated One-Carbon Metabolism. *Annu. Rev. Nutr.* (2010). doi:10.1146/annurev.nutr.012809.104810

15. Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* (2011). doi:10.1038/nature10350

16. Jain, M. *et al.* Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science (80-.).* (2012). doi:10.1126/science.1218595

17. Tjong, E. & Mohiuddin, S. S. *Biochemistry, Tetrahydrofolate*. *StatPearls* (2019).

18. Singhal, N. K. *et al.* Changes in methionine metabolism and histone H3 trimethylation are linked to mitochondrial defects in multiple sclerosis. *J. Neurosci.* (2015). doi:10.1523/JNEUROSCI.4349-14.2015

19. Bhagavan, N. V & Ha, C.-E. Chapter 25 - Nucleotide Metabolism. in *Essentials of Medical Biochemistry* (2015). doi:https://doi.org/10.1016/B978-0-12-416687-5.00025-7

20. McCarty, M. F., O'Keefe, J. H. & DiNicolantonio, J. J. Dietary glycine is rate-limiting for glutathione synthesis and may have broad potential for health protection. *Ochsner Journal* (2018). doi:10.1043/TOJ-17-0022

21. Tramonti, A. *et al.* Human Cytosolic and Mitochondrial Serine Hydroxymethyltransferase Isoforms in Comparison: Full Kinetic Characterization and Substrate Inhibition Properties. *Biochemistry* (2018).

doi:10.1021/acs.biochem.8b01074

22. Anderson, D. D. & Stover, P. J. SHMT1 and SHMT2 are functionally redundant in nuclear de novo thymidylate biosynthesis. *PLoS One* (2009). doi:10.1371/journal.pone.0005839

23. Budai, B. *et al.* Impact of SHMT1 polymorphism on the clinical outcome of patients with metastatic colorectal cancer treated with first-line FOLFIRI+bevacizumab. *Pharmacogenet. Genomics* (2012). doi:10.1097/FPC.0b013e32834d8376

24. Ducker, G. S. *et al.* Human SHMT inhibitors reveal defective glycine import as a targetable metabolic vulnerability of diffuse large B-cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* (2017). doi:10.1073/pnas.1706617114

25. Mattaini, K. R., Sullivan, M. R. & Vander Heiden, M. G. The importance of serine metabolism in cancer. *Journal of Cell Biology* (2016). doi:10.1083/jcb.201604085

26. Curtin, S. C. Trends in cancer and heart disease death rates among adults aged 45–64: United States, 1999–2017. *Natl. Vital Stat. Reports* (2019).

27. Tanday, S. Cancer overtakes heart disease as biggest killer in 12 countries. *The Lancet. Oncology* (2016). doi:10.1016/S1470-2045(16)30403-X

28. National Cancer Institute. Defining Cancer. *National Cancer Institute* (2018). doi:10.4172/2167-0420.100016

29. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* (2011). doi:10.1016/j.cell.2011.02.013

30. Nijman, S. M. B. Synthetic lethality: General principles, utility and detection using genetic screens in human cells. *FEBS Letters* (2011). doi:10.1016/j.febslet.2010.11.024

31. Ward, P. S. & Thompson, C. B. Metabolic Reprogramming: A Cancer Hallmark Even Warburg Did Not Anticipate. *Cancer Cell* (2012). doi:10.1016/j.ccr.2012.02.014

32. De Berardinis, R. J. & Chandel, N. S. Fundamentals of cancer metabolism. *Science Advances* (2016). doi:10.1126/sciadv.1600200

33. Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends in Biochemical Sciences* (2016). doi:10.1016/j.tibs.2015.12.001

34. Locasale, J. W. Serine, glycine and one-carbon units: Cancer metabolism in full circle. *Nature Reviews Cancer* (2013). doi:10.1038/nrc3557

35. Visentin, M., Zhao, R. & Goldman, I. D. The Antifolates. *Hematology/Oncology Clinics of North America* (2012). doi:10.1016/j.hoc.2012.02.002

36. Takemura, Y., Kobayashi, H. & Miyachi, H. Cellular and molecular mechanisms of resistance to antifolate drugs: New analogues and approaches to overcome the resistance. *Int. J. Hematol.* (1997). doi:10.1016/S0925-5710(97)00058-3

37. Zhao, R. & Goldman, I. D. Resistance to antifolates. *Oncogene* (2003). doi:10.1038/sj.onc.1206946

38. Gonen, N. & Assaraf, Y. G. Antifolates in cancer therapy: Structure, activity and mechanisms of drug resistance. *Drug Resist. Updat.* (2012). doi:10.1016/j.drup.2012.07.002

39. Norris, R. E. & Adamson, P. C. Clinical potency of methotrexate, aminopterin, talotrexin and pemetrexed in childhood leukemias. *Cancer Chemother. Pharmacol.* (2010). doi:10.1007/s00280-009-1120-8

40. Canter, P. Herb, Nutrient and Drug Interactions: Clinical Implications and Therapeutic Strategies. *Focus Altern. Complement. Ther.* (2010). doi:10.1111/j.2042-7166.2008.tb01840.x

41. Mehrshahi, P. *et al.* Functional analysis of folate polyglutamylation and its essential role in plant metabolism and development. *Plant J.* (2010). doi:10.1111/j.1365-313X.2010.04336.x

42. Daidone, F. *et al.* In silico and in vitro validation of serine hydroxymethyltransferase as a chemotherapeutic target of the antifolate drug pemetrexed. *Eur. J. Med. Chem.* (2011). doi:10.1016/j.ejmech.2011.02.009

43. Luqmani, Y. A. Mechanisms of drug resistance in cancer chemotherapy. in *Medical Principles and Practice* (2005). doi:10.1159/000086183

44. Mansoori, B., Mohammadi, A., Davudian, S., Shirjang, S. & Baradaran, B. The different mechanisms of cancer drug resistance: A brief review. *Advanced Pharmaceutical Bulletin* (2017). doi:10.15171/apb.2017.041

45. Assaraf, Y. G. Molecular basis of antifolate resistance. *Cancer and Metastasis Reviews* (2007). doi:10.1007/s10555-007-9049-z

46. Zhao, R., Gao, F., Hanscom, M. & Goldman, I. D. A Prominent Low-pH Methotrexate Transport Activity in Human Solid Tumors: Contribution to the Preservation of Methotrexate Pharmacologic Activity in HeLa Cells Lacking the Reduced Folate Carrier. *Clin. Cancer Res.* (2004). doi:10.1158/1078-0432.CCR-1066-03

47. Raz, S., Stark, M. & Assaraf, Y. G. Folylpoly-γ-glutamate synthetase: A key determinant of folate homeostasis and antifolate resistance in cancer. *Drug Resist. Updat.* (2016). doi:10.1016/j.drup.2016.06.004

48. Liu, S. G. *et al.* FPGS rs1544105 polymorphism is associated with treatment outcome in pediatric B-cell precursor acute lymphoblastic leukemia. *Cancer Cell Int.* (2013). doi:10.1186/1475-2867-13-107

49. Hanauske, A.-R. Pemetrexed Disodium: A Novel Antifolate Clinically Active Against Multiple Solid Tumors. *Oncologist* (2001). doi:10.1634/theoncologist.6-4-363

50. Liani, E. *et al.* Loss of folylpoly-γ-glutamate synthetase activity is a dominant mechanism of resistance to polyglutamylation-dependent novel antifolates in multiple human leukemia sublines. *Int. J. Cancer* (2003). doi:10.1002/ijc.10829

51. Raz, S., Stark, M. & Assaraf, Y. G. Binding of a Smad4/Ets-1 complex to a novel intragenic regulatory element in exon12 of FPGS underlies decreased gene expression and antifolate resistance in leukemia. *Oncotarget* (2014). doi:10.18632/oncotarget.2399

52. D'Argenio, V. The high-throughput analyses era: Are we ready for the data

struggle? *High-Throughput* (2018). doi:10.3390/ht7010008

53. Hodge, K., Have, S. Ten, Hutton, L. & Lamond, A. I. Cleaning up the masses: Exclusion lists to reduce contamination with HPLC-MS/MS. *J. Proteomics* (2013). doi:10.1016/j.jprot.2013.02.023

54. Horgan, R. P. & Kenny, L. C. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. *Obstet. Gynaecol.* (2011). doi:10.1576/toag.13.3.189.27672

55. Leon-Mimila, P., Wang, J. & Huertas-Vazquez, A. Relevance of Multi-Omics Studies in Cardiovascular Diseases. *Front. Cardiovasc. Med.* (2019). doi:10.3389/fcvm.2019.00091

56. Ioannidis, J. P. A. & Khoury, M. J. Improving validation practices in 'omics' research. *Science* (2011). doi:10.1126/science.1211811

57. Hassani, S. *et al.* Analysis of -omics data: Graphical interpretation- and validation tools in multi-block methods. *Chemom. Intell. Lab. Syst.* (2010). doi:10.1016/j.chemolab.2010.08.008

58. Suravajhala, P., Kogelman, L. J. A. & Kadarmideen, H. N. Multi-omic data integration and analysis using systems genomics approaches: Methods and applications In animal production, health and welfare. *Genetics Selection Evolution* (2016). doi:10.1186/s12711-016-0217-x

59. Zubarev, R. A. The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics* (2013). doi:10.1002/pmic.201200451

60. Marcotte, E. M. Measuring the dynamics of the proteome. *Genome Research* (2001). doi:10.1101/gr.178301

61. Ghazalpour, A. *et al.* Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet.* (2011). doi:10.1371/journal.pgen.1001393

62. Wang, D. *et al.* A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Mol. Syst. Biol.* (2019). doi:10.15252/msb.20188503

63. Xie, H. *et al.* Proteomics analysis to reveal biological pathways and predictive proteins in the survival of high-grade serous ovarian cancer. *Sci. Rep.* (2017). doi:10.1038/s41598-017-10559-9

64. Fonseca, L. D. *et al.* Liver proteomics unravel the metabolic pathways related to Feed Efficiency in beef cattle. *Sci. Rep.* (2019). doi:10.1038/s41598-019-41813-x

65. Martens, L. Proteomics databases and repositories. *Methods Mol. Biol.* (2011). doi:10.1007/978-1-60761-977-2_14

66. Ong, S. E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* (2002). doi:10.1074/mcp.M200025-MCP200

67. Zhu, W., Smith, J. W. & Huang, C. M. Mass spectrometry-based label-free quantitative proteomics. *Journal of Biomedicine and Biotechnology* (2010). doi:10.1155/2010/840518

68. Deracinois, B., Flahaut, C., Duban-Deweer, S. & Karamanos, Y. Comparative and quantitative global proteomics approaches: An overview. *Proteomes* (2013). doi:10.3390/proteomes1030180

69. Mustafa, S. *et al.* Comparison of the tumor cell secretome and patient sera for an accurate serum-based diagnosis of pancreatic ductal adenocarcinoma. *Oncotarget* (2017). doi:10.18632/oncotarget.14449

70. Alhamdani, M. S. S., Schröder, C. & Hoheisel, J. D. Analysis conditions for proteomic profiling of mammalian tissue and cell extracts with antibody microarrays. *Proteomics* (2010). doi:10.1002/pmic.201000170

71. Hoheisel, J. D., Alhamdani, M. S. S. & Schröder, C. Affinity-based microarrays for proteomic analysis of cancer tissues. *Proteomics - Clinical Applications* (2013). doi:10.1002/prca.201200114

72. Betzen, C. *et al.* Clinical proteomics: Promises, challenges and limitations of affinity arrays. *Proteomics - Clinical Applications* (2015). doi:10.1002/prca.201400156

73. Powers, R. & Riekeberg, E. New frontiers in metabolomics: From measurement to insight. *F1000Research* (2017). doi:10.12688/f1000research.11495.1

74. Antoniewicz, M. R., Kelleher, J. K. & Stephanopoulos, G. Elementary metabolite units (EMU): A novel framework for modeling isotopic distributions. *Metab. Eng.* (2007). doi:10.1016/j.ymben.2006.09.001

75. Wiechert, W., Möllney, M., Isermann, N., Wurzel, M. & De Graaf, A. A. Bidirectional reaction steps in metabolic networks: III. Explicit solution and analysis of isotopomer labeling systems. *Biotechnol. Bioeng.* (1999). doi:10.1002/(SICI)1097-0290(1999)66:2<69::AID-BIT1>3.0.CO;2-6

76. Price, N. D., Papin, J. A., Schilling, C. H. & Palsson, B. O. Genome-scale microbial in silico models: The constraints-based approach. *Trends in Biotechnology* (2003). doi:10.1016/S0167-7799(03)00030-1

77. Jang, C., Chen, L. & Rabinowitz, J. D. Metabolomics and Isotope Tracing. *Cell* (2018). doi:10.1016/j.cell.2018.03.055

78. Paul Lee, W. N., Wahjudi, P. N., Xu, J. & Go, V. L. Tracer-based metabolomics: Concepts and practices. *Clinical Biochemistry* (2010). doi:10.1016/j.clinbiochem.2010.07.027

79. Maguire, G., Boros, L. & Lee, P. Development of tracer-based metabolomics and its implications for the pharmaceutical industry. *International Journal of Pharmaceutical Medicine* (2007). doi:10.2165/00124363-200721030-00004

80. Zuckerkandl, E. & Cavalli, G. Combinatorial epigenetics, 'junk DNA', and the evolution of complex organisms. *Gene* (2007). doi:10.1016/j.gene.2006.12.001

81. Šerman, A., Vlahović, M., Šerman, L. & Bulić-Jakuš, F. DNA methylation as a regulatory mechanism for gene expression in mammals. *Collegium Antropologicum* (2006).

82. Gaj, T., Gersbach, C. A. & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology* (2013). doi:10.1016/j.tibtech.2013.04.004

83. Humbert, O., Davis, L. & Maizels, N. Targeted gene therapies: Tools,

applications, optimization. *Critical Reviews in Biochemistry and Molecular Biology* (2012). doi:10.3109/10409238.2012.658112

84. Heynen-Genel, S., Pache, L., Chanda, S. K. & Rosen, J. Functional genomic and high-content screening for target discovery and deconvolution. *Expert Opinion on Drug Discovery* (2012). doi:10.1517/17460441.2012.711311

85. Guichard, S. M. CRISPR–Cas9 for Drug Discovery in Oncology. in *Annual Reports in Medicinal Chemistry* (2017). doi:10.1016/bs.armc.2017.08.006

86. Sanchez-Vega, F. *et al.* Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* (2018). doi:10.1016/j.cell.2018.03.035

87. Espinal-Enríquez, J. *et al.* Genome-wide expression analysis suggests a crucial role of dysregulation of matrix metalloproteinases pathway in undifferentiated thyroid carcinoma. *BMC Genomics* (2015). doi:10.1186/s12864-015-1372-0

88. Han, J. *et al.* Genome-wide CRISPR/Cas9 Screen Identifies Host Factors Essential for Influenza Virus Replication. *Cell Rep.* (2018). doi:10.1016/j.celrep.2018.03.045

89. Callow, M. G. *et al.* CRISPR whole-genome screening identifies new necroptosis regulators and RIPK1 alternative splicing article. *Cell Death Dis.* (2018). doi:10.1038/s41419-018-0301-y

90. Pertesi, M. *et al.* Essential genes shape cancer genomes through linear limitation of homozygous deletions. *Commun. Biol.* (2019). doi:10.1038/s42003-019-0517-0

91. Rao, D. D., Vorhies, J. S., Senzer, N. & Nemunaitis, J. siRNA vs. shRNA: Similarities and differences. *Advanced Drug Delivery Reviews* (2009). doi:10.1016/j.addr.2009.04.004

92. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* (2016). doi:10.1038/nprot.2016.136

93. Tyanova, S. & Cox, J. Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. in *Methods in Molecular Biology* (2018).

doi:10.1007/978-1-4939-7493-1_7

94. Huang, D. W. *et al.* The DAVID Gene Functional Classification Tool: A novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.* (2007). doi:10.1186/gb-2007-8-9-r183

95. Ingenuity Systems. Ingenuity Pathway Analysis. *Www.Ingenuity.Com* (2013).

96. Clasquin, M. F., Melamud, E. & Rabinowitz, J. D. LC-MS data processing with MAVEN: A metabolomic analysis and visualization engine. *Curr. Protoc. Bioinforma.* (2012). doi:10.1002/0471250953.bi1411s37

97. Heigwer, F., Kerr, G. & Boutros, M. E-CRISP: Fast CRISPR target site identification. *Nature Methods* (2014). doi:10.1038/nmeth.2812

98. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* (2009). doi:10.1186/gb-2009-10-3-r25

99. Salzberg, S. L. Next-generation genome annotation: We still struggle to get it right. *Genome Biology* (2019). doi:10.1186/s13059-019-1715-2

100. Rardin, M. J. Rapid Assessment of Contaminants and Interferences in Mass Spectrometry Data Using Skyline. *J. Am. Soc. Mass Spectrom.* (2018). doi:10.1007/s13361-018-1940-z

101. Rexach, J., Swarup, V., Chang, T. & Geschwind, D. Dementia risk genes engage gene networks poised to tune the immune response towards chronic inflammatory states. *bioRxiv* (2019). doi:10.1101/597542

102. Guo, S. *et al.* CRISPR/Cas9 Systems: The Next Generation Gene Targeted Editing Tool. *Proceedings of the National Academy of Sciences India Section B - Biological Sciences* (2015). doi:10.1007/s40011-014-0362-3

103. Mei, Y., Wang, Y., Chen, H., Sun, Z. S. & Ju, X. Da. Recent Progress in CRISPR/Cas9 Technology. *Journal of Genetics and Genomics* (2016). doi:10.1016/j.jgg.2016.01.001

104. Eid, A. & Mahfouz, M. M. Genome editing: The road of CRISPR/Cas9 from bench

to clinic. *Experimental and Molecular Medicine* (2016). doi:10.1038/emm.2016.111

105. Burgio, G. Redefining mouse transgenesis with CRISPR/Cas9 genome editing technology. *Genome Biol.* (2018). doi:10.1186/s13059-018-1409-1

106. Gurumurthy, C. B. *et al.* Reproducibility of CRISPR-Cas9 methods for generation of conditional mouse alleles: A multi-center evaluation. *Genome Biol.* (2019). doi:10.1186/s13059-019-1776-2

107. Gurumurthy, C. B. & Kent Lloyd, K. C. Generating mouse models for biomedical research: Technological advances. *DMM Dis. Model. Mech.* (2019). doi:10.1242/dmm.029462

108. Qin, W. *et al.* Generating Mouse Models Using CRISPR-Cas9-Mediated Genome Editing. *Curr. Protoc. Mouse Biol.* (2016). doi:10.1002/9780470942390.mo150178

109. Jiang, F. & Doudna, J. A. CRISPR–Cas9 Structures and Mechanisms. *Annu. Rev. Biophys.* (2017). doi:10.1146/annurev-biophys-062215-010822

110. Gleditzsch, D. *et al.* PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. *RNA Biology* (2019). doi:10.1080/15476286.2018.1504546

111. Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* (2008). doi:10.4161/cc.7.18.6679

112. Cubbon, A., Ivancic-Bace, I. & Bolt, E. L. CRISPR-Cas immunity, DNA repair and genome stability. *Bioscience Reports* (2018). doi:10.1042/BSR20180457

113. Liu, M. *et al.* Methodologies for improving HDR efficiency. *Frontiers in Genetics* (2019). doi:10.3389/fgene.2018.00691

114. Swarts, D. C. & Jinek, M. Cas9 versus Cas12a/Cpf1: Structure–function comparisons and implications for genome editing. *Wiley Interdisciplinary Reviews: RNA* (2018). doi:10.1002/wrna.1481

115. Dominguez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: Repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nature Reviews Molecular Cell Biology* (2016). doi:10.1038/nrm.2015.2

116. Kampmann, M. CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem. Biol.* (2018). doi:10.1021/acschembio.7b00657

117. Dhanjal, J. K., Radhakrishnan, N. & Sundar, D. Identifying synthetic lethal targets using CRISPR/Cas9 system. *Methods* (2017). doi:10.1016/j.ymeth.2017.07.007

118. Pacold, M. E. *et al.* A PHGDH inhibitor reveals coordination of serine synthesis and one-carbon unit fate. *Nat. Chem. Biol.* (2016). doi:10.1038/nchembio.2070

119. Millard, P., Letisse, F., Sokol, S. & Portais, J. C. IsoCor: Correcting MS data in isotope labeling experiments. *Bioinformatics* (2012). doi:10.1093/bioinformatics/bts127

120. Melamud, E., Vastag, L. & Rabinowitz, J. D. Metabolomic analysis and visualization engine for LC - MS data. *Anal. Chem.* **82**, 9818–9826 (2010).

121. Bennett, B. D., Yuan, J., Kimball, E. H. & Rabinowitz, J. D. Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach. *Nat. Protoc.* (2008). doi:10.1038/nprot.2008.107

122. Chen, L., Ducker, G. S., Lu, W., Teng, X. & Rabinowitz, J. D. An LC-MS chemical derivatization method for the measurement of five different one-carbon states of cellular tetrahydrofolate. *Anal. Bioanal. Chem.* (2017). doi:10.1007/s00216-017-0514-4

123. Alhamdani, M. S. S. *et al.* Single-step procedure for the isolation of proteins at near-native conditions from mammalian tissue for proteomic analysis on antibody microarrays. *J. Proteome Res.* (2010). doi:10.1021/pr900844q

124. Alhamdani, M. S. S. *et al.* Immunoassay-based proteome profiling of 24 pancreatic cancer cell lines. *J. Proteomics* (2012). doi:10.1016/j.jprot.2012.04.042

125. Boettcher, M. *et al.* Dual gene activation and knockout screen reveals

directional dependencies in genetic networks. *Nat. Biotechnol.* (2018). doi:10.1038/nbt.4062

126. Meiser, J. *et al.* Increased formate overflow is a hallmark of oxidative cancer. *Nat. Commun.* (2018). doi:10.1038/s41467-018-03777-w

127. Meiser, J. *et al.* Serine one-carbon catabolism with formate overflow. *Sci. Adv.* (2016). doi:10.1126/sciadv.1601273

128. Ducker, G. S. *et al.* Reversal of Cytosolic One-Carbon Flux Compensates for Loss of the Mitochondrial Folate Pathway Article Reversal of Cytosolic One-Carbon Flux Compensates for Loss of the Mitochondrial Folate Pathway. *Cell Metab.* **23**, 1140–1153 (2016).

129. Hou, Z. & Matherly, L. H. Biology of the major facilitative folate transporters SLC19A1 and SLC46A1. in *Current Topics in Membranes* (2014). doi:10.1016/B978-0-12-800223-0.00004-9

130. Rothem, L. *et al.* Resistance to multiple novel antifolates is mediated via defective drug transport resulting from clustered mutations in the reduced folate carrier gene in human leukaemia cell lines. *Biochem. J.* (2002). doi:10.1042/BJ20020801

131. Yang, S., Zhou, X., Li, R., Fu, X. & Sun, P. Optimized PEI-based Transfection Method for Transient Transfection and Lentiviral Production. *Curr. Protoc. Chem. Biol.* (2017). doi:10.1002/cpch.25

132. Andrianifahanana, M. *et al.* Mucin (MUC) gene expression in human pancreatic adenocarcinoma and chronic pancreatitis: A potential role of MUC4 as a tumormarker of diagnostic significance. *Clin. Cancer Res.* (2001).

133. Seppen, J. *et al.* Apical Gene Transfer into Quiescent Human and Canine Polarized Intestinal Epithelial Cells by Lentivirus Vectors. *J. Virol.* (2000). doi:10.1128/jvi.74.16.7642-7645.2000

134. Gerhards, N. M. & Rottenberg, S. New tools for old drugs: Functional genetic screens to optimize current chemotherapy. *Drug Resist. Updat.* (2018). doi:10.1016/j.drup.2018.01.001

135. Banerjee, D. *et al.* Molecular mechanisms of resistance to antifolates, a review. *Acta biochimica Polonica* (1995).

136. Wu, C., Sun, M., Liu, L. & Zhou, G. W. The function of the protein tyrosine phosphatase SHP-1 in cancer. *Gene* (2003). doi:10.1016/S0378-1119(03)00400-1

137. Chim, C. S., Wong, K. Y., Loong, F. & Srivastava, G. SOCS1 and SHP1 hypermethylation in mantle cell lymphoma and follicular lymphoma: Implications for epigenetic activation of the Jak/STAT pathway [2]. *Leukemia* (2004). doi:10.1038/sj.leu.2403216

138. Zhang, X. *et al.* Research on the epigenetic regulation mechanism of the PTPN6 gene in advanced chronic myeloid leukaemia. *Br. J. Haematol.* (2017). doi:10.1111/bjh.14739

139. Deininger, M., Buchdunger, E. & Druker, B. J. The development of imatinib as a therapeutic agent for chronic myeloid leukemia. *Blood* (2005). doi:10.1182/blood-2004-08-3097

140. Aarabi, M. *et al.* High-dose folic acid supplementation alters the human sperm methylome and is influenced by the MTHFR C677T polymorphism. *Hum. Mol. Genet.* (2015). doi:10.1093/hmg/ddv338

141. Yip, K. W. & Reed, J. C. Bcl-2 family proteins and cancer. *Oncogene* (2008). doi:10.1038/onc.2008.307

142. Pham-Ledard, A. *et al.* Diagnostic and prognostic value of BCL2 rearrangement in 53 patients with follicular lymphoma presenting as primary skin lesions. *Am. J. Clin. Pathol.* (2015). doi:10.1309/AJCP4SUBR4NPSPTN

143. Ding, J. *et al.* Bcl-2 and Bax interact via the BH1-3 groove-BH3 motif interface and a novel interface involving the BH4 motif. *J. Biol. Chem.* (2010). doi:10.1074/jbc.M110.148361

144. Gross, A., Jockel, J., Wei, M. C. & Korsmeyer, S. J. Enforced dimerization of BAX results in its translocation, mitochondrial dysfunction and apoptosis. *EMBO J.* (1998). doi:10.1093/emboj/17.14.3878

145. Chipuk, J. E. *et al.* Direct Activation of Bax by p53 Mediates Mitochondrial Membrane Permeabilization and Apoptosis. *Science (80-.).* (2004). doi:10.1126/science.1092734

146. Miyashita, T. *et al.* Tumor suppressor p53 is a regulator of bcl-2 and bax gene expression in vitro and in vivo. *Oncogene* (1994).

147. Zhan, Q. *et al.* Induction of bax by genotoxic stress in human cells correlates with normal p53 status and apoptosis. *Oncogene* (1994).

148. Toshiyuki, M. & Reed, J. C. Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell* (1995). doi:10.1016/0092-8674(95)90412-3

149. Wang, D. *et al.* OIP5 Promotes growth, metastasis and chemoresistance to cisplatin in bladder cancer cells. *J. Cancer* (2018). doi:10.7150/jca.27381

150. Chun, H. K. *et al.* OIP5 is a highly expressed potential therapeutic target for colorectal and gastric cancers. *BMB Rep.* (2010). doi:10.5483/BMBRep.2010.43.5.349

151. Li, Y. *et al.* Overexpression of Opa interacting protein 5 increases the progression of liver cancer via BMPR2/JUN/CHEK1/RAC1 dysregulation. *Oncol. Rep.* (2019). doi:10.3892/or.2019.7006

152. Gatta, E. *et al.* Emerging role of one-carbon metabolism and DNA methylation enrichment on δ-containing GABAA receptor expression in the cerebellum of subjects with Alcohol Use Disorders (AUD). *Int. J. Neuropsychopharmacol.* (2017). doi:10.1093/ijnp/pyx075

153. Serefidou, M., Venkatasubramani, A. V. & Imhof, A. The Impact of One Carbon Metabolism on Histone Methylation. *Front. Genet.* **10**, 764 (2019).

154. Parkhitko, A. A. *et al.* Tissue-specific down-regulation of S-adenosyl-homocysteine via suppression of dAhcyL1/dAhcyL2 extends health span and life span in Drosophila. *Genes Dev.* (2016). doi:10.1101/gad.282277.116

155. Belužić, L. *et al.* Knock-down of AHCY and depletion of adenosine induces DNA damage and cell cycle arrest. *Sci. Rep.* (2018). doi:10.1038/s41598-018-32356-

8

156. Ulrey, C. L., Liu, L., Andrews, L. G. & Tollefsbol, T. O. The impact of metabolism on DNA methylation. *Human Molecular Genetics* (2005). doi:10.1093/hmg/ddi100

157. Crider, K. S., Yang, T. P., Berry, R. J. & Bailey, L. B. Folate and DNA Methylation: A Review of Molecular Mechanisms and the Evidence for Folate's Role. *Adv. Nutr.* (2012). doi:10.3945/an.111.000992

158. Chen, N. C. *et al.* Regulation of homocysteine metabolism and methylation in human and mouse tissues. *FASEB J.* (2010). doi:10.1096/fj.09-143651

159. Soda, K. Polyamine metabolism and gene methylation in conjunction with one-carbon metabolism. *International Journal of Molecular Sciences* (2018). doi:10.3390/ijms19103106

160. Coleman, C. S., Stanley, B. A., Jones, A. D. & Pegg, A. E. Spermidine/spermine-N1-acetyltransferase-2 (SSAT2) acetylates thialysine and is not involved in polyamine metabolism. *Biochem. J.* (2004). doi:10.1042/BJ20040790

161. Lagziel, S., Lee, W. D. & Shlomi, T. Inferring cancer dependencies on metabolic genes from large-scale genetic screens. *BMC Biol.* (2019). doi:10.1186/s12915-019-0654-4

162. Jiang, W. *et al.* An optimized method for high-titer lentivirus preparations without ultracentrifugation. *Sci. Rep.* (2015). doi:10.1038/srep13875