Master's Thesis

Investigation of Paralog Dependencies for Targeted Cancer Therapy

carried out for the purpose of obtaining the degree of Master of Science (MSc)

submitted at TU Wien, Faculty of Technical Chemistry

Fiona SPREITZER

Mat. Nr.: 01525661

Under the supervision of

Dr. Ralph Neumüller

Boehringer Ingelheim RCV GmbH & Co KG, Vienna, Austria

and

Univ.Prof. Mag. Dr.rer.nat. Robert Mach

Institute of Chemical, Environmental and Bioscience Engineering, TU Wien

This master's thesis was conducted at Boehringer-Ingelheim RCV in the department Cancer Research – Cancer Cell Signaling.

The thesis will be locked and hence not publicly available for 2 years, effective from March 2021 onwards.

1 Abstract

Synthetic lethality is a concept where the simultaneous loss of two genes leads to cell death while deficiency of only one of the genes has no negative effect on cell survival. Paralog dependency constitutes a special form of synthetic lethality. Paralog genes in a paralog dependency, these two genes have emerged from gene duplication and are believed to be functionally redundant. Paralog dependencies represent a promising opportunity for targeted cancer therapy because if one gene is lost due to a cancer-specific alteration, a drug inhibiting the corresponding paralog only affects cancer cells. Healthy cells are spared as they do not rely on only one paralog. Therefore, it is anticipated that targeted cancer therapy elevates problems caused by chemotherapy. In this master's thesis, six potential paralog pairs were assessed: DNAJC15/19, RPP25/25L, PAPSS1/2, PRPS1/2, SLC25A28/37 and VPS4A/B. The goal was to verify whether these gene pairs are functionally redundant and to assess their potential as therapeutic targets for cancer therapy. Two paralog dependencies could be confirmed. DNAJC19 inhibition can be used to target cancers, in which the biomarker DNAJC15 is silenced through DNA methylation. Both proteins, DNAJC15 and DNAJC19, are part of a mitochondrial complex, required to transport precursor proteins across the inner mitochondrial membrane. Interestingly, the loss of DNAJC15 through DNA methylation has been associated with chemotherapeutic resistance. The second confirmed paralog dependency is RPP25 and RPP25L. These proteins are subunits of RNase P, which processes precursor tRNA as well as other cellular RNA. Both targets led to a block in cell proliferation when knocked out in cancer cell lines which do not express the respective paralog. This sensitivity can be mitigated when the biomarker paralog is overexpressed.

Synthetische Letalität beschreibt das Verhältnis zweier Gene, in welchem der gleichzeitige Verlust beider Gene zum Zelltod führt, während es keine negativen Effekte für die Zelle gibt, wenn nur eines der Gene fehlt. Man spricht von einer paralogen Abhängigkeit, wenn dieses Genpaar durch Genduplikation entstanden ist. Paraloge Abhängigkeiten stellen eine vielversprechende Gelegenheit für die gezielte Krebstherapie dar, denn wenn eines der Gene durch eine krebsspezifische Veränderung verloren geht, schädigt ein Wirkstoff, welcher das andere Paralog inhibiert, nur diese Krebszellen. Gesunde Zellen werden nicht beeinträchtigt, da diese nicht von nur einem Paralog abhängig sind. Daher können mit einer gezielten Krebstherapie Probleme, welche in der klassischen Chemotherapie auftreten, vermieden werden. In dieser Masterarbeit wurden sechs potentielle Paralogpaare untersucht: DNAJC15/19, RPP25/25L, PAPSS1/2, PRPS1/2, SLC25A28/37 und VPS4A/B. Das Ziel war die Verifikation einer paralogen Abhängigkeit zwischen diesen Genen und die Untersuchung des Potentials der jeweiligen Target-Gene für die gezielte Krebstherapie. Zwei paraloge Abhängigkeiten konnten bestätigt werden: Durch die Inhibition von DNAJC19 können gezielt Tumore behandelt werden, in welchen die Expression des Biomarkers DNAJC15 durch DNA Methylierung verhindert wird. Diese beiden Proteine sind Teil des Transportproteins TIM23, welches Precursor-Proteine durch die innere Mitochondrienmembran befördert. Zusätzlich wurde der Verlust von DNAJC15 durch DNA Methylierung mit Resistenz gegen Chemotherapie assoziiert. Die zweite bestätigte paraloge Abhängigkeit besteht zwischen RPP25 und RPP25L. Diese Proteine sind Untereinheiten von RNase P, welche für die Prozessierung von verschiedenen RNAs, unter anderem Precursor-tRNAs, zuständig ist. In beiden Fällen konnte beobachtet werden, dass Krebszellenlinien, welche das jeweilige Biomarker-Gen nicht exprimieren, einen erheblichen Proliferationsnachteil haben, wenn das Target-Gen ausgeschalten wird. Diese Empfindlichkeit kann jedoch verhindert werden, wenn das entsprechende Biomarker-Gen überexprimiert wird.

2   Table of Contents

3   List of Abbreviations

2,3-DPG	2,3-Diphosphoglycerate	MRP	mitochondiral RNA processing
ADP	adenosin diphosphate	MYC	myelocytomatosis proto-oncogene protein
AMP	adenosin monophosphate	NHEJ	non-homologous end joining
APP-MP	4-amino-8-(beta-D-ribofuranosylamino)pyrimido[5,4-d]pyrimidine-5'-monophosphate	OE	Overexpression
ATM	Ataxia Telangiectasia Mutated protein	PARK2	Parkinson protein 2
BRCA	breast cancer tumor supressor	PARP	poly-ADP-ribose-polymerase
Cas9	CRISPR associated protein 9	PINK1	PTEN-induced kinase 1
CRISPR	clustered regular interspaced short palindromic repeats	PRPP	5-phosphoribosyl 1-pyrophosphate
CRISPRi	CRISPR interference	PRPS	phosphoribosyl pyrophosphate synthetase
dCas9	deactivated Cas9	PTEN	phosphatase and tensin homolog
DNA	deoxyribonucleic acid	RAD52	DNA repair protein RAD52 homolog
DNAJ	protein family orthologous to Escherichia coli DnaJ/HSP40	RISC	RNA-induced silencing complex
DNMT	DNA methyltransferase	RNA	ribonucleic acid
DSB	double strand break	RNAi	RNA interference
ENO	enolase (2-phospho-D-glycerate hydrolases)	RPP	ribonuclease P protein subunit
GDP	guanosin diphosphate	rRNA	ribosomal RNA
gRNA	guide RNA	shRNA	short hairpin RNA
HR	homologous recombination	siRNA	short interferening RNA
HSP	Heat shock protein	SLC	solute carrier protein
KO	Knock out	SSB	single strand break
Kras	Kirsten Rat Sarcoma virus oncogene	TIM	Mitochondrial import inner membrane translocase subunit
MBP-1	MYC promotor binding protein 1	tRNA	transfer RNA
MCJ	methylation controlled J-domain protein (DNAJC15)	WT	wild type
Mfrn	mitoferrin		

4 Introduction

4.1 Importance of targeted cancer therapy

According to current reports of the World Health Organization, cancer causes one in six deaths worldwide. 30% of all deaths between the ages of 30 and 69 are caused by cancer, making this disease a leading cause of premature death in a majority of countries all over the world. As the number of patients suffering from cancer is predicted to increase in the future, effective prevention and treatment is of utmost importance (1). "Cancer" is an umbrella term, representing a large group of diseases characterized by uncontrolled cell proliferation. Cancerous cells can spread to other tissues and organs and interfere with organ function and homeostasis. These metastases lead in many cases to death (1). At its core, cancer is a genetic disease (2). Tumors develop through stepwise acquirement of genetic alterations finally leading to neoplastic transformation (3; 4; 5; 6). These so called driver mutations can be classified into two distinct categories: Oncogenes refer to genes upregulated in cancer cells stimulating proliferation and improving survival while tumor suppressor genes are genes which promote neoplastic transformation when lost. Tumor suppressor genes are often involved in regulating cell growth and promoting DNA repair (7).

Various approaches have been developed to treat cancer including surgery, radiotherapy, chemotherapy and, more recently, targeted therapy and immunotherapy (8; 9). Many chemotherapeutics were developed aiming to kill rapidly growing cancer cells (10). Unfortunately, these drugs can be toxic to healthy cells, especially to other frequently dividing cells (11). Although cytotoxic chemotherapy has been beneficial for large number of patients, in more recent years treatment effects have plateaued due to a lack of selectivity between normal and cancer cells (8). Hence, the focus shifted to more personalized approaches (4). The goal of targeted cancer therapy is to provide drugs specific to cancer cells without affecting healthy cells. Approaches for this involve monoclonal antibodies, prodrug therapies and small molecule drugs interacting with cancer-specific essential processes (8; 12). In the past, the choice of novel drug targets for cancer therapy was based on genes and pathways that are mutated in cancer cells (13). This, however, reduces the number of possible targets severely, especially in the light of the fact that not every oncogene is easily druggable. In particular, loss-of-function mutations in tumor suppressor genes are hard to tackle as the genes are either largely inactive or completely absent and therefore not targetable with compounds (3). One way of increasing the pool of promising drug targets is investigating synthetic lethality in cancer cells (14).

4.2 Synthetic lethality and paralog dependencies

4.2.1 Concept of synthetic lethality and paralog dependencies

The term synthetic lethality describes the relationship of two or more genes of which a simultaneous loss of two genes results in cell death, whereas alterations in either gene alone do not impede viability (Figure 1) (4; 3; 11). Generally, synthetic lethality is associated with loss-of-function mutations even though the concept can be extended to gain-of-function mutations and overexpression of a certain gene, termed synthetic dosage lethality. Here, the overexpression of one gene is not compatible with the loss of another gene (15; 16).

There are multiple models of synthetically lethal interactions (Figure 1): Firstly, if two pathways exist which promote cell survival, mutations in either pathway can be compensated by the other. However, if a second mutation in the now essential pathway occurs, it will be lethal (14; 11). Genes of multi-protein

5

complexes may be in a synthetic lethality relation, as one altered gene does not necessarily hinder the functionality of the complex while multiple changes do (17). Finally, synthetic lethality also occurs in the form of paralog dependencies on which this study will be focused on.
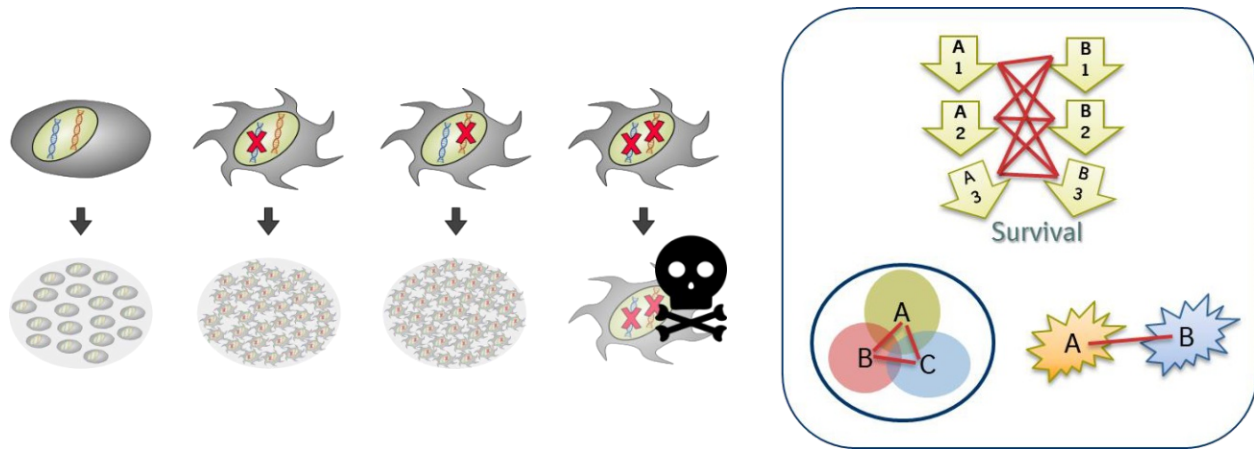
Figure 1: Basic Principal of Synthetic Lethality
Left: Loss of either gene is tolerated whereas simultaneous loss-of-function is lethal. Right: Models for synthetic lethality, red lines symbolize synthetic lethality interactions

Paralogs are genes that were duplicated in their evolutionary history. Due to their shared past these genes are often redundant (18). This way, mutations in either gene could be tolerated by the cells as they can rely on the other gene, however, the loss of both genes is lethal (19; 20). Paralogs are generally less likely to be essential than singletons, however, this depends on the genetic background as through mutations either gene can become crucial for survival (19). In the course of evolution, paralogs evolve independently from each other, therefore the later a paralog pair emerged the higher their similarity and their ability to buffer each other's loss of function. It is therefore speculated that evolutionary distance negatively correlates with functional redundancy. As the overlapping function does not offer a bona fide selection advantage, one of the two genes often functionally diverges or loses its function over time (21). Older paralogs on the other hand may have diverged so much, that both genes become essential. For instance, heteromeric paralogs emerging from ancestral homomers are less likely to show synthetic lethality (22). Nevertheless, as shown by De Kegel and Ryan, the vast majority of human genes is never or only sometimes essential which can be explained by paralog buffering (19). Similar results emerged from studies of model organisms like *Saccharomyces cerevisiae* and *Caenorhabditis elegans* (3). Paralog pairs that participate in protein complexes are more likely to show synthetic lethality (19).

Synthetic lethality is gaining attention as a concept for cancer therapy (11; 4). In theory, all genes that have a synthetically lethal relationship with a mutated gene in the cancer genome are possible targets. This is especially promising if the mutated synthetically lethal gene is a tumor suppressor gene or otherwise difficult to target directly. (3; 14; 11; 23) Besides mutations in tumor suppressor or oncogenic genes any other tumor-specific alteration (24) could cause synthetic lethality and hence lead to novel therapeutics (4). This also includes changes in expression and epigenetic modifications. As the mutation frequency in cancer is considerably increased, loss-of-function mutations in non-essential genes like paralogs can accumulate (20). Healthy cells should remain viable as they do not have these mutations

(4). Hence, the therapeutic index[i] of drugs exploiting synthetic lethality should be significantly improved compared to traditional chemotherapeutic agents (11).

In order to exploit synthetic lethalities for cancer therapy, the drug target does not need to be mutated. For this reason, these genes might not have been considered in the past as cancer-specific targets. Hence, there is the potential to increase the number of cancer relevant drug targets significantly (14). Furthermore, synthetic lethality offers a knowledge-based approach for combination therapy (16). This can enhance already existing as well as possibly abandoned drugs, which were only targeting a single pathway, however, in the right combination could proof successful. Combination therapy can also deal with secondary mutations which render the initial medication ineffective (14; 11).

4.2.2   Exploration of synthetic lethality

Initially, synthetic lethality was observed in *Drosophila melanogaster*, where certain combinations of homologous chromosomes, each viable as homozygotes, are not viable as heterozygotes (25; 26). Early genetic screens investigated synthetic lethality in model organisms *Saccharomyces cerevisiae*, *Drosophila melanogaster* or *Caenorhabditis elegans* by random mutagenesis and cross-breading followed by phenotypic screening (11; 19). Genes found in model organisms have often paved the way for the identification of new therapeutic targets in the space of synthetic lethality with tumor-associated mutations (27; 16).

In chemical screens, wild type and cancer cells are treated with various chemical compounds. If a certain drug-like chemical specifically kills only the cancer cells, the compound likely targets a synthetically lethal partner to a mutated gene in the cancer cell. This method prerequisites a stock of chemicals to apply and therefore is relatively limited. Additionally, target identification and, consequently, knowledge-based approaches to further improve this therapy, can be challenging. (11; 16; 28)

Nowadays, dependence on individual genes can be investigated with genetic screens applying libraries of short interfering RNAs (siRNAs), short hairpin RNAs (shRNAs) or guide RNAs (gRNAs) for CRISPR-Cas9 screens (14; 4; 16). Both siRNA and shRNA screens are based on RNA interference (RNAi) where these small synthetic RNAs cause translational silencing or degradation of matching mRNAs through base pairing and followed by endonucleolytic cleavage by the RNA-induced silencing complex (RISC) (29). On the other hand, CRISPR-Cas9 technologies exploit the ability of the Cas9 protein to create double-strand breaks in the DNA corresponding to a given gRNA. Without a template for homologous recombination (HR), these breaks are repaired by error-prone non-homologous end joining (NHEJ) frequently leading to deletions or insertions (30; 19). Consequently, gene disruption is complete with CRISPR, whereas RNAi causes merely posttranscriptional silencing which can vary greatly in efficiency between different genes. This makes CRISPR screens more straightforward when it comes to interpreting screening results, however, one can argue that RNAi screens more closely mimic the hypomorphic effect of a drug (14). Furthermore, shRNA screens suffer more from off-target effects (30; 4). One possibility of combining gene silencing with the accuracy of CRISPR-Cas9 is CRISPR interference (CRISPRi). Here, the catalytically inactive Cas9 (dCas9) is applied to transcriptionally silence the gene of interest indicated by the gRNA (30; 14; 4).

---

[i] Therapeutic index: Ratio of drug dosage (concentration) causing toxic effects to dosage (concentration) for therapeutic effects (11)

Genetic screens can be set up as arrayed screens similar to compound screens, while shRNAs and gRNAs can also be used for pooled screens (30; 11). In a pooled screen, the whole library is added to the examined cells, however, each cell should be effected by only one shRNA or gRNA. Afterwards, the relative abundance of individual shRNAs or gRNAs is compared before and after prolonged cell culture. Depleted shRNAs or gRNAs should correspond to essential genes. This approach only became feasible with the development of high-throughput sequencing (14).

Specifically for paralog dependencies, mining of expression databases or of cancer genomics databases might reveal genes that are members of paralog families that are underexpressed in cancer cells. These cells then may rely on a corresponding paralog (20). Combining this information with results from genome-wide shRNA or CRISPR screens can corroborate possible paralog dependency hits (18; 31; 28). The first therapeutic that was approved based on the principal of synthetic lethality, the PARP inhibitor olaparib, was only driven by hypothesis of two different pathways in DNA repair compensating each other (23; 32; 33).

4.2.3 Established synthetically lethal gene PARP in BRCA-deficient tumor cells

The PARP (Poly-ADP-ribose-polymerase) protein family consists of 18 nuclear enzymes, of which PARP-1 is the major player in total activity. PARP-1 activates the enzyme ATM (essential for HR), inactivates pathways leading to NHEJ, acts as a nick sensor and, most importantly for its synthetic lethality, recruits DNA repair proteins to single strand breaks (SSBs) (34). As SSBs frequently occur spontaneously in each cell, inhibition of PARP consequently leads to an accumulation of SSBs (35). These SSBs are in turn converted into double strand breaks (DSBs) when encountered by the replication fork ultimately leading to cell death (17; 34; 23). In order to avoid this, DSBs are repaired by HR, a process in which BRCA gene products play a vital role (14; 33). Hence, as shown in mice, PARP-1 deficient cells can survive (35).

Familial breast cancer, is commonly associated with an inherited defect in either the BRCA1 or the BRCA2 gene. When the other BRCA gene (Breast cancer tumor suppressor) loses its function through random mutation, it leads to a cancerous state (35). Consequently, these tumor cells lacking in BRCA genes are sensitive to PARP inhibition. Healthy cells, on the other side, still have at least one functioning BRCA gene and are therefore not affected by the effects of PARP inhibitors (3; 32). Thus, BRCA and PARP constitute a synthetic lethality which can be exploited in BRCA-deficient cancer cells through small molecule PARP inhibitors (23; 17). Furthermore, as this concept does not solely rely on BRCA mutations but on insufficient HR, the use of PARP inhibitors can be extended to other cancer forms with impeded HR (32). This could be for example relevant for cancers associated with "BRCAness", referring to epigenetic silencing of BRCA, upregulated transcription suppressors of BRCA or defects in other genes involved in HR like ATM (23; 34). The knowledge-based approach of this synthetic lethality also enabled additional research such as Sullivan-Reed et al.'s proposal to target RAD52 simultaneously to PARP in a dual synthetic lethality concept as this protein can serve as a backup for BRCA in HR (33). Similarly, Mendes-Pereira et al. showed that PARP inhibitors can be applied for treatment of PTEN-deficient tumors, yet another tumor suppressor involved in HR (36).

4.2.4 Established paralog dependency ENO1 and ENO2

Enolases (2-phospho-D-glycerate hydrolases) are crucial metabolic enzymes, which catalyze the reaction of 2-phospho-D-glycerate (PGA) to phosphoenolpyruvate (PEP) in glycolysis as well as the reverse reaction during glycogenesis (37; 38). There are three isoenzymes in the human body: ENO1 (alpha-enolase), ENO2 (gamma-enolase) and ENO3 (beta-enolase). ENO1 is ubiquitously expressed, while ENO2

and ENO3 are tissue-specifically expressed in neurons or muscle, respectively (37; 38). The active form of enolase consist of two ENO subunits either as a homodimer or as a heterodimer in the corresponding tissues (37). With the exception of heat stability, the various enolase forms show similar properties (38). The ENO1 gene has a second, smaller product, called MBP-1 (MYC promotor binding protein) which binds to the tumor suppressor MYC (37; 39) and acts as a tumor suppressor by regulating growth and metastasis of cancer cells (40). The expression of enolases in cancer varies depending on the type of tumor (37) and sometimes even the clinical stage with reports of both overexpression (41; 42; 43; 44) and underexpression (45; 46).

In 1-5% of glioblastoma, however, the 1p36 locus is homozygously deleted (47). This locus contains several tumor suppressors (48; 49; 50) as well as the ENO1 gene. Consequently, ENO1 is lost in many of these tumor cells. Instead, they rely on ENO2 for their metabolism. This paralog dependency was confirmed via shRNA-mediated knockout of ENO2, which lead to profound inhibition of cell growth in cells without ENO1 while wild type cells were unaffected (47). ENO1-deficient cells have a 90% lower total enolase activity than wild-type cells. Therefore, an enolase inhibitor like PhAH, even though it does not distinguish between ENO1 and ENO2, shows significantly higher toxicity towards ENO1-deficient cells (47). As PhAH cannot be used for the treatment of patients with ENO1 null tumors due to poor cell permeability, new compounds like POMHEX have been developed which kills ENO1-deficient glioblastoma cells in the low nM range whereas normal cells can tolerate µM doses (51). Aside from glioblastoma, ENO1 homozygous deletions occur as well in hepatocellular carcinoma and cholangiocarcinoma (51).

4.3 Examined paralog dependencies in this study

4.3.1 RPP25 and RPP25L

RNase P is an endoribonuclease that processes precursor tRNA as well as other RNA substrates as for example long noncoding RNAs, rRNA and mRNA (52; 53; 54). Insufficient pre-tRNA processing and accumulation, caused by knockout of RPP30, a RNase P subunit, results in severe replication stress and triggers DNA damage response (55). This replication stress could be caused by conflicts between transcription and replication processes. Alternatively, precursor tRNA might form RNA:DNA hybrids. Additionally, RNase P is involved in chromatin regulation and HR and acts as a transcription factor for 5S rRNA gene (52). Via neofunctionalization processes, RNase MRP was evolved from RNase P (52). Both enzyme complexes share common protein subunits (56; 57). RNase MRP generates through endoribonucleolytic cleavage RNA primer for mitochondrial DNA replication (58; 59). Additionally, it was observed in *S. cerevisiae* that RNase MRP specifically cleaves mRNA of Clb2 protein, which is required for initiation and completion of mitosis. Possibly, levels of other mRNAs, most likely of they possess unusually long 5'-UTR like Clb2 mRNA, are controlled similarly (56). Both RNase P and RNase MRP participate in pre-rRNA processing (59).

RPP25 is one of at least 10 protein subunits of RNase P and RNase MRP (58; 53; 60; 54). Composition of these complexes is not constant but varies depending on the pending molecular task (52). Together with RPP20, RPP25 forms a strong 1:1 heterodimer, which subsequently binds to RNase P or RNase MRP via RPP25 (58; 61; 57). RPP25 is located in the nucleus and nucleosomes (60) together with RPP20, whereby RPP20 requires RPP25 for efficient nuclear entry and accumulation in the nucleoli (58). RPP25 and RPP20 belong to the Alba-like superfamily of protein, which are predicted to interact with both DNA and RNA (54; 61; 57). Therefore, this heterodimer might be involved in directing RNase P to chromatin of

target genes (54). RPP25 and RPP20 are part of RNase MRP pre-rRNA complexes, indicating that they dissociate before or during the assembly with preribosomes and reassociate after completed catalysis (58; 59). RPP20 alone is not able to bind to RNase MRP P3 domain contrary to RPP25, however, the strongest interaction comprises the RPP20-RPP25 heterodimer. Expression levels of RPP20 and RPP25 are correlated most likely regulated on protein level (58). Beside lowered RPP20 levels, RPP25 knockout led to inhibition of 5S rRNA gene transcription, while RNase P activity processing pre-tRNAs was unaffected. RPP25 deficient cells were not able to initiate 5S rRNA transcription correctly (53).

RPP25 has a paralog, RPP25L, with 43% amino acid sequence identity, which is assumed to share similar functions. In order to establish a paralog dependency on RPP25L, RPP25 does not necessarily have to be knocked out. Low expression levels of RPP25 through promotor hypermethylations, which occur in bladder, ovary, endometrium and glioma lineages, could suffice for a synthetic lethal interaction (62).

4.3.2   PRPS1 and PRPS2

Phosphoribosyl pyrophosphate synthetases, a family of enzymes essential to nucleotide synthesis, consist of three highly similar and conserved genes: PRPS1, PRPS2 and PRPS1L1 (63; 64). These enzymes catalyze the production of 5-phosphoribosyl 1-pyrophosphate (PRPP) from ATP and ribose-5-phosphate. PRPP is a central substrate for nucleotide synthesis, de novo synthesis of purines and pyrimidines as well as salvage synthesis of purine nucleotides (65; 66).

While PRPS1 and PRPS2 occur ubiquitously, the relative expression levels of PRPS variants are depending on the examined tissue (65; 67; 68; 64). Studies in rats showed that both, PRPS1 and PRPS2, are highly expressed in thymus, adipose tissues and testis while PRPS2 is more prominent in lung and spleen and PRPS1 more in brain and adrenal gland (67). PRPS1L1 on the other side is specifically expressed in testis (67). However, it might be redundant as PRPS1L1 KO mice showed no changes in viability or fertility when compared to WT mice (69). This tissue-differential expression is also true for PAP39, a PRPP synthetase associated protein, indicating that composition and, consequently, enzyme properties might differ in various tissues (65).

The more important paralogs PRPS1 and PRPS2 are both located on the X chromosome (66; 68). Although coding sequences of each gene are highly similar, differences are seen in promotor regions potentially causing tissue-differential expression (68). Cunningham et al. showed, that PRPS2, but not PRPS1 is translationally regulated (70). PRPS1, however, is transcriptionally regulated by microRNA-376 (71). The two paralogs PRPS1 and PRPS2 share 95% identity on amino acid level, nevertheless, differ in some physical and kinetic properties (64). PRPS2 is less stable than PRPS1 with regards to heat treatment and disaggregates more easily in the absence of protective levels of Mg^{2+} and ATP (65; 64). The two enzymes can be separated based on their isoelectric points and differ in optimal pH value for maximal activity (64).

PRPP synthetases are regulated by feedback inhibition through various nucleotides such as AMP, ADP or GDP (65; 64; 71). PRPS1 is significantly more sensitive to this kind of inhibition than PRSP2 (65; 64). This reduced feedback inhibition of PRPS2 is of particular interest when considering that PRPS2 is the more dominant isoform in rapidly proliferating cells (64). 2,3-Diphosphoglycerate (2,3-DPG), a regulatory molecule for hemoglobin oxygen affinity, inhibits both PRPS1 and PRPS2, although here PRPS2 is more sensitive. At higher concentrations of phosphate (5 mM) and lower concentrations of 2,3-DPG (1-10 mM) PRPS1 activity can even be stimulated (64). 4-amino-8-(beta-D-ribofuranosylamino)pyrimido-[5,4-d]-

pyrimidine-5'-monophosphate (APP-MP) acts as an inhibitor to PRPS1 and PRPS2, however, again, this inhibition is more potent towards PRPS1. In both cases IC_{50} values of APP-MP heavily depend on free phosphate concentration. APP-MP most likely acts as an allosteric inhibitor in the same way as ADP (72).

PRPS2 is frequently upregulated in colorectal cancer cells (73), melanoma (74) and prostate cancer and is linked to malignancy (75). In prostate cancer, PRPS2 silencing triggered cell cycle arrest and apoptosis (75). Cancer cells that emerge upon MYC oncogene hyperactivation rely on PRPS2 for de novo purine biosynthesis (70; 73; 74). In these cells, PRPS2 levels are elevated through upregulation of eIF4E, which is a translation initiation factor for PRPS2 (but not PRPS1) and a direct transcriptional target of MYC. As PRPS2 is not as much affected by feedback regulation, it may be more suitable for the rapidly growing cancer cells (70). In their study, Cunningham et al. describe a synthetically lethal interaction between MYC hyperactivation and PRPS2 knockout. Healthy cells were able to mitigate PRPS2 loss with PRPS1. Furthermore, they demonstrated that PRPS2 deficient mice are viable and fertile and did not show disadvantageous phenotypes. Even in spleen tissue, where in rats PRPS2 is the more prominent enzyme (67), no alterations were observed (70).

PRPS1, on the other side, has been assigned an important role in neuroblastoma (76) and glioblastoma (77), connecting PRPS1 expression with tumor malignancy. Knockout of PRPS1 led to lowered proliferation and apoptosis (76; 77). Furthermore, PRPS1 overexpression is correlated with and possibly regulated by MYCN, a MYC-related proto-oncogene (76). In acute lymphoblastic leukemia, PRPS1 mutations were observed in relapsed patients, which circumvent normal feedback inhibition of PRPS1 (78).

Beside their role in cancer, PRPS1 and PRPS2 are also associated with other medical conditions. Both gain-of-function and loss-of-function mutations of PRPS1 can result in various disorders. X-Linked Nonsyndromic Sensorineural Hearing Deafness, Charcot-Marie-Tooth Disease and Arts syndrome, leading to various degrees of hearing loss hypotonia and neuropathy,   are results of loss-of-function mutations of PRPS1 (71; 63). PRPS1 superactivity, however, also can manifest similar neurological symptoms (63; 71). This superactivity can be caused by either gain of function mutations or increased PRPS1 expression (63). PRPS2 depletion is associated with hypospermatogenesis and apoptosis of spermatogenic cells (79).

4.3.3   DNAJC15 and DNAJC19

DNAJC15 and DNAJC19 belong to the HSP40 (also: DNAJ) heat shock protein family (80). Heat shock protein expression is pronounced under stress conditions such as elevated temperature, oxygen deprivation, ultra-violet light or chemical agents which can cause protein denaturation or misfolding (81; 82; 83; 84; 85; 86). Nevertheless HSP are present in cells under normal conditions as well (81). They are separated in to groups based on molecular weight (83; 84; 85). HSP40 is a group of protein orthologous to *Escherichia coli*'s DnaJ protein, which are conserved throughout evolution and like other HSP serve important functions in protein translation, folding, unfolding, translocation and degradation (81; 84). Their most noted role is as a co-chaperone to HSP70 chaperones, stimulating HSP70's ATPase activity (81; 82; 87; 88). For this purpose, HSP40 proteins bind to HSP70s via the J-domain, the common denominator within the HSP40 group (80; 81). Based on their domain structures, HSP40 protein are split into three subclasses DNAJA, DNAJB and DNAJC (also DNAJ type I, II and III), of which DNAJC is the most loosely defined with only the J-domain occurring anywhere within the protein (81; 87). Unlike other DNAJ proteins, the J-domain of DNAJC15 is located at the C-terminal (81; 89).

DNAJ15 is highly expressed in testis, but occurs in all tissues (81; 90). In cancer cells, HSP expression is frequently altered (81; 85). Interestingly, HSP40 are involved in both pro- and anticancer processes (85). Extensive methylation of DNAJC15 (DNAJD1) of CpG islands leads to transcriptional silencing (91; 92; 93), hence DNAJC15 is also known as MCJ (methylation-controlled J-protein) (89). This hypermethylation was observed in Wilms tumors (94), ovarian tumors (91; 90) and primitive neuroectodermal tumor (PNTE) cell lines (91). Lindsey et al. noticed biallelic hypermethylation for DNAJC15 in PNTE cell lines, implying that this epigenetic silencing may suffice without the need of deletions (91). In patients with childhood neuroblastoma, the level of promotor methylation of DNAJC15 is significantly increased when MYCN is amplified and in patients that were older at time of diagnosis (95). However, DNAJC15 methylation and consequent lowered expression is not exclusive to cancerous tissues; for example it was also observed in normal ovarian samples (92).

Loss of DNAJC15 has been associated with increased resistance to chemotherapeutics like paclitaxel, topotecan, cisplatin and doxorubicin (81; 91; 89; 93; 96; 84; 90). In further consequence this chemoresistance can lead to treatment failure and recurrence. In the case of doxorubicin, direct induction of the transcription factor ETV7 upon drug treatment has been observed (Figure 2). ETV7 is a negative regulator of DNAJC15, repressing its expression through DNA methylation by recruiting DNA methyltransferase DNMT3A (93). Supporting the importance of methylation in DNAJC15 expression, inhibitors of methyltransferases can activate DNAJC15 expression in otherwise non-expressing cells (90; 93). DNAJC15 seems to be involved in proteasomal degradation of c-Jun, hence, DNAJC15 knockout causes higher levels of c-Jun. The transcription factor AP-1, consisting of a c-Jun homomer, upregulates the expression of the ABCB1 transporter protein. (89) This links the absence of DNAJC15 to an increased expression of ABCB1, which in turn most likely causes the enhanced drug resistance through excretion (80; 89; 96).
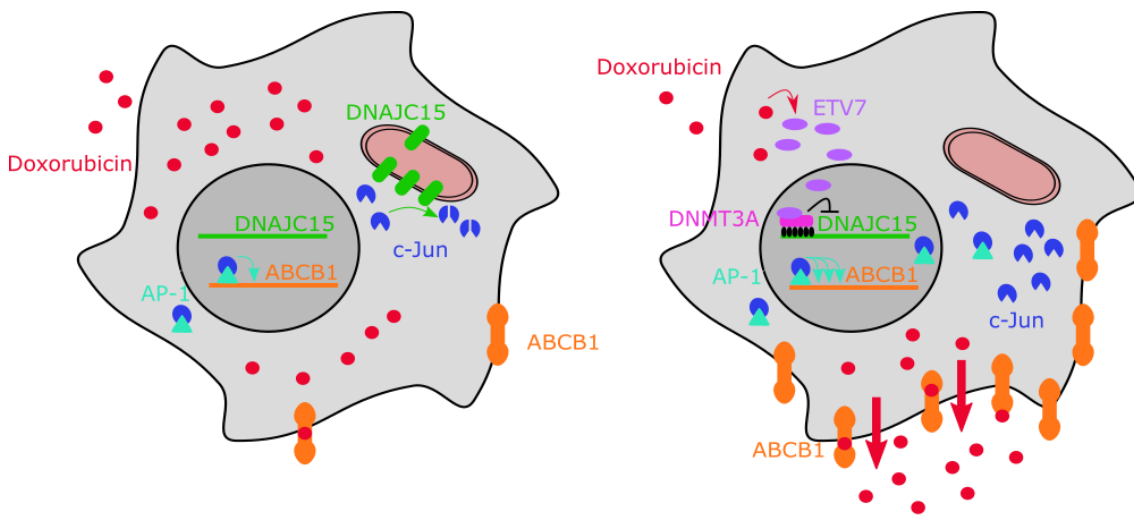

Figure 2: Chemotherapeutic resistance through DNAJC15 silencing.
Left: In a non-resistant cell, c-Jun levels, which is part of the transcription factor AP-1 regulating ABCB1 expression, are regulated trough DNAJC15. Right: Resistance to chemotherapeutics can emerge from upregulation of ETV7, which recruits DNMT3A to methylate DNAJC15. Consequently, c-Jun is unregulated overexpressing ABCB1, which then deposits chemotherapeutics like doxorubicin.

The sequences of both, DNAJC15 and DNAJC19, are similar to the TIM14 protein (Table 1) which plays a vital part a translocase of *S. cerevisiae* (80; 89). This ATP-dependent translocase TIM23 transports

precursor protein with an amino-terminal mitochondrial targeting signal across the inner mitochondrial membrane (80). In their study, Schusdziarra et al., demonstrate that DNAJC15 take on TIM14's role in the human TIM23 mitochondiral translocase (80). In MCF7 breast cancer cells, knockout of DNAJC15 led to reduced import of pre-protein into mitochondria, however, cell viability was retained most likely through DNAJC19 (80). Both DNAJC15 and DNAJC19 stimulate the ATPase activity of the mitochondrial HSP70 mortalin, another component of TIM23, to a similar degree (80). In either instance, this stimulating effect can be negated by MAGMAS, a regulatory protein for the ATPase activity of human import motors (80; 97; 82). These observations reinforce the hypothesis that this might be a paralog pair with overlapping functionality.

Table 1: Amino acid sequence similarity of DNAJC15, DNAJC19 and TIM14

	Amino acid sequence identity
DNAJC15 – DNAJC19	58%
DNAJC15 – TIM14	51%
DNAJC19 – TIM14	58%

However, there are structural differences between DNAJC15 and DNAJC19, with DNAJC19 lacking a N-terminal segment present in DNAJC15 whose functionality is not fully understood yet (80). Additionally, DNAJC19 mutations can lead to dilated cardiomyopathy and cerebellar ataxia, which cannot be rescued simply by DNAJC15 (82; 98).

One possible starting point for future drug development may be phenoxy-N-arylacetamides, which have been reported to inhibit HSP40 proteins (99). Unfortunately, attempts of developing marketable HSP targeting drug thus far failed due to adverse toxic effects (86). If DNAJC15/DNAJC19 turns out to be a paralog dependencies and specific inhibition is feasible, this toxicity might be avoided.

4.3.4 SLC25A28 and SLC25A37

Iron is both essential for all eukaryotic cells as well as toxic at high concentration. Therefore, iron levels in cytosol and various organelles are tightly regulated (100). In mitochondria, iron is required for the synthesis of heme and iron-sulfur (Fe-S) clusters, which are utilized either within the mitochondria themselves or exported for cytosolic and nuclear proteins (101; 100).

The genes SLC25A28 and SLC25A37 encode for the mitochondrial solute carrier proteins mitoferrin 2 and mitoferrin 1 (102), located at the inner mitochondrial membrane (100). Both mitoferrins are responsible for iron transport into mitochondria in non-erythroid cells (101). SLC25A28 is expressed ubiquitously, while SLC25A37 occurs predominantly in blood cells (103). In erythroid cells, Mfrn1 (SLC25A37) accumulates due to increased protein half-live (101) possibly by a Mfrn1 specific binding to the mitochondrial transporter Abcb10. Even though Mfrn1 and Mfrn2 share 70% amino acid sequence similarity, a 25 amino acid sequence unique to Mfrn1 enables this interaction (102). SLC25A37 frequently exhibits intron retention which may be a posttranscriptional control mechanism for this protein (104). Additionally, there are erythroid-specific regulatory cis-elements in the SLC25A37 promotor (105; 103).

Decreasing either SLC25A28 or SLC25A37 expression by RNAi in mammalian cells leads to reduced heme synthesis (SLC25A28: 58%, SLC25A37: 27%), simultaneous silencing even more severely (84%). Overexpression of the respective paralog restores heme synthesis to regular levels. In wild type cells,
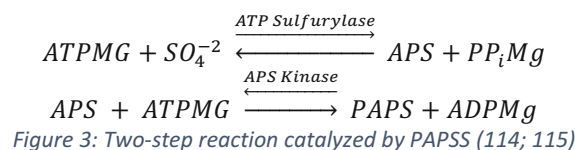
infection with overexpression vectors only led to a minor increase in heme synthesis as mitoferrins are regulated posttranslationally. Nevertheless, in non-erythroid cells one mitoferrin is sufficient for iron acquisition for heme synthesis. When mitoferrin is silenced, protoporphyrin accumulates in mitochondria and cytosolic iron concentrations increase. Downstream, further effects may occur like for example severely reduced activity of xanthine oxidase, a cytosolic enzyme containing Fe-S clusters. (101)

In a meta-analysis, reduced SLC25A37 expression in the brain was linked to major depressive disorder (106). In zebrafish, loss-of-function mutation of SLC25A37 causes severe anemia (105). Neuronal SLC25A37 knockout in mice causes decreased oxygen consumption and complex-I dependent electron transport, further leading to hippocampus-dependent spatial learning and memory deficiency (107). In invertebrates, only one mitoferrin exists, which is most likely a homolog to Mfrn2 (108; 109). In *D. melanogaster* mitoferrin is essential for spermatogenesis, a requirement which might similarly occur in vertebrates (108). Mitoferrin deficiency in *C. elegans* leads to reduced body size and fertility, slowed movements, increased sensitivity to paraquat, a toxin interfering with the electron transport chain, as well as, surprisingly, extended lifespan (109).

Li et al. investigated the mechanisms of PINK1 and PARK2, two tumor suppressor genes that, when either is lost, promote oncogenic Kras-driven pancreatic tumorigenesis in mice. They linked depletion of either gene with increased pancreatic iron levels, a higher ratio of mitochondrial iron to cytosolic iron as well as increased mitoferrin levels due to reduced protein degradation. The increased mitochondrial iron levels cause oxidative stress and lead to the Warburg effect, a metabolic abnormality common in cancer cells where glucose is converted to lactate despite the presence of oxygen, ultimately providing a rational for tumorigenesis. (110)

4.3.5   PAPSS1 and PAPSS2

Sulfation is one of the fundamental metabolic processes occurring in living cells. Sulfonated biomolecules include posttranslationally modified proteins, macromolecules such as glycosaminoglycan and proteoglycans, sulfolipides and small molecules like neurotransmitters and hormones. Many drugs and other xenobiotics are inactivated and released from the body via sulfation (111). In order to enable sulfation, inorganic sulfate is activated in the form of 3'-phosphoadenosine 5'-phosphosulfate (PAPS) (111; 112). In mammals, the reaction of ATP and sulfate to PAPS (Figure 3) is catalyzed by PAPS synthase, a bifunctional protein containing a C-terminal ATP sulfurylase and a N-terminal APS kinase activity, which is considered the rate-limiting step of PAPS formation (111; 113; 112).

$$ATPMG + SO_4^{-2} \xleftrightarrow[\text{APS Kinase}]{\text{ATP Sulfurylase}} APS + PP_iMg$$

$$APS \ + \ ATPMG \ \xrightarrow{} \ PAPS + ADPMg$$

Figure 3: Two-step reaction catalyzed by PAPSS (114; 115)

Two isoforms of PAPS synthases occur in humans: PAPSS1 and PAPSS2, sharing 78% amino acid identity (116; 112). PAPSS2 has a second splice variant, PAPSS2b, which only differs in an additional 5 amino acid sequence GMALP in the ATP sulfurylase domain (116). PAPSS1 is expressed ubiquitously, with higher levels in brain and skin (116), while PAPSS2 expression is more tissue-specific being primarily found in liver, cartilage and adrenal gland (117).

The PAPSS2 gene is located on chromosome 10 between MINPP1 (multiple inositol polyphosphate phosphatase 1) and PTEN (phosphatase and tensin homolog) in a highly conserved gene locus (116). This

is especially relevant, as PTEN is a known tumor suppressor (118). Thus, PTEN deletions may lead to bystander deletions of PAPSS2. PAPSS1 is placed on chromosome 4 (117). Neither PAPSS1 nor PAPSS2 are housekeeping genes, indicated by several promotor elements and the lack of a TATA box (116). PAPSS2, but not PAPSS1, is induced by the zinc finger protein Snail, a known as a inducer of tumor metastasis. Zhang et al. showed a connection between PAPSS2 upregulation through Snail and breast cancer cell migration, proliferation and invasion ability (119). Furthermore, TGF-β was identified to regulate PAPSS2 expression in mouse articular cartilage (120). As TGF-β signaling plays a central role in some cancers, both acting as a tumor suppressor and if mutated supporting tumor progression (121), levels of PAPSS2 may be altered as well.

PAPSS2 is crucial for skeletal and cartilage development (122; 112). PAPSS1 cannot compensate adequately for loss-of-function mutations of PAPSS2, observed in the context of skeletal malformations (122; 116) and androgen metabolism (123; 112; 124). Multiple possible causes have been postulated: PAPSS2 has been reported to have a significantly higher catalytic efficiency than PAPSS1 (117) which may result insufficient PAPS amounts in tissues with a high sulfation rate like adrenal glands and liver when PAPSS2 is mutated (123). However, when only the rate-limiting step of PAPS synthesis is examined, this difference vanished (112). Additionally, PAPSS1 is affected by APS substrate inhibition, involving 50 amino acids at the N-terminal (114). Apart from that, differences in subcellular localization were observed, with PAPSS1 being a predominantly nuclear enzyme while PAPSS2 mainly occurs in the cytoplasm, however, both enzymes are able to shuttle between nucleus and cytosol (113). For nuclear localization, an intact N-terminal KKxK motif is required, present in both isoforms (113). Furthermore, the two isoforms are reported to form a heterodimer when co-expressed, which may cause relocalization of PAPSS2 into the nucleus (125). In regards of PAPS synthesis for sulfation of dehydroepiandrosterone (DHEA), an androgen precursor, PAPSS2, is reported to electrostatically interact with the sulfontransferase SULT2A1, which catalyzes the subsequent reaction of PAPS with DHEA. This transient protein interaction is isoform-specific, which may explain some phenotypes associated with PAPSS2 mutations (112).

Leung et al. discovered, that PAPSS1 knockout sensitizes NSCLC cells for cisplatin treatment, which is part of the standard combination chemotherapy for advanced NSCLC patients (126; 127). This correlation could also be observed *in vivo* in zebrafish and mice (127). They observed low-level DNA damage associated with PAPSS1 knockdown and showed that the sensitizing effect can be extended to other DNA crosslinking agents like carboplatin, oxaliplatin, mitomycin C and some topoisomerase I inhibitors (126) as well as DNA-damaging radiation . However, cytotoxicity of mitotic and topoisomerase II inhibitors did not change upon PAPSS1 knockout, possibly due to the different type of DNA lesions caused by topoisomerase II inhibitors which are in turn repaired via different pathways (126). Under stress conditions, the sensitizing effect of PAPSS1 knockout is further increased (127). Thus, PAPSS1 may be involved in a certain DNA repair mechanism which when impeded leads to higher sensitivity towards corresponding chemotherapeutics (126). For this reason, development of a PAPSS1 specific inhibitor is even more promising, as not only a potential paralog dependency with PAPSS2 may be exploited for treatment, but it also could support conventional therapies (127). Chlorate is already reported as a PAPSS inhibitor blocking ATP sulfurylase activity, however, it does not distinguish between the two isoforms (126).

4.3.6 VPS4A and VPS4B

VPS4 is an AAA-ATPase associated with various cellular processes like intercellular protein transport, lysosomal degradation or cytokinesis, of which in mammals two paralogs available: VPS4A and VPS4B with 81% amino acid sequence identity (128; 129; 130). The genes for VPS4A and VPS4B are located on chromosome 16 and chromosome 18, respectively (128). In rats, each paralog are expressed in a tissue-specific way, both in terms of abundance and paralog ratio (128).

VPS4A and VPS4B occur in the cytosol (128). VPS4 variants are the only known nucleotide hydrolases in the ESCRT-III pathway (Endosomal Sorting Complex Required for Transport) (131). This complex allows membrane fission for various cellular processes like the generation of mulitvesicular bodies, cytokinetic abscission, nuclear envelope sealing, plasma membrane repair or microvesicle shedding (131). Here, VSP4 is required for ESCRT turnover at the midbody (131). A lack of both VPS4 results in abscission failure, demonstrated in HeLa cells (131).

Xu et al. observed high levels of VPS4B in the articular cartilage of an osteoarthritis rat model. VSPB4 was induced in the inflammatory microenvironment, which might be causal for coincident upregulation of apoptotic markers as IL-1β-induced apoptosis was partially alleviated upon VPS4B knockout (132). Similarly, upregulation of VPS4B combined with increased apoptosis was observed in intestinal epithelial cells of patients with Morbus Crohn, another inflammatory disease (133).

Upon middle cerebral artery occlusion, a model for ischemia, VPS4B is significantly upregulated in the adult rat brain and presumably plays a role in subsequent caspase-3 activated apoptosis (130). In breast cancer hypoxia-mediated degradation of VPS4B via the ubiquitin-proteasome system has been observed. This in turn impeded the degradation of activated EGFR, a known oncogene, leading to increased EGFR abundance and tumor aggressiveness. Congruently, lysosomal degradation of EGFR requires ESCRT protein complexes, in which VPS4 is involved as stated above (134). Further, in breast cancer cells VPS4B depletion leads to a metabolic shift downregulating glycolysis and fatty acid synthesis while fatty acid β-oxidation is significantly upregulated (135). In NSCLC, high expression of VPS4B is correlated with poor patient survival. Here, VPSB4's role in cell cycle progression pivotal and loss of VSPB4 delayed proliferation (136). Taken all together, the role of VPS4B in cancer remains debatable, as it is both involved in the degradation of known oncogenes and rapid proliferation.

In a recent study, Szymanska et al. described a paralog dependency between VPS4A and VPS4B. TCGA database analysis revealed a downregulation of VPS4B in various tumor types, most prominently in colorectal cancer. This loss of VPS4B in a multitude of cancer cells most likely represents a passenger alteration accompanying the loss of nearby tumor suppressors DCC, SMAD2 or SMAD4, also located on chromosome 18. Using RNAi-based loss-of-function, they demonstrated that both, VPS4A loss in VPS4B deficient cancer cells and simultaneous loss of VPS4A and VPS4B in other cancer cells, reduces cell viability via upregulation of inflammation and apoptotic pathways. Further, it was suggested that apoptosis occurs both via caspase-dependent and caspase-independent RIPK1 pathways and accompanied by the secretion of damage-associated molecular patterns which in turn are recognized by the adaptive immune system. Importantly, the sensitivity of VPS4B deficient tumor cells to RNAi-mediated VPS4A knockout could be confirmed in mouse models. (137)

5 Aim

The aim of this thesis is to validate or de-validate the putative paralog dependencies described above and to assess their potential for targeted cancer therapy. The putative paralog dependencies for this study were selected due to their strong correlation between low expression of the biomarker paralog and sensitivity to knocking out the target paralog. Publicly available CRISPR/Cas9 screening data (AVANA Screen) served as the starting point for a bioinformatics-based identification of putative paralog pairs (138).

In order to investigate the potential paralog dependency, cell lines with low expression of the biomarker paralog are selected and their sensitivity to CRISPR/Cas9 mediated knockout of the target paralog is tested. Subsequently, we explored whether overexpression of the biomarker paralog causes resistance to knockout of the target paralog to confirm the assumed paralog dependency.

In the assessment of the paralog dependency for targeted cancer therapy, multiple bioinformatics tools were used and developed. The aim was to investigate expression in healthy and tumor tissues as well as to explore whether the loss of biomarker expression can be explained through deletions od DNA methylations.

6 Material

6.1 Media

DMEM High Glucose with L-glutamine	Sigma-Aldrich, #D6429
DMEM / F-12 (1:1)	Gibco, #31330-038
EMEM	Sigma, #M5650
F-12 Nut Mix (1X) + GlutaMAX™-I	Gibco, #31765-027
IMDM	PAN Biotech, #P04-20150
McCoy's 5A (1X) + GlutaMAX™-I	Gibco, #36600-021
Opti-MEM® (1x) + GlutaMAX™-I	Gibco, #51985-026
RPMI-1640 Medium (ATCC Modification)	Gibco, #A10491-01
Waymouth's MB 752/1 Medium	Gibco, #31220-023
William's Medium E (1X) + GlutaMAX™-I	Gibco, #32551-020

6.2 Media supplements

Foetal Bovine Serum	HyClone #SH30071.03
GlutaMAX™-I (100x)	Gibco, #35050-061
Hydrocortisone	Sigma, #H0888
Insulin solution human	Sigma, #I9278
Sodium Pyruvate 100 mM (100X)	Gibco, #11360-039
Tet System Approved FBS	TAKARA, #631106

6.3 Antibiotics

Blasticidin	InVivoGen, #BLL-37-04A
Geneticin	Gibco, #2027021A
Hygromycin B Gold	InVivoGen, #HGG-37-06A
Penicilin/Streptomycin	Gibco, # 15140122
Puromycin	Sigma, #P9620

6.4 Cell lines

Cell line	Organ	Medium	FCS, %	Supplements	Split Rate
59M	ovary	DMEM	10	20 µg/mL Insulin	1:2 - 1:4, 1x / week
AsPC-1	pancreas	RPMI-1640	10		1:3 - 1:4
AU565	breast	RPMI-1640	10		1:4 - 1:6
CAL-12T	lung	DMEM	10		1:5 - 1:10
DK-MG	brain	RPMI-1640	10		1:2 - 1:4, 1x / week
DMS 53	lung	Waymouth's MB	10		1:2 - 1:4, 1x / week
H4	brain	DMEM	10		1:5 - 1:10
HCC1500	breast	RPMI-1640	10		1:3 - 1:5
Hep G2	liver	EMEM	10	1x Glutamax, 1x Pyruvate	1:5 - 1:10
HLF	liver	DMEM	10		1:5 - 1:10

HuH-6 Clone 5	liver	EMEM	10	1x Glutamax, 1x Pyruvate	1:5 - 1:8
JHH-2	liver	Williams'E Medium	10		1:2 - 1:4, 1x / week
JOPACA-1	pancreas	IMDM	20		1:3 - 1:5, 1x / week
KYM-1	lung	DMEM	10		1:5 - 1:10
KYSE-150	esophagus	RPMI-1640	10		1:8 - 1:15
KYSE-270	esophagus	50% RPMI-1640 + 50% F12K Nut Mix	10		1:2 - 1:3
293T-LentiX	kidney	DMEM	10	1x Pyruvate	1:8 - 1:10
NCI-H187	lung	RPMI-1640	10		1:2 - 1:5
NCI-H1975	lung	RPMI-1640	10		1:5 - 1:10
NCI-H2110	lung	RPMI-1640	10		1:2 - 1:4, 1x / week
NCI-H2170	lung	RPMI-1640	10		1:3 - 1:8
NCI-H716	colon	RPMI-1640	10		1:3 - 1:6
PANC 08.13	pancreas	RPMI-1640	15	10 µg/mL Insulin	1:2 - 1:4, 1x / week
Platinum E	kidney	DMEM	10	1 µg/mL Puromycin, 10 µg/mL Blasticidin	1:5 - 1:10
RT-112	bladder	RPMI-1640	10		1:5 - 1:10
SCC-25	oral cavity	DMEM:HAM's F12	10	400 ng/mL Hydrocortisone	1:3 - 1:6
SHP-77	lung	RPMI-1640	10		1:5 - 1:10
SK-BR-3	breast	DMEM	10	1x Pyruvate	1:3 - 1:6, 1x / week
SK-N-MC	brain	EMEM	10		1:3 - 1:5
SNU-761	liver	RPMI-1640	10		1:3 - 1:5
TT	thyroid	F12K Nut Mix	10		1:2, 1x / week
U-2 OS	bone	McCoy's 5A	10		1:2 - 1:4

6.5 Plasmids

All plasmids have been ordered and produced by Genscript.
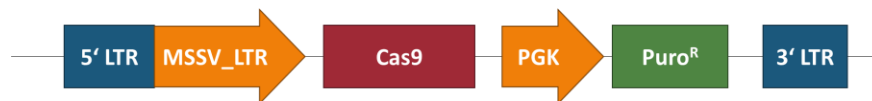
Figure 4: Simplified Cas9 vector MP110_Lenti_Cas9_puro_(cc60)
5' LTR: truncated 5' long terminal repeat, MSSV_LTR: Promotor, Cas9: Cas9 (Csn1) endonuclease from the Streptococcus pyogenes Type II CRISPR/Cas system with SV40 and nucleoplasmin nuclear localization sequences, PGK: promotor for mouse phosphoglycerate kinase 1, Puro^R: : puromycin N-acetyltransferase, 3' LTR: self-inactivating 3' long terminal repeat

Figure 5: Simplified RIEH vector for ecotropization 2nd_gen_virus_pRRL-RIEH
5' LTR: truncated 5' long terminal repeat, EF-1α: Promotor for human elongation factor EF-1α, rtTA3: Reverse tetracycline-dependent transactivator, IRES: Internal ribosomal entry site, EcoR: Ecotropic receptor Cationic amino acid transporter SLC7A1 (Cat-1), PGK: promotor for mouse phosphoglycerate kinase 1, Hygro^R: : Hygromycin phosphotransferase, 3' LTR: self-inactivating 3' long terminal repeat

19

*Figure 6: Simplified RT3REN vector for introduction of inducible overexpression constructs
dsRed has been removed from the vector*
5' LTR: *truncated 5' long terminal repeat*, **TRE3G:** *3rd-generation Tet-responsive promoter*, **GOI**: *Human codon-optimized
gene of interest*, **PGK**: *promotor for mouse phosphoglycerate kinase 1*, **Gen**[R] : *Aminoglycoside phosphotransferase (Geniticin
resistance)*, **3' LTR**: *self-inactivating 3' long terminal repeat*

The nucleotide sequence for the respective gene of interest (GOI) translated from the corresponding amino acid sequence (UniProt https://www.uniprot.org/) and human codon-optimized by Genscript.

ID	Overexpression construct (GOI)	Amino acid sequence
RN179	inducible RPP25	MENFRKVRSEEAPAGCGAEGGGPGSGPFADLAPGAVHMRVKEGSKIRNLMAFAT ASMAQPATRAIVFSGCGRATTKTVTCAEILKRRLAGLHQVTRLRYRSVREVWQS LPPGPTQGQTPGEPAASLSVLKNVPGLAILLSKDALDPRQPGYQPPNPHPGPSS PPAAPASKRSLGEPAAGEGSAKRSQPEPGVADEDQTA
RN183	inducible DNAJC15	MAARGVIAPVGESLRYAEYLQPSAKRPDADVDQQRLVRSLIAVGLGVAALAFAG RYAFRIWKPLEQVITETAKKISTPSFSSYYKGGFEQKMSRREAGLILGVSPSAG KAKIRTAHRRVMILNHPDKGGSPYVAAKINEAKDLLETTTKH
RN184	inducible DNAJC19	MASTVVAVGLTIAAAGFAGRYVLQAMKHMEPQVKQVFQSLPKSAFSGGYYRGGF EPKMTKREAALILGVSPTANKGKIRDAHRRIMLLNHPDKGGSPYIAAKINEAKD LLEGQAKK
RN186	inducible SLC25A28	MELEGRGAGGVAGGPAAGPGRSPGESALLDGWLQRGVGRGAGGGEAGACRPPVR QDPDSGPDYEALPAGATVTTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQPDPA ARYRNVLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYEKLKKTLSDV IHPGGNSHIANGAAGCVATLLHDAAMNPAEVVKQRMQMYNSPYHRVTDCVRAVW QNEGAGAFYRSYTTQLTMNVPFQAIHFMTYEFLQEHFNPQRRYNPSSHVLSGAC AGAVAAAATTPLDVCKTLLNTQESLALNSHITGHITGMASAFRTVYQVGGVTAY FRGVQARVIYQIPSTAIAWSVYEFFKYLITKRQEEWRAGK
RN187	inducible SLC25A37	MELRSGSVGSQAVARRMDGDSRDGGGGKDATGSEDYENLPTSASVSTHMTAGAM AGILEHSVMYPVDSVKTRMQSLSPDPKAQYTSIYGALKKIMRTEGFWRPLRGVN VMIMGAGPAHAMYFACYENMKRTLNDVFHHQGNSHLANGIAGSMATLLHDAVMN PAEVVKQRLQMYNSQHRSAISCIRTVWRTEGLGAFYRSYTTQLTMNIPFQSIHF ITYEFLQEQVNPHRTYNPQSHIISGGLAGALAAAATTPLDVCKTLLNTQENVAL SLANISGRLSGMANAFRTVYQLNGLAGYFKGIQARVIYQMPSTAISWSVYEFFK YFLTKRQLENRAPY
RN190	inducible PRPS1	MPNIKIFSGSSHQDLSQKIADRLGLELGKVVTKKFSNQETCVEIGESVRGEDVY IVQSGCGEINDNLMELLIMINACKIASASRVTAVIPCFPYARQDKKDKSRAPIS AKLVANMLSVAGADHIITMDLHASQIQGFFDIPVDNLYAEPAVLKWIRENISEW RNCTIVSPDAGGAKRVTSIADRLNVDFALIHKERKKANEVDRMVLVGDVKDRVA ILVDDMADTCGTICHAADKLLSAGATRVYAILTHGIFSGPAISRINNACFEAVV VTNTIPQEDKMKHCSKIQVIDISMILAEAIRRTHNGESVSYLFSHVPL

6.6   gRNAs

All gRNAs have been cloned into the following construct by Genscript:

Figure 7: Simplified vector for packaging gRNAs

5' LTR: *truncated 5' long terminal repeat,* **U6**: *RNA polymerase III promoter for human U6 snRNA,* **gRNA**: *guide RNA sequence targeting specific genes (see table below),* **tracrRNA**: *guide RNA scaffold for the Streptococcus pyogenes CRISPR/Cas9 system,* **EF-1α**: *Promotor for human elongation factor EF-1α,* **EGFP**: *mammalian codon-optimized enhanced GFP* **3' LTR**: *self-inactivating 3' long terminal repeat*

Target	ID	Amino acid gRNA start	gRNA sequence	Length, bp
PCNA (positive control)	gRNA_MP043	92	CGAAGATAACGCGGATACCT	20
POLR2A (positive control)	gRNA_MP046	491	GTACAATGCAGACTTTGACG	20
Negative Control	gRNA_MP049	-	GGCAGTCGTTCGGTTGATAT	20
Negative Control	gRNA_MP053	-	GATACACGAAGCATCACTAG	20
RPP25L	gRNA_RN290	26	ATCTCGGACCCGCATCTCAA	20
RPP25L	gRNA_RN291	73	ATTGTCAAGCGGCGGGTCCC	20
RPP25L	gRNA_RN292	52	GCTCGGCATGTAGTGTTCTC	20
RPP25L	gRNA_RN293	35	AATTCGCAACCTGCTGGGGT	20
RPP25L	gRNA_RN294	43	GCTCTGGGTCGGTTGGAGGG	20
DNAJC19	gRNA_RN318	63	AGCATTAATACTAGGTGTA	19
DNAJC19	gRNA_RN319	10	TTTAAATCATCCTGACAA	18
DNAJC19	gRNA_RN320	86	CTTTTAAATCATCCTGACAA	20
PAPSS1	gRNA_RN328	82	TCACCATCCAGAGTGTAGCA	20
PAPSS1	gRNA_RN329	141	AATGCAAGGCAAATTCATGA	20
PAPSS1	gRNA_RN330	333	TGCGAAGAATGGCCACACGG	20
PAPSS1	gRNA_RN331	277	TCTCATAAAGCCATTCAATG	20
PAPSS1	gRNA_RN332	495	GATGTATGCTGGACCAACTG	20
PAPSS1	gRNA_RN333	430	ATGGGTATCCTGCATTAACA	20
SLC25A28	gRNA_RN334	108	ACATTGCGATAGCGGGCAGC	20
SLC25A28	gRNA_RN335	90	GTACCCATCGACTGCGTCA	20
SLC25A28	gRNA_RN336	224	GGTGGTGTAGCTGCGGTAAA	20
SLC25A28	gRNA_RN337	206	GTACACAGTCTGTCACCCGG	20
SLC25A28	gRNA_RN338	320	GGGTGACCGCCTATTTCCGA	20
SLC25A28	gRNA_RN339	266	CTCCTGCGCAAGCTCCAGAG	20
PRPS1	gRNA_RN344	104	GCTTGGCTGAGATTGGCGCC	20
PRPS1	gRNA_RN345	82	GATGACTGCAGTAACCCGGC	20
PRPS1	gRNA_RN346	44	TGGTGAAAGTGTACGTGGAG	20
PRPS1	gRNA_RN347	254	GATGCGAGAAATAGCAGGAC	20
PRPS1	gRNA_RN348	212	GGATCGGGTGGCCATCCTTG	20
VPS4A	gRNA_RN361	64	AGACCGGGCCGAGAAGCTGA	20

VPS4A	gRNA_RN362	29	GCGGCTGTACCAGCATGCGG	20
VPS4A	gRNA_RN363	242	GGCCGCCCGGAGGATCAAAA	20
VPS4A	gRNA_RN364	189	TGAGGAGGACACAGAGAAGA	20
VPS4A	gRNA_RN365	417	GAGGTCGTCTGCATTCACCG	20
VPS4A	gRNA_RN366	394	GGCTCTAAGAGTTTGTCCCC	20

6.7 Antibodies

Primary Antibodies		
Rabbit anti-Actin	42 kDa	Sigma, A2066
Mouse anti-Actin	42 kDa	Sigma, A5441
Mouse anti-DNAJC15	16,5 kDa	Abnova, H00029103_B01P
Rabbit anti-DNAJC15	16,5 kDa	Novus Biologicals, NBP2-67439
Rabbit anti-DNAJC19	13 kDa	Protein Tech, 12096-1-AP
Rabbit anti-DNAJC19	13 kDa	Abcam, ab230187
Rabbit anti-PRPS2	35 kDa	Abcam, ab234886
Rabbit anti-RPP25	21 kDa	Abcam, ab229612
Rabbit anti-RPP25	21 kDa	Invitrogen, PA5-97315
Rabbit anti-RPP25	21 kDa	Thermo Fisher, A305-093A
Rabbit anti-SLC25A28	39 kDa	Novus Biologicals, NBP2-20390
Rabbit anti-SLC25A37	38 kDa	Novus Biologicals, NBP1-91570
Rabbit anti-SLC25A37	38 kDa	Abnova, PAB5900

Secondary Antibodies	
IRDye 680LT Goat anti-rabbit	Li-Cor, #926-68021
IRDye 800CW Goat anti-mouse	Li-Cor, #925-32210

6.8 Other chemicals and reagents

1,4 – Dithiothreitol	Sigma Aldrich, #3483-12-3
Accumax solution	SIGMA, #SLB70721
Bovine Serum Albumine	SIGMA, #SLBX7498
Cell culture grade water	CORNING, #30917005
CellTiter-Glo 2.0 Reagent	Promega, #0000378050
Cryostor	Biolife Solutions, #19092
Doxycycline	SIGMA, # 24390-14-5
DPBS (10x)	Gibco, #1880286
DPBS (1x)	Gibco, #2098592
Halt Protease Inhibitor Cocktail (100x)	ThermoScientific #UD281587A
Laemmli buffer (4x)	Bio-Rad, #161-0747
Lenti-X™ Packaging Single Shots (VSV-G)	Clontech, #631276
Lipofectamine LTX Reagent	InVitrogen, #1859370
Page Ruler Plus Prestained Protein Ladder	ThermoScientific, #00661116
Plus Reagent	InVitrogen, #1861755
Polybrene	Santa Cruz Biotechnology, # sc-134220

Quick Start Bradford 1x Dye Reagent	Bio-Rad, #500-0205
Rely+On™ Virkon® tablets	Lanxess, #27632
RIPA Buffer	Sigma, #R0278
Roti®-Block	Roth, #A151.3
TrypLE™ Express	Gibco, #12604-013
Trypsin EDTA	Gibco, #043-90317H
Tween 20 solution (10 %)	Bio-Rad, #161-0781
XT MOPS running buffer (20x)	Bio-Rad, #161-0788

6.9   Devices

Accuri C6 Flow Cytometer	BD Biosciences
Axiovert 100	Zeiss
BBD 6220 CO2-Incubator	ThermoFisher Scientific, #51020241
BD FACSVerse™	BD Bioscience, #651155
Centrifuge 5424 R	Eppendorf
Galaxy MiniStar microcentrifuge	VWR
Gel tray	Bio-Rad Laboratories, Inc.
HERAcell 240 CO2 incubator	ThermoFisher
LI-COR Odyssey imaging platform	LI-COR BioSciences
Mr. Frosty	Thermo Scientific, # 5100-0001
PG2002-S scale	Mettler Toledo
PowerPac HC power supply	Bio-Rad, #1645052
P-touch 9500pc	Brother
RCT basic	IKA
Thermomixer comfort	Eppendorf, #5355 000.011
Trans-Blot Turbo transfer system (semi-dry)	Bio-Rad, #170-4159
Ultrospec 3000 UV/Visible Spectrophotometer	Pharmacia Biotech, #80-2106-20
ViCell XR (Cell Viability Analyser)	Beckman Coulter, #731050
Victor X4 Multilable Plate Reader	PerkinElmer
Vortex Genie 2	Scientific Industries, #SI-0236
ZOE™ Fluorescent Cell Imager	Bio-Rad, #1450031

6.10 Consumables

1.5 mL cuvettes	BRAND, #10151710
1.8 mL Nunc cryotube vials	Thermo Scientific, #375418
12-well plate	Corning, #3336
25 cm² cell culture flask	Corning, #3289
4 mL sample cup (XR ViCell)	Beckman Coulter, #12458306
6-well plate	Corning, #3335
75 cm² cell culture flask	Corning, # 3290
96-well plate (clear)	Corning, # 243656
Cell Scraper	Corning, #3010

Criterion XT Precast Gel 4-12% 18 Well Comb, 30 µL	Bio-Rad, #3450124
Criterion XT Precast Gel 4-12% 26 Well Comb, 15 µL	Bio-Rad, #3450125
Eppendorf tubes® 1.5 mL	Eppendorf AG
Eppendorf tubes® 2 mL	Eppendorf AG
Eppendorf tubes® 5 mL	Eppendorf AG
Falcon 15mL High clarity PP Centrifuge tube	Corning, #352096
Falcon 175cm² Rectangular straight neck cell culture flask with vented cap	Corning, #353112
Falcon 50mL High Clarity PP Centrifuge Tube	Corning, #352070
Inject solo syringe	BRAUN, #4606736V
Pipette tips (10 µL – 1000 µL ART tips)	Thermo Scientific
Pipettes (P10 – P1000)	Eppendorf AG
Round-bottom polystrene tube	Falcon , #352052
Surgical disposable scalpels	B.Braun, #5518083
Syringe filter 0.22 µM	SIGMA, #SLGVS25LS
Trans-Blot Turbo Transfer, Midi Format, 0.2 µm Nitrocellulose	Bio-Rad, #1704159
Trans-Blot Turbo Transfer, Midi Format, 0.2 µm PVDF	Bio-Rad, #1704157

6.11 Software

BD Accuri C6 Flow cytometry software	BD cytometers
BD FACSuite	BD, version 1.0.5.3841
Image Studio	LI-COR Biosciences, version 2.0.38
RStudio, Version 1.2.5019	RStudio, Inc.
Ordino, Version 8.0.0	Boehringer Ingelheim RCV GmbH & Co KG

24

7 Methods

7.1 Cell Culture

7.1.1 Thawing

Cells were stored in Cryostor at -80°C. For cultivation, cells were thawed at room temperature and transferred into a 10 mL of the corresponding medium. Next, they were centrifuged (1200 rpm, 5 minutes) and the supernatant was discarded to remove remaining Cryostor. Finally, cells were resuspendend in medium and transferred into cell culture flasks.

7.1.2 Cultivation and splitting

All cells were cultivated at 37°C and 5% CO_2. In order to maintain high cell viability, cultured cells were split twice a week if not stated otherwise. Medium was removed from the cells and adherent cells were detached using trypsin. For very sensitive cell lines (SHP-77 and RT-112) the more gentle reagent Tryple Express was used instead of trypsin. After approximately 10 minutes incubation at 37°C, the single cells were suspended in medium, stopping the lysis reaction. According to the split ratio (e.g. 1:10), a fraction of the cells were transferred into a new cell culture flask and filled up with fresh medium. Finally, required antibiotics were added.

7.1.3 Seeding

For experiments, residual harvested cells after splitting were seeded into the corresponding plates. For this purpose 500 µL cell suspension were used for Vi-CELL™ cell counting to determine the number of viable cells per milliliter. This way, a certain known cell amount could be applied in each experiment. Fresh Medium was added to reach the desired volume. For experiments, no antibiotics were added to the medium.

7.1.4 Freezing

In order to store cell lines over extended periods of time, cells were detached and diluted with medium to stop typsinization. Cells were harvested by centrifugation (1200 rpm, 5 minutes). The supernatant was removed and the cells were resuspended in CryoStor freeze medium. Aliquots of 1 mL were transferred into Cryotubes and put into Mr. Frosty freezing containers filled with 2-propanol. This way, cells were frozen gently at -80°C. After at least two days, aliquots were transferred into boxes for storage.

7.2 Generation of Stable Cell Lines and Viruses

For safety reasons, all steps involving virus particles were carried out in a separate S2+ laboratory. Liquids containing viruses as well as all consumables that were possibly contaminated were inactivated with 5% Virkon prior to discarding in specific waste bags.

7.2.1 Virus production

In order to generate lentivirus for introducing gRNAs to the target cells, $4 \cdot 10^6$ 293T-LentiX cells were seeded in 10 cm petri dishes in 8 mL medium and incubated overnight. The next day, 7 µg gRNA were diluted with water to a total volume of 600 µL, added to LentiX packaging single shot (VSV-G) tubes and vortexed until the pellet was dissolved. After 10 min incubation at room temperature to allow nanoparticle complexes to form, samples were spun down and the entire 600 µL solution were added dropwise to the cultivated 293T-LentiX. After 4 h incubation the cells were supplemented with additional 6 mL medium, and for another 48°C incubated in the S2+ laboratory. The formed pantropic

viral particles were harvested by collecting the medium and passing it through a 0.45 µM PVDF filter. Aliquots of 1 mL were stored at -80°C.

7.2.2 Introduction of Cas9

For CRISPR depletion assays, stable Cas9 expression is required, hence it was introduced into the cells. For this purpose, 400 000 cells were seeded in 500 µL medium into two wells of a 6-well-plate and equipped with 0.8 µL polybrene. In the S2+ lab, 500 µL MP110_Lenti_Cas9_puro_(cc60) virus solution was added to one well, the other well was supplemented with 500 µL medium as a control. The Cas9 virus production was done by colleagues at Boehringer Ingelheim RCV. After incubation overnight, the medium was discarded, cells were washed with 1 mL PBS, 1 mL fresh medium was added and cells were incubated for another 48 h. Then again, medium was removed and cells were washed with 1 mL PBS to remove any remaining virus. This time, the added 2 mL fresh medium contained a lethal concentration of puromycin, which was beforehand determined. Therefore, cells which were not successfully transduced as well as the control cells died during subsequent incubation, while cells which could integrate the insert survived.

7.2.3 Ecotropization of cell lines

Cell lines proven to be sensitive to knocking out the biomarker paralog in the CRISPR depletion assays were transduced with the RIEH construct (MSCV-rtTA3-IRES-EcoR-PGK-Hygro^r). This was done for two reasons: EcoR, also known as CAT-1 (High affinity cationic amino acid transporter 1), is the receptor for the ecotropic murine retroviral leukemia virus found in rats. Under normal conditions ecotropic viruses cannot infect human cells due to the lack of corresponding receptors. Through ecotropization the target cell line becomes susceptible to these viruses. This system was used to introduce the overexpression constructs via ecotropic viruses as it highly reduces the risk of self-infection and thereby increases work-safety. Additionally, the reverse tetracycline-dependent transactivator rtTA3 is required for the doxycycline induced expression of the overexpression constructs via the RT3REN plasmid (139). When rtTA3 binds to doxycycline, it activates genes under control of the tet-responsive promotor.

Similarly to the introduction of Cas9, 400 000 cells were seeded in 500 µL medium in a-6-plate and 0.8 µL polybrene were added. One well was transduced with 500 µL 2nd_gen_virus_pRRL-RIEH virus solution while one well was used as a control. The RIEH virus was packaged by colleagues at Boehringer Ingelheim RCV. After incubation overnight, cells were washed with PBS, supplemented with 1 mL fresh medium and incubated for further two days. Cells were washed again with PBS and 2 mL medium containing a lethal concentration of hygromycin were added. This way, only transduced cells survived.

7.2.4 Introduction of overexpression constructs

For retroviral packaging, 800 000 Platinum- E cells were plated in 6-well-plates and incubated for 24 hours. The Platinum-E cell line was created from the cell line 293T and expresses viral envelope, matrix and enzyme proteins for ecotropic helper virus production (140). For each transduction, two reagent mixtures were prepared: (1) 30 µL serum-free OptiMem medium mixed with 6.4 µL LTX reagent and (2) 30 µL serum-free OptiMem medium mixed 2 µg target plasmid DNA and 2.6 µL DNA Plus reagent. Both mixtures are combined, incubated for 5 minutes at room temperature and added to the Platinum-E cells. The next day, 200 000 cells per well of the target cell line were seeded in a 6-well-plate in Tet-free medium. Per transduction one well is required as well as an additional well for the selection control. Furthermore, the medium of the transduced Platinum-E cells is changed to 2 mL medium of the target cells. After another 24 hours, the supernatant of the transduced Platinum-E cells is harvested and

passed through 0.45 µM filter. Medium is removed from the target cells and replaced with the freshly harvested virus solution. To each well 8 µg/mL polybrene is added to aid transduction. The next day, medium is exchanged and a lethal concentration geneticin is added to select for successfully transduced cells.

7.3 Assays and Analyses

7.3.1 CRISPR/Cas9 depletion assay

CRISPR/Cas9 depletion assay were used to investigate sensitivity of cell lines to knockout of certain proteins. For this purpose, 200 000 cells were seeded in 500 µL medium in a 12-well-plate. One well per assayed gRNA was prepared and 500 µL gRNA virus and polybrene to a final concentration of 8 µg/mL were added. Alongside gRNAs targeting the queried protein, two positive controls targeting the essential genes PCNA and POLR2A and two nontargeting negative controls were used. One well was not transduced. After incubation overnight, the medium was removed, cells were washed with 1 mL PBS and fresh medium was added. After another 48 h, three days after transduction, cells were detached with 500 µL AccuMax. Of that, 200 µL were transferred into a new plate containing 1 mL medium. The remaining cell suspension was analyzed with flow cytometry. The non-transduced sample was used to gate viable cells and set the threshold for background fluorescence. This way, transduced viable cells can be identified due to their fluorescence from the GFP which is co-expressed with the gRNA. Flow cytometry was performed at a flow rate of 66 µL/min and at least 5000 viable cells are registered per measurement. The initially measured transduction rate on day 3 is set as a reference point. In the following two weeks, cells were split and analyzed twice a week, monitoring percentage of transduced cells. When the gRNA targets an essential protein, the amount of transduced cells should decrease over time similarly to the positive control. On the other hand, if the gRNA disrupts a nonessential gene, the ratio of transduced cells should remain constant.

Once sensitivity of a cell line to the KO of the target paralog was confirmed and inducible overexpression (OE) constructs were introduced, the CRISPR depletion assay was repeated with and without the addition of doxycycline. If the transduced cells no longer deplete when the OE construct is induced, it can be concluded that the OE construct rescued the cells from the negative effects of the KO of the target paralog.

7.3.2 CellTiter-Glo® assay

The CellTiter-Glo® assay uses a thermostable luciferase to measure ATP concentrations after cell lysis. This metabolic assay can be utilized as proxy for overall biological activity and cell viability. After cultivation CellTiter-Glo® reagent is added to the cells to a final ratio of 1:4 reagent to medium. After mixing and 10 minutes incubation in the dark at room temperature, luminescence is recorded using the multilabel Plate Reader VICTOR X4 with a measurement time of 0.2 seconds.

For antibiotic kill curves, 5000 cells were seeded in 140 µL medium per well in a 96-well plate and incubated for 4 h to give the cells time to settle. In the meantime, a 1:2 dilution series of the antibiotic was prepared (Table 2). In each well, 10 µL of diluted antibiotics were added. After cultivation for 4 to 7 days, CellTiter Glo® assay was performed. Measured luminescence values were normalized to untreated cells.

Table 2: Antibiotic Concentrations for Kill Curves

Antibiotic	Starting Concentration Dilution Series; µg/mL	Final Concentration on the Cells; µg/ml							
Puromycin	300	20	10	5	2.5	1.25	0.625	0.313	0
Hygromycin	12000	800	400	200	100	50	25	12.5	0
Geneticin	12000	800	400	200	100	50	25	12.5	0

7.3.3 Cell number measurements

Vi-CELL™ was used to perform cell number measurements. For this purpose, 500 µL cell suspension were transferred into a Vi-CELL™ sample cup and put into the Vi-CELL™ sampler. Real-time cellular imaging is applied to determine cell concentrations, distinguishing viable from dead cells using the trypan blue vital dye method (141), which colors cells with leaky membranes (142).

7.3.4 Western blot

Western blots were used to examine protein expression of the investigated paralogs. To this end, 400 000 cells were seeded in 6-well-plates in Tet-free medium. When the induction of overexpression constructs was investigated, doxycyline was added to the cells. After cultivation at 37°C and 5% CO_2 for 3 – 4 days until confluence is reached, the cells are washed with PBS, put on ice and lysed for 5 minutes in 100 µL RIPA buffer supplemented with HALT protease/phosphatase inhibitor cocktail. All subsequent steps were performed on ice to prevent protein degradation. The samples were scraped, transferred into Eppendorf tubes, sonicated at 35 Amp for 40 seconds and stored at -80°C.

For western blotting, samples were thawed on ice and centrifuged (13 minutes, 4°C and 13 000 rpm). The total protein concentration was determined using the Bradford assay. Here, 2 µL of the supernatant is mixed with 1 mL Bradford reagent, incubated for 5 minutes at room temperature and the absorbance is measured at 595 nm using Bradford reagent without added samples as a blank. The protein concentration is calculated using a BSA calibration curve. The samples are diluted with lysis buffer and 4x laemmli buffer supplemented with 0.05 M DTT so that the same amount of total protein (e.g. 20 µg) is loaded in the loading volume. After incubation for 5 minutes at 95°C samples are loaded onto the electrophoresis gels. For the electrophoresis, MOPS running buffer and PageRuler prestained protein ladder were used. During the first 10 minutes the voltage is set to 80 V in order to pack the gel and then is increased to 130 V for the remainder of the runtime (approximately 1 h).

In order to optimize western blotting for the various investigated proteins, testing western under different conditions blots were conducted using samples expressing and not-expressing the inquired protein. The consequently selected western blot conditions are summarized in Table 3.

The electrophoresis gel is blotted onto the membrane (2.5 A, 7 min). The membrane was treated with blocking buffer for at least one hour before being incubated with first antibodies in the first antibody solution overnight at 4°C. Additionally to the antibody targeting the questioned protein an anti-actin-antibody is added to use actin as a loading control. The next day, the membrane was washed three times with PBST for 15 minutes and incubated with a solution containing the two 1:10000 diluted secondary antibodies for 1 h at room temperature. After washing the membrane again three times for 15 minutes with PBST, the membrane was scanned using the LI-COR measuring the fluorescence of the secondary antibodies at 800 and 700 nm.

Table 3: Western Blot conditions

Antigen	First Antibody	1st antibody dilution	Membrane	Blocking Buffer	1st antibody solution	2nd antibody solution
DNAJC15	H00029103_B01P Mouse polyclonal	1:500	PVDF	Roti-Block	Roti-Block	Roti-Block
RPP25	A350-93A Rabbit polyclonal	1:2000	Nitrocellulose	10% BSA, 10% PBST in H_2O	5% BSA in PBST	0.1% Casein in PBST
PRPS2	ab234886 Rabbit polyclonal	1:500	Nitrocellulose	10% BSA, 10% PBST in H_2O	5% BSA in PBST	0.1% Casein in PBST
DNAJC19	12096-1-AP Rabbit polyclonal	1:500	Nitrocellulose	Roti-Block	Roti-Block	Roti-Block
SLC25A37	NBP1-91570 Rabbit polyclonal	1:1000	Nitrocellulose	10% BSA, 10% PBST in H_2O	5% BSA in PBST	0.1% Casein in PBST
SLC25A28	NBP2-20390 Rabbit polyclonal	1:500	PVDF	10% BSA, 10% PBST in H_2O	5% BSA in PBST	0.1% Casein in PBST

7.4 Bioinformatics

7.4.1 Amino acid sequence identity

In order to investigate sequence identity, the amino acid sequence was extracted from Uniprot protein knowledgebase (https://www.uniprot.org/) (143). Next, sequences of two paralogs were aligned using the pBLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi) of the National Center of Biotechnology Information NCBI (144).

7.4.2 Cell line selection

Cells for experiments were selected upon following criteria: Firstly, cell lines must be either sensitive or resistant to CRISPR-Cas9 mediated knockout of one paralog according to AVANA CERES screens. In this screens a sgRNA library was applied to analyze various cancer cell lines in regard to their susceptibility to specific gene knockouts. The resulting CERES scores represent cell proliferation capabilities upon depletion of a certain gene computationally decoupled from gene copy number effects. Thereby, CERES score are scaled from 0 to -1, corresponding to knockouts auf nonessential and common core essential genes respectively. Without decoupling depletion results from gene copy number, genes with higher copy numbers would be overrepresented among detrimental knockout not because of high essentiality of the corresponding gene product but due to multiple CRISPR-Cas9 induced DNA double strand breaks leading to gene-independent DNA damage responses and G2 cell-cycle arrest. (138) Cell lines with CERES scores below -0.5 or above 0.15 were considered sensitive or resistant, respectively. Secondly, normalized gene expression of the corresponding paralog was considered using TPM-values (Transcript per million), which describe the abundance of certain mRNAs in the cultured cell lines (145). As a paralog dependency was assumed, cell lines with low TPM-values (ideally 1) for the biomarker paralog and higher TPM-values for the target paralog were preferred. At the same time, the targeted paralog should be expressed at least at a basal level (TPM > 10) to avoid selecting a false positive sensitive cell line. For resistant cell lines, both paralogs should be expressed. Finally, factors like in-house cell line availability and growth type, preferring adherent cells, were considered in order to simplify and accelerate experiments.

7.4.3 Gene expression in normal tissue and in tumor tissue

For utilizing a paralog dependency for targeted cancer therapy, a lack of biomarker gene expression is required while the cancer cells rely on the other paralog. To investigate, whether this is the case in

cancer patient samples, the expression of each paralog pair in various tumor types was analysed using the **c**ancer **g**enome **a**tlas (TCGA Research Network: https://www.cancer.gov/tcga).

Furthermore, the expression of each paralog in healthy tissues was investigated to assess possible risks of therapies exploiting one of the found paralog dependencies. When the biomarker gene is low expressed not only in cancer but also in some normal tissues, targeted knockout of the other paralog might cause adverse effects in the corresponding organs. Conversely, if the biomarker is the prominent version of this gene throughout the body while the targeted gene is sparsely expressed in healthy tissues, toxicity towards normal cells appears less likely. For this purpose, each investigated gene was queried using the GTex (Genotype-tissue expression) Portal (https://GTexportal.org/) (146). This way, TPM-values of each paralog gene across various human tissues were acquired.

7.4.4 Prevalence of deep deletions in tumor cells

Genetic deletions are one plausible explanation for the loss of biomarker paralog expression. The prevalence of deep, possibly homozygous, deletions of paralogs was investigated using the cBioPortal (https://www.cbioportal.org/) (147; 148). For this purpose, TCGA PanCancer Atlas Studies (TCGA Research Network: https://www.cancer.gov/tcga) were selected and queried by each gene. Using the GISTIC algorithm (Genomic Identification of Significant Targets in Cancer), tumor relevant copy number alterations were identified (149). Using a stringent cutoff of -2 for identifying putative deep deletions, the relative deep deletion frequencies for each gene in each tumor type as well as in all tumor types combined was calculated. In order to evaluate this data further, the deep deletion frequencies of each gene was plotted against their chromosomal starting position. Furthermore, genes deemed essential in at least 8 out of 12 cell lines examined essentiality studies (150; 151; 152) and known tumor suppressor genes collected by Zhao et al. (https://bioinfo.uth.edu/TSGene/) (153) were highlighted.

7.4.5 DNA methylation

Another reason for the lack of biomarker expression may be silencing through DNA methylation. In order to investigate the relationship between DNA methylation and gene expression the Shiny Methylation Analysis Resource Tool (SMART) has been employed (http://www.bioinfo-zs.com/smartapp/) (154). With this application, biomarker paralog expression and methylation of CpG islands of the corresponding gene of TCGA samples was investigated. Additionally, the correlation between these two variables was investigated using the Spearman's rank correlation test.

8 Results

8.1 Experimental workflow and cell line production

The same overall workflow was used to investigate the potential paralog dependencies (Figure 8). For each target paralog, two sensitive and two resistant cell lines were selected and Cas9_Puro was introduced. The puromycin, hygromycin and geneticin concentrations necessary for selection were determined with antibiotic kill curves (Table 4). After the sensitivity was confirmed in CRISPR depletion assays, ecotropization via the RIEH construct was conducted. This way, overexpression constructs of the biomarker paralog or the target paralog were introduced and selected using geneticin.
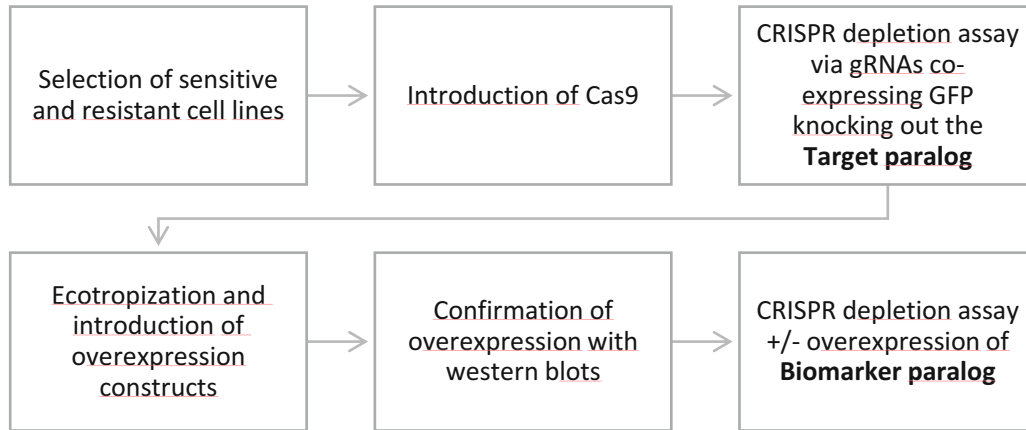
Figure 8: Experimental setup for verification of a paralog dependency.

Table 4: Concentration of Antibiotics for Selection

Cell line	Biomarker / Target	Puromycin µg/mL	Hygromycin µg/mL	Geneticin µg/mL
AU565	PRPS2 / PRPS1	1.5	NA	NA
JOPACA-1	PRPS2 / PRPS1	1	400	800
SHP-77	PRPS2 / PRPS1	3	NA	NA
JHH-2	PRPS2 / PRPS1	1.5	NA	NA
DK-MG	PRPS2 / PRPS1	1	NA	NA
SK-BR-3	PRPS2 / PRPS1 VPS4B / VPS4A	4	800	800
NCI-H2110	VPS4B / VPS4A	3	100	100
59M	VPS4B / VPS4A	10	400	800
DMS 53	VPS4B / VPS4A	5	NA	NA
KYSE-150	RPP25 / RPP25L	0.25	400	800
U-2 OS	RPP25 / RPP25L	1	100	200
SK-N-MC	RPP25 / RPP25L	1	NA	NA
NCI-H2170	RPP25 / RPP25L	1	NA	NA
CAL-12T	DNAJC15 / DNAJC19	1	200	400
NCI-H1975	DNAJC15 / DNAJC19	1	100	400
RT-112	DNAJC15 / DNAJC19	1	NA	NA
SCC-25	DNAJC15 / DNAJC19	1	NA	NA
H4	PAPSS2 / PAPSS1	1	NA	NA
AsPC-1	PAPSS2 / PAPSS1	3	NA	NA
HLF	PAPSS2 / PAPSS1	2	NA	NA
Hep G2	PAPSS2 / PAPSS1 SLC25A37 / SLC25A28	4	1000	2000
HuH-6	SLC25A37 / SLC25A28	2	400	800
SNU-761	SLC25A37 / SLC25A28	5	100	400
NCI-H716	SLC25A37 / SLC25A28	1	NA	NA
KYSE-270	SLC25A37 / SLC25A28	4	NA	NA

8.2 DNAJC15 and DNAJC19

Using the AVANA CRISPR screening data, the cell lines were selected by their CERES scores, which quantify their sensitivity to DNAJC15 or DNAJC19 knockout respectively (Screen). For both genes constitute context dependencies. This refers to a dependency in a subset of cancer cell lines as opposed to pan-essential or never-essential genes. Therefore, the essentiality of these genes is observable in a subset of cell lines and is determined by a the specific (epi-)genetic background. The color of the individual data points corresponds to the expression of the putative paralog gene. This way, DNAJC19 sensitivity was associated with low expression of DNAJC15. Hence, DNAJC15 was termed the *biomarker* paralog for the *target* paralog DNACJ19.
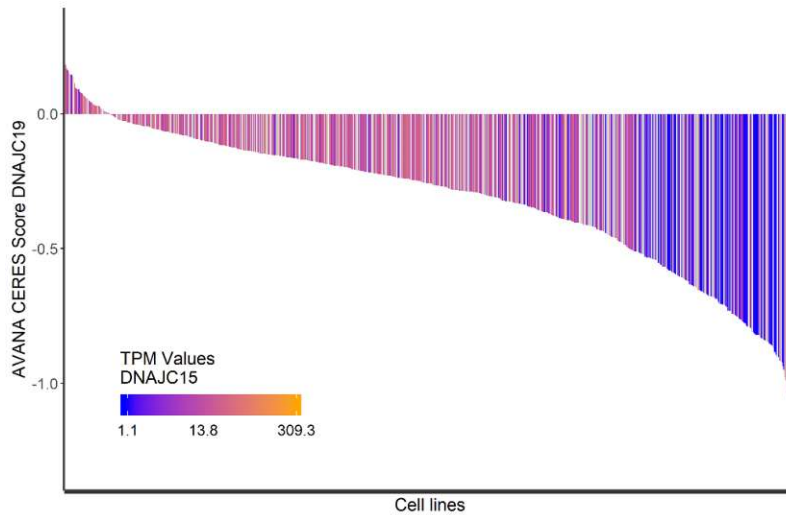
Figure 9: Sensitivity of cancer cell lines to DNAJC19 knockout (AVANA CERES screen results)

Using gene expression data in TCGA samples from **t**he **c**ancer **g**enome **a**tlas (TCGA Research Network: https://www.cancer.gov/tcga), we were able to identify tumors which do not express DNAJC15. Namely, DNAJC15 expression is absent in some prostate, uterus, ovary, lung and bladder cancers (Figure 10a). These patients might benefit from a therapy targeting DNAJC19.

Next, we examined whether the lack of DNAJC15 in tumors can be explained by a genetic lesion. According to the GISTIC results, DNAJC15 is deep (homozygous) deleted in 1.84% from all TCGA samples (https://www.cbioportal.org/). For specific cancer types, this number is even higher e.g. 16.6% in prostate cancer (PRAC), 4.66% in bladder cancer (BLCA) or 4.17% in lymphoma (DLBC). DNAJC15 resides in relatively close proximity to the chromosomal position of RB1, a tumor suppressor which is frequently deleted (Figure 10b). Thus, DNAJC15 deletion may occur as a bystander deletion to RB1. Bioinformatics-based in depth analysis of the genomic locus casts doubt on the reported deep deletion frequencies as the core-essential gene TPT1 is located between DNAJC15 and RB1. The GISTIC algorithm (Genomic Identification of Significant Targets in Cancer) predicts that the TPT1 gene is deep deleted in 1.87% of TCGA samples as well. This however seems highly implausible: If TPT1 is truly essential, cells lacking this gene would not be expected to survive. Conversely, two other tumor suppressors are located in proximity to DNAJC15: FOXO1 and TSC22D1. Deep deletions of these genes are less frequent than RB1, but could cause a passenger deletion of DNAJC15 as no essential genes are located in-between.

As DNAJC15 silencing through DNA methylation had been reported multiple times in literature (80; 89; 92; 94; 95), we analyzed whether low expression of DNAJC15 in tumor samples can be explained by this

mechanism (Figure 10c). Indeed, a strong correlation between DNAJC15 transcripts per million values (TPM) and the degree of DNA methylation on the CpG islands was observed in TCGA samples. Hardly any tumor samples were identified in which DNAJC15 was not methylated and not expressed. Hence, it appears that silencing through DNA methylation is the primary reason for DNAJC15 deficiency in cancer. Consistently, in the cell lines sensitive to DNAJC19 KO selected for laboratory experiments, DNAJC15 is also methylated.
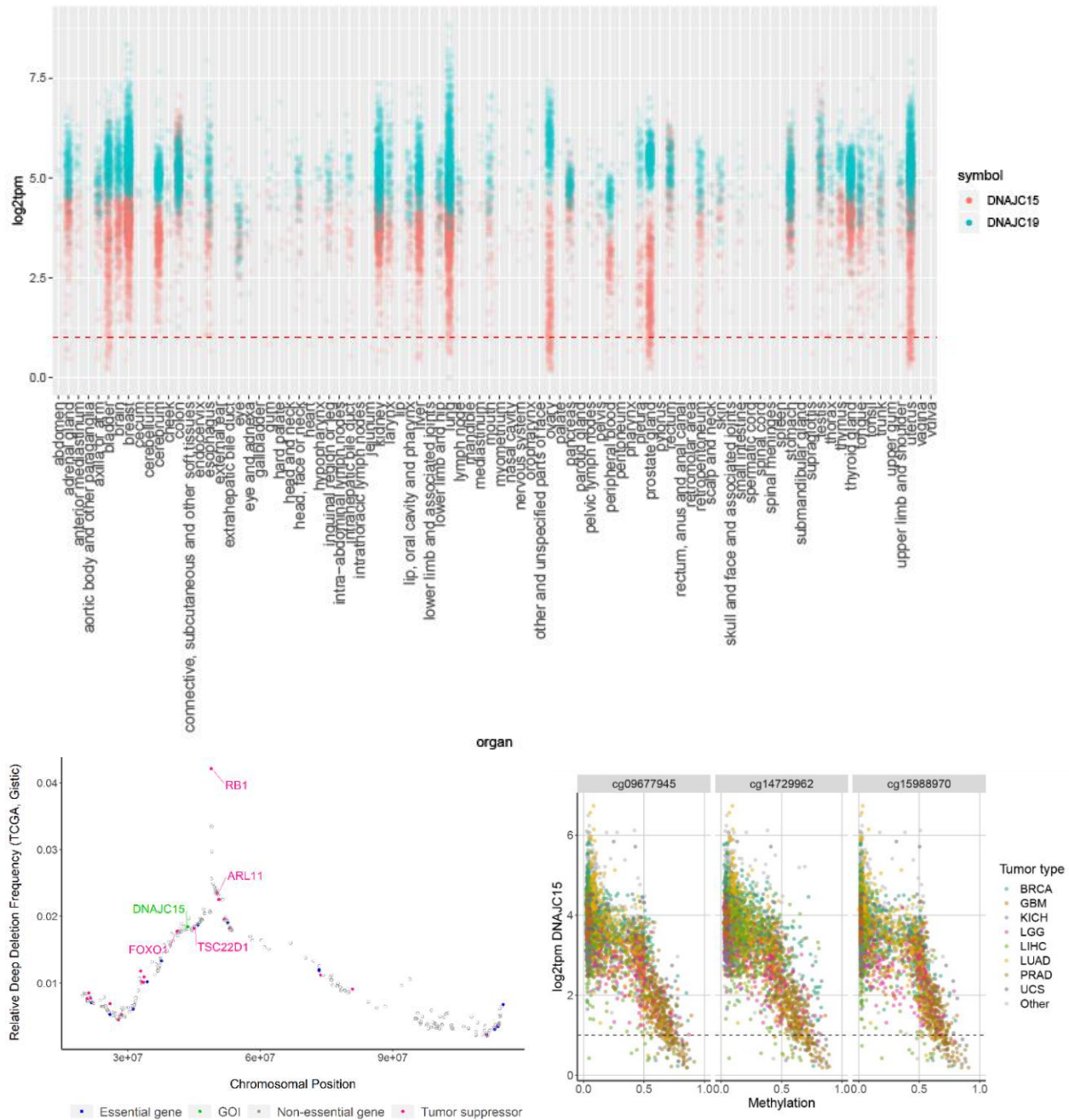
Figure 10: DNAJC15 Expression in cancer (TCGA).
Top: Expression in various cancer types, Bottom left: Deep deletion frequency in all cancer types of genes encoded on chromosome 13, Bottom right: DNA Methylation of DNAJC15 CpG islands correlates strongly with gene expression. Tumor types with low expression of DNAJC15 are colored. Pearson's product moment correlation for the examined CpG islands: cg09677945: $\rho = -0.75$ and $p < 2.2 \cdot 10^{-16}$, cg1472996: $\rho = -0.73$ and $p < 2.2 \cdot 10^{-16}$, cg15988970: $\rho = -0.72$ and $p < 2.2 \cdot 10^{-16}$

34

In normal tissue, both genes are relatively equally expressed (Figure 11). It is especially important, that the biomarker paralog DNAJC15 is expressed in healthy tissues so that a potential therapeutic targeting DNAJC19 does not cause unwanted side-effects. Such effects might occur in ovarian or heart tissues due to low DNAJC15 expression, which need to be monitored closely in further studies. In blood, both genes are frequently not expressed, thus they presumably do not fulfill an essential function in these cells and targeting DNAJC19 should not adversely affect blood cells.
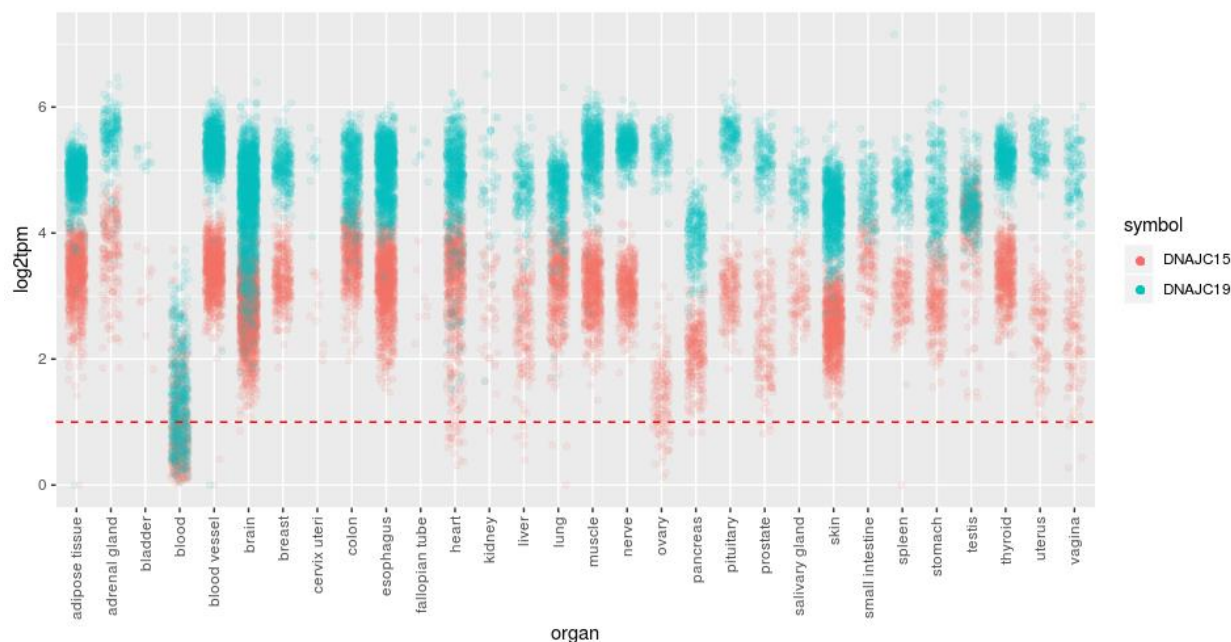
Figure 11: DNAJC15 and DNAJC19 expression in healthy tissues (GTex)

The initial CRISPR depletion assay confirmed the sensitivity of CAL-12T_Cas9_Puro and NCI-H1975_Cas9_Puro cell lines to the knockout of DNAJC19 (Table 5, Figure 12). Only gRNA_RN318 recapitulated the expected knockout phenotype of DNACJ19 as sensitive cells (reported in the public domain) transduced with the other to tested gRNAs did not deplete. Cell lines resistant to DNAJC19 KO (RT-112_Cas9_Puro and SCC-25_Cas9_Puro) did not deplete upon transduction with gRNA_RN318 indicating that depletion is not caused by an unintentional KO of some other essential gene.

Table 5: Selected Cell lines for investigating DNAJC15 and DNAJC19 paralog dependency

	Cell lines	Gene Expression DNAJC19 (TPM)	Gene Expression DNAJC15 (TPM)	DNAJC19 AVANA CERES Score
Sensitive cell lines	CAL-12T	48.6	1.2	-1.09
	NCI-H1975	86.8	1.3	-1.10
Resistant cell lines	RT-112	32.8	16.5	0.17
	SCC-25	64.0	32.3	0.16
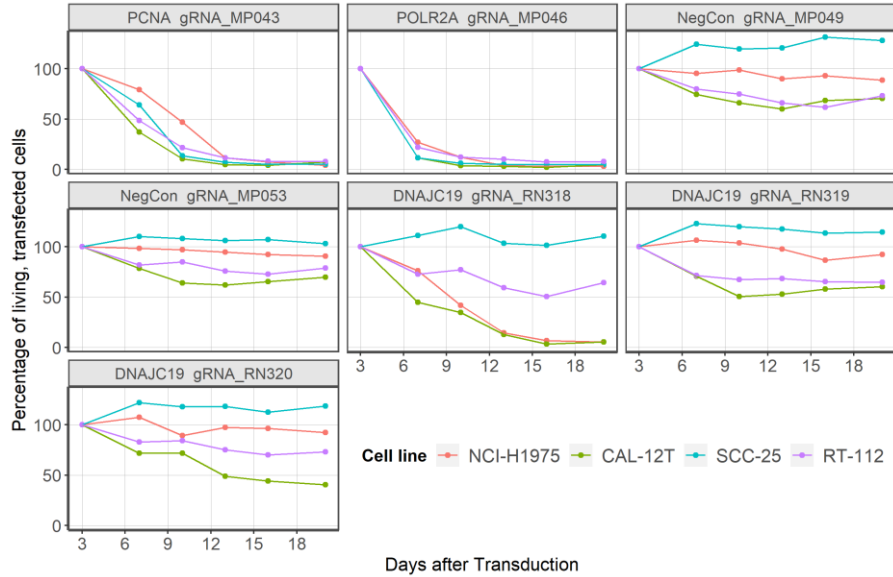
Figure 12: CRISPR depletion assay shows sensitivity of NCI-H1975 and CAL-12T to DNAJC19 KO
Cells have been transduced with gRNAs targeting DNAJC19, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction.

Next, the RIEH construct, for the ectopic expression of the ecotropic retrovirus receptor, was introduced into the cell lines for which sensitivity to DNAJC19 KO was confirmed. This receptor allows for the viral transduction of overexpression (OE) constructs. Next, overexpression constructs encoding for codon-optimized, doxycycline-inducible DNAJC15 and DNAJC19 were transduced. Successful transduction with the inducible OE construct was verified through western blotting. Additionally, the expression levels were examined with doxycycline concentrations between 0.1 and 2 µg/mL to select the optimal condition for further experiments (Figure 13). Since DNAJC19 is already present in the selected, DNAJC19 sensitive cells before doxycycline introduction of the OE construct, successful overexpression was assumed if the band specific for DNAJC19 gets stronger upon doxycycline induction. All tested doxycycline levels sufficed for induction, hence, a concentration of 0.5 µg/mL doxycycline was chosen for subsequent experiments.
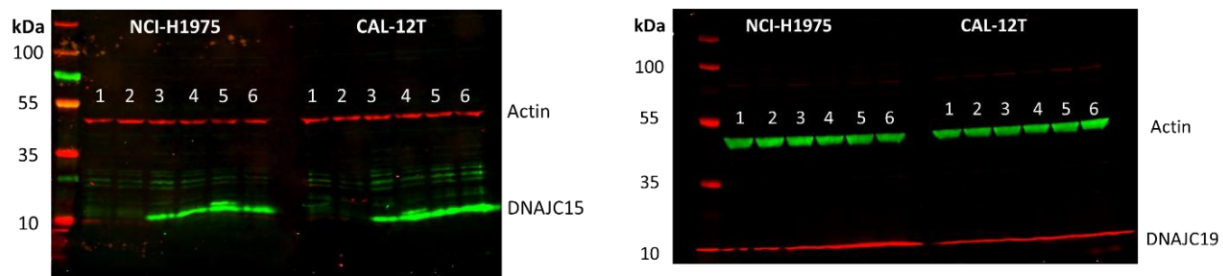
Figure 13: Western blot to confirm overexpression of DNAJC15 (left) and DNAJC19 (right)
Cells were cultivated for 4 days with different levels of doxycycline. Lane 1: Parental cell line, Lane 2 – 6: Cell line with overexpression construct at various doxycycline concentrations (0 µg/mL, 0.1 µg/mL, 0.5 µg/mL, 1 µg/mL, 2 µg/mL)

When the CRISPR depletion assay was performed with and without induction of DNAJC15, an unambiguous rescue of NCI-H1975_Cas9_Puro_RIEH_RN183 was achieved (Figure 14). The sensitivity to DNAJC19 KO was no longer observable when DNAJC15 was overexpressed, therefore we conclude that these two genes are indeed functionally redundant and can be described in the context of a paralog

36

dependency. Unfortunately, the paralog dependency could not be observed in the second selected cell line CAL-12T.
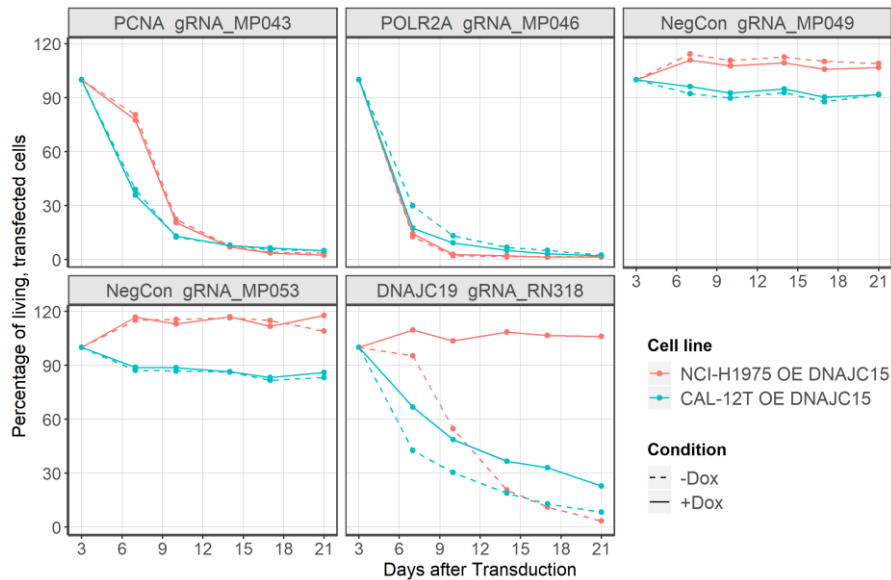
Figure 14: CRISPR depletion assay with inducible overexpression of DNAJC15.
NCI-H1975 can be rescued from DNAJC19 sensitivity through overexpression of DNAJC15. Cells have been transduced with gRNAs targeting DNAJC19, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction. Overexpression of DNAJC15 was induced with 0.5 µg/mL doxycycline.

The CRISPR depletion assay was repeated with and without the induction of codon-optimized DNAJC19 using the effective gRNA_RN318 targeting DNAJC19 (Figure 16). As the DNA sequence of the OE construct differs from the endogenous DNAJC19 gene, the gRNA is not be able to knock out the ectopically expressed DNAJC19 (Figure 15) and consequently, transduced cells should no longer deplete upon CRISPR mediated loss-of-function. This was the case for the cell line NCI-H1975_Cas9_Puro_RIEH_RN184, demonstrating that the CRISPR-mediated depletion phenotype is on-target and specifically caused by the knockout of DNAJC19. However, transduced CAL-12T_Cas9_Puro_RIEH_RN184 were not rescued by the overexpression of codon-optimized DNAJC19. Therefore, the observed phenotype in CAL-12T cells may not be specific to the loss of DNAJC19 but instead stems from an off-target effect in this specific cell line. The sequence of gRNA_RN318 was reassessed using the Human BLAT search to identify matching sequences in the human genome (https://www.genome.ucsc.edu/cgi-bin/hgBlat). One off-target match was indeed found in a non-coding region on chromosome 14. This might be one possible explanation for the apparent sensitivity of CAL-12T, however, also other unknown effects may be underlying the observed effects in this particular cell line. As the loss-of-function phenotype of DNAJC19 could be rescued by overexpression of DNAJC19 in NCI-H1975_Cas9_Puro_RIEH_RN184 cells, the potential off-target effect does not seem to affect this cell line. This experiment on the one hand validates the paralog dependency of DNAJC15 and DNAJC19 as both genes were able to rescue NCI-H1975. On the other, it explains why the rescue through DNAJC15 was not achieved in CAL-12T: The initially reported sensitivity was not caused by a KO of DNAJC19, ergo it is not expected that DNAJC15 is able to rescue the transduced cells.

```
                                           190         200         210
               ....|....|....|....|....|....|....|....|....|....
Endogenous DNAJC19   CGGGAAGCAGCATTAATACTAGGTGTAAGCCCTACTGCC
RN_318                          AGCATTAATACTAGGTGTA
Codon-optimized DNAJC19  AGGGAGGCCGCCCTGATCCTGGGCGTGTCTCCAACAGCC
```

Figure 15: Sequence alignment of endogenous DNAJC19, used DNAJC19 gRNA and codon-optimized DNAJC19 overexpression construct

Figure 16: CRISPR depletion assay with inducible overexpression of codon-optimized DNAJC19.
NCI-H1975 can be rescued from DNAJC19 sensitivity through overexpression of codon-optimized exogenous DNAJC19. Cells have been transduced with gRNAs targeting DNAJC19, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction.

8.3   RPP25 and RPP25L

With the AVANA CERES screening data, a strong correlation between sensitivity to RPP25L KO and low expression of RPP25 was observed (Figure 17), suggesting that RPP25 might serve as the biomarker paralog explaining the sensitivity profile across hundreds of cancer cell lines for the target paralog RPP25. Conversely, only very few cell lines in the AVANA screen exhibit known low expression of RPP25L or sensitivity to RPP25 KO.
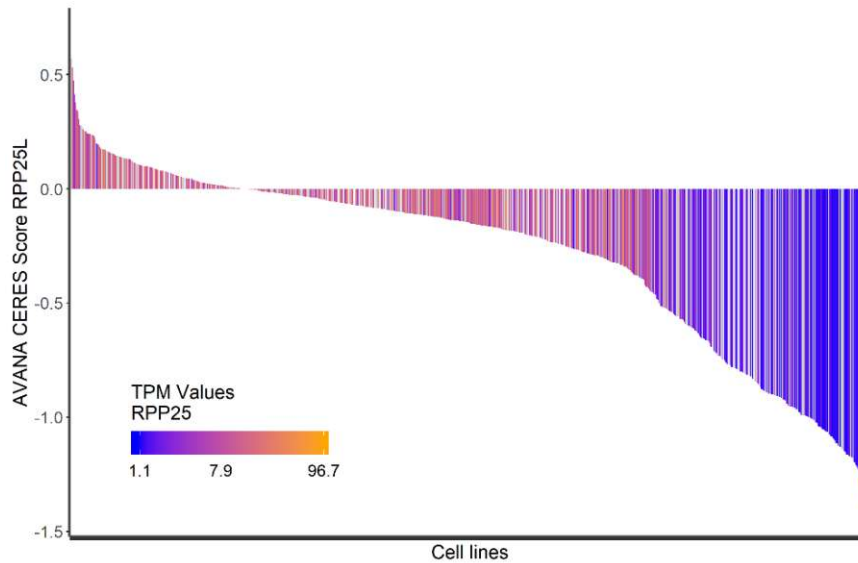
Figure 17: Sensitivity of cancer cell lines to RPP25L knockout (AVANA CERES screen results)

According to the TCGA data, RPP25 is not expressed in samples from multiple different tumor types including brain, cerebrum, kidney, liver, peripheral blood, prostate and uterine cancers (Figure 18a). However, the GISTIC results indicate a very low deep deletion frequency in only 0.16% in all TCGA samples (https://www.cbioportal.org/). The highest deep deletion frequency in the various cancer types can be found in ovarian carcinoma (OV) accounting for 1.04% of the samples. Consistently, there is no obvious frequently deleted tumor suppressor in close proximity to the chromosomal position of RPP25, which might entail higher than average deep deletions of this gene. Furthermore, core-essential genes reside on the same chromosome with higher deep deletion frequencies, raising doubt whether even these low numbers are overestimations (Figure 18b). Gandhi et al. suggested, that RPP25 loss might also occur due to promotor hypermethylation (62). Indeed, we found a correlation between RPP25 expression and CpG island methylation in TCGA samples (Figure 18c). Different to DNAJC15, not every site described as CpG island is correlated with RPP25 expression (Appendix Figure 48). Furthermore, there are cancer samples that do not express RPP25 despite an un-methylated state at the CpG otherwise correlated with expression. Hence, we concluded that DNA methylation is one of multiple mechanisms regulating RPP25 expression. As stated above, deep deletions do not explain the loss of RPP25 sufficiently. Therefore, another mechanism like histone modifications or transcription factor expression and binding might be responsible for the differences in expression.
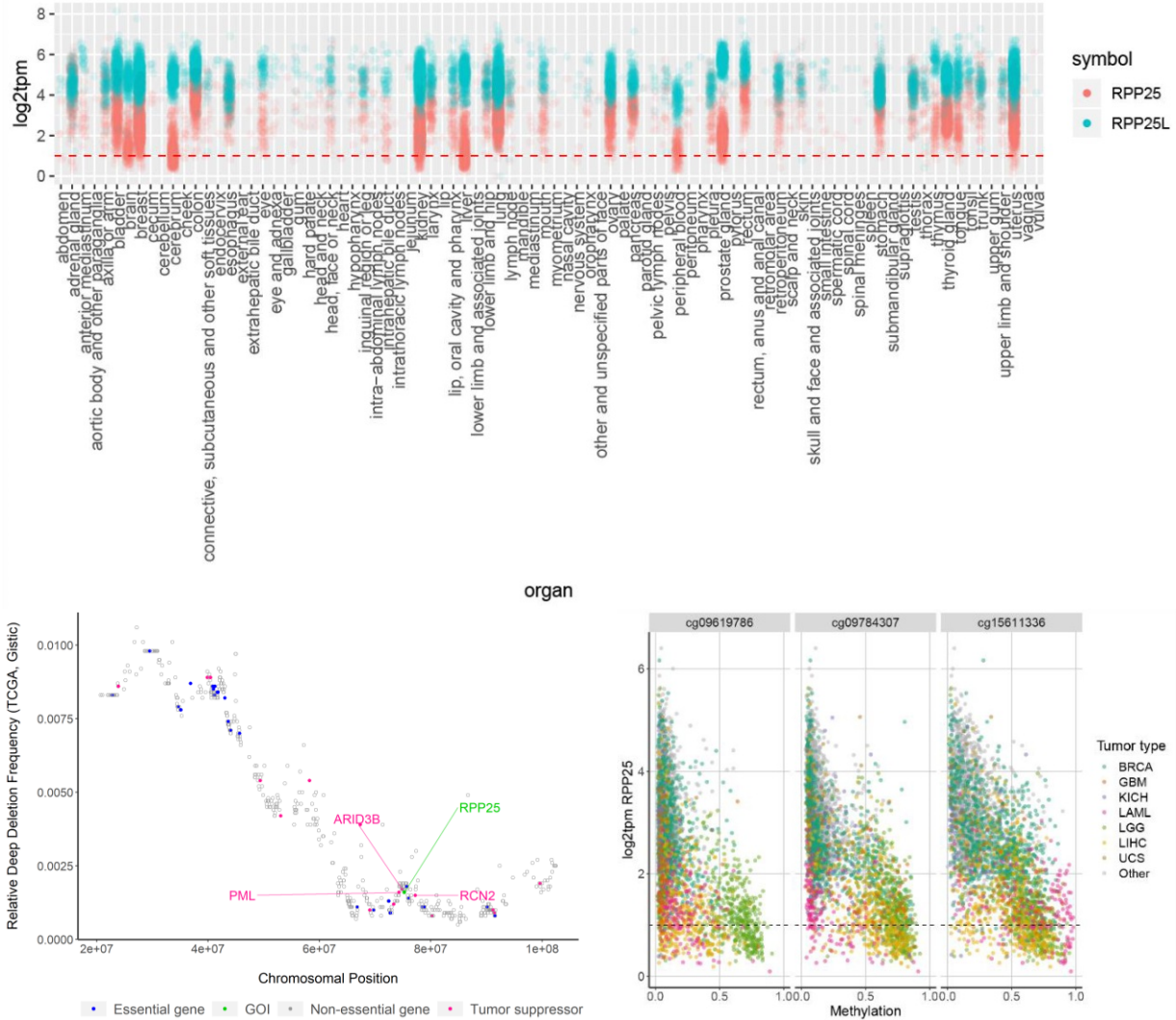
Figure 18: RPP25 Expression in cancer (TCGA).
Top: Expression in various cancer types, Bottom left: Deep deletion frequency in all cancer types of genes encoded on chromosome 15, Bottom right: DNA Methylation of RPP25 CpG islands correlates with gene expression. Tumor types with low expression of RPP25 are colored. Pearson's product moment correlation for the examined CpG islands: cg09619786: $\rho = -0.54$ and $p < 2.2 \cdot 10^{-16}$, cg09784307: $\rho = -0.68$ and $p < 2.2 \cdot 10^{-16}$, cg15611336: $\rho = -0.65$ and $p < 2.2 \cdot 10^{-16}$

According to the GTex data (Genotype-tissue expression, https://GTexportal.org/ (146)), the expression in normal tissue for these two putative paralogs appears to be tissue-dependent (Figure 19). Especially in blood, brain, heart, muscle and liver cells, RPP25L is the dominant form, while RPP25 is underrepresented. This complicates potential drug development targeting RPP25L, as tissue-specific drug delivery might be required to avoid negative side effects.
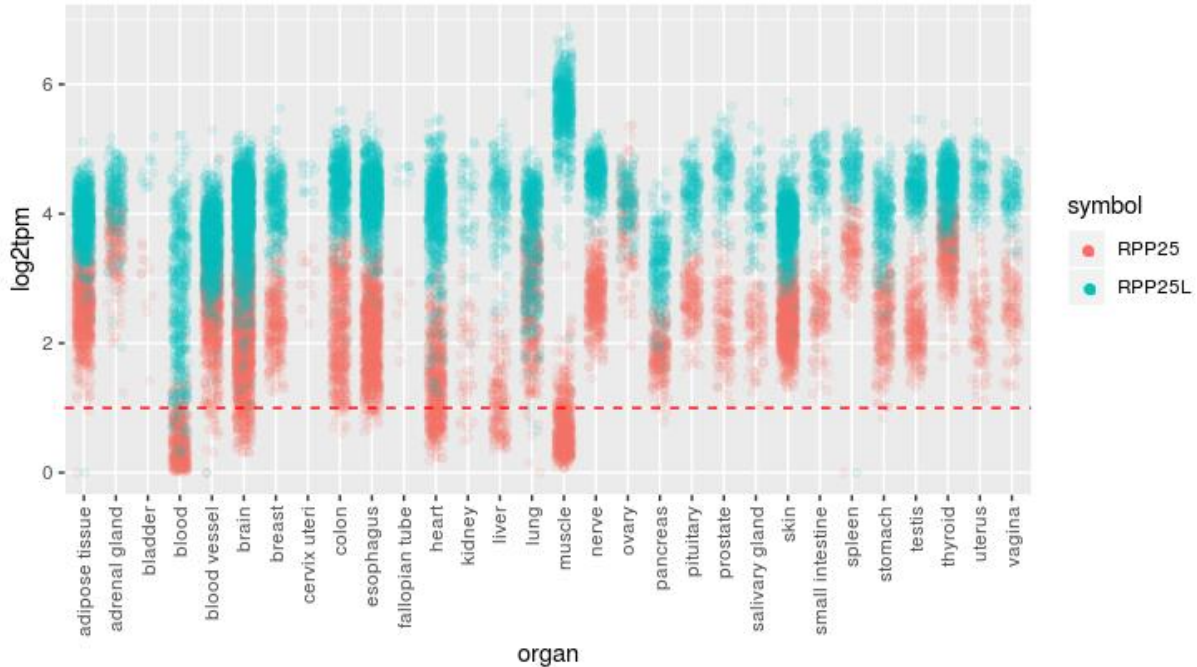
Figure 19: RPP25 and RPP25L expression in healthy tissues (GTex)

As a first step in investigating the putative paralog functional redundancy between RPP25 and RPP25L, the sensitivity of the selected cell lines KYSE-150_Cas9_Puro and U-2 OS_Cas9_Puro to the KO of RPP25L was assessed in CRSPR depletion experiments (Table 6, Figure 20). According to the CCLE (https://portals.broadinstitute.org/ccle), RPP25 is methylated in U-2 OS and not methylated in KYSE-150. The resistant cell line SK-N-MC_Cas9_Puro consistently did not deplete upon RPP25L CRISPR-mediated loss-of-function. These data suggest that the tested gRNAs specific for RPP25L do not cause general cytotoxicity and that potential sensitivity observed in sensitive cell lines is not caused by an off-target effect. Unfortunately, the CRISPR depletion assay with a second selected resistant cell line NCI-H2170_Cas9_Puro was inconclusive, as the positive controls did not deplete as expected. As the resistant cell lines were primarily tested in order to detect possible off-target effects and the depletion experiment with SK-N-MC_Cas9_Puro raised no such concern, we continued using gRNA_RN290 and gRNA_RN291 in further CRISPR experiments. These two gRNAs were selected for their strong depletion capability in the sensitive cell lines while additionally their target locations in RPP25L are relatively far apart but still within the ALBA (acetylation lower binding affinity) domain.

Table 6: Selected Cell lines for investigating RPP25 and RPP25L paralog dependency

	Cell lines	Gene Expression RPP25L (TPM)	Gene Expression RPP25 (TPM)	RPP25L AVANA CERES Score
Sensitive cell lines	U-2 OS	36.7	2.2	-1.31
	KYSE-150	40.9	1.0	-1.25
Resistant cell lines	SK-N-MC	24.1	18.3	0.36
	NCI-H2170	21.9	16.0	0.25
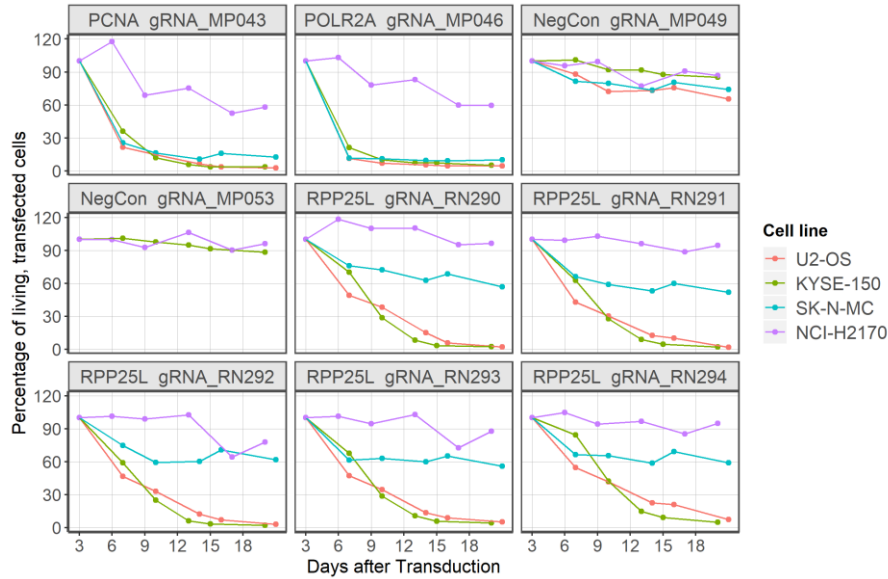
Figure 20: CRISPR depletion assay shows sensitivity of U-2 OS and KYSE-150 to RPP25L KO.
Cells have been transduced with gRNAs targeting RPP25L, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction.

After ecotropization and introduction of the RPP25 OE constructs into the sensitive cell lines, the CRISPR depletion assay was repeated (Figure 22). All tested doxycycline concentrations led to a strong induction of RPP25 validated by western blotting. Subsequently, a concentration of 0.5 µg/mL doxycycline was selected for the CRISPR experiments (Figure 21). In the CRISPR depletion assay, there was no obvious difference in RPP25L sensitivity between doxycycline induced and not-induced cells. Therefore, it seemed as if the rescue was not successful. However, no clear conclusions can be drawn, as the Western Blot suggested that the overexpression of RPP25 was leaky and therefore happens in the absence of doxycycline (Figure 21).
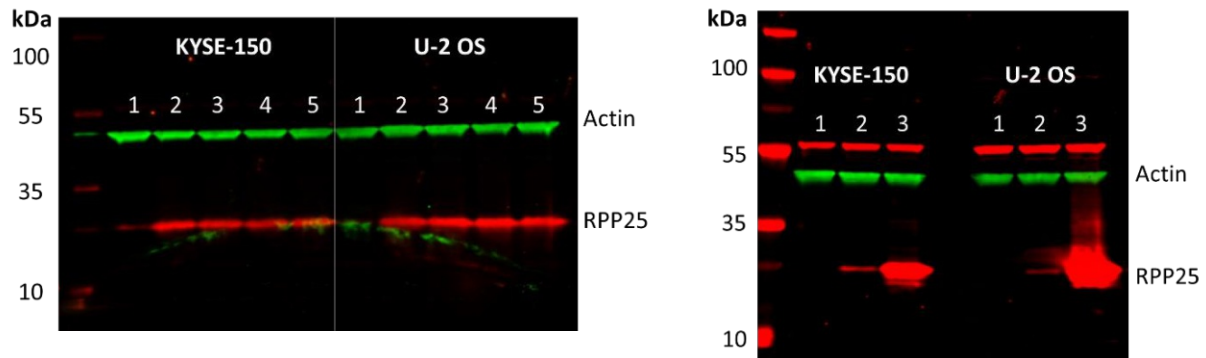
Figure 21: Western blot to confirm overexpression of RPP25
Cells were cultivated for 3 days with different levels of doxycycline. Left: Doxycycline induction titration (lane 1 – 5 : 0 µg/mL doxycycline, 0.1 µg/mL doxycycline , 0.5 µg/mL doxycycline, 1 µg/mL doxycycline, 2 µg/mL doxycycline)
Right: Expression of RPP25 is leaky. (lane 1: parental, 2: no doxycycline addition, 3: 0.5 µg/mL doxycycline)
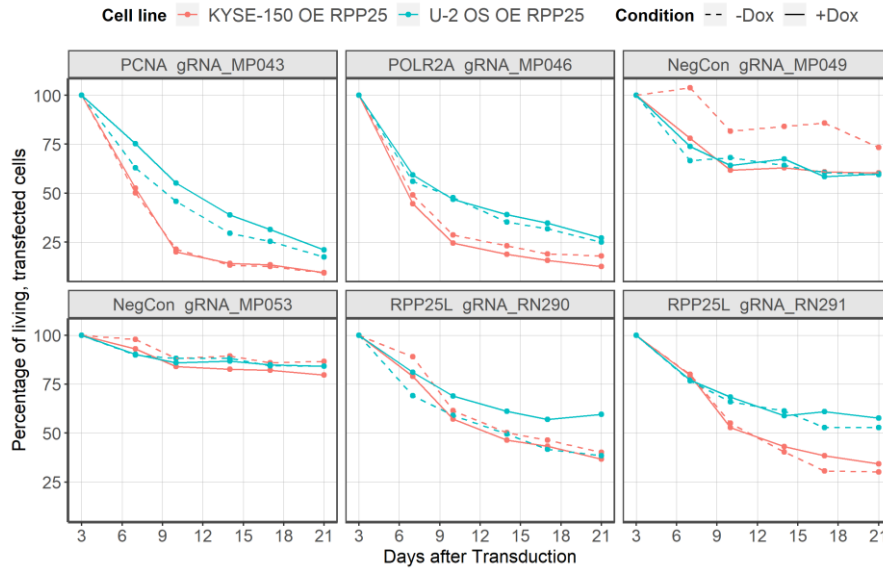
Figure 22: CRISPR depletion assay with inducible RPP25 overexpression.
Cells have been transduced with gRNAs targeting RPP25L, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction. Overexpression of RPP25 was induced with 0.5µg/mL doxycycline. As RPP25 expression was leaky, a more pronounced difference between induced and not-induced cell might be achievable.

In order to circumvent the leaky expression observed in the pool of transduced cells, single cell clones were produced and tested via western blot (Figure 23). The clones KYSE-150_Cas9_Puro_RIEH_RN179_#1 and U-2 OS_Cas9_Puro_RIEH_RN179_#1 displayed the desired expression behavior and were selected for further experiments.
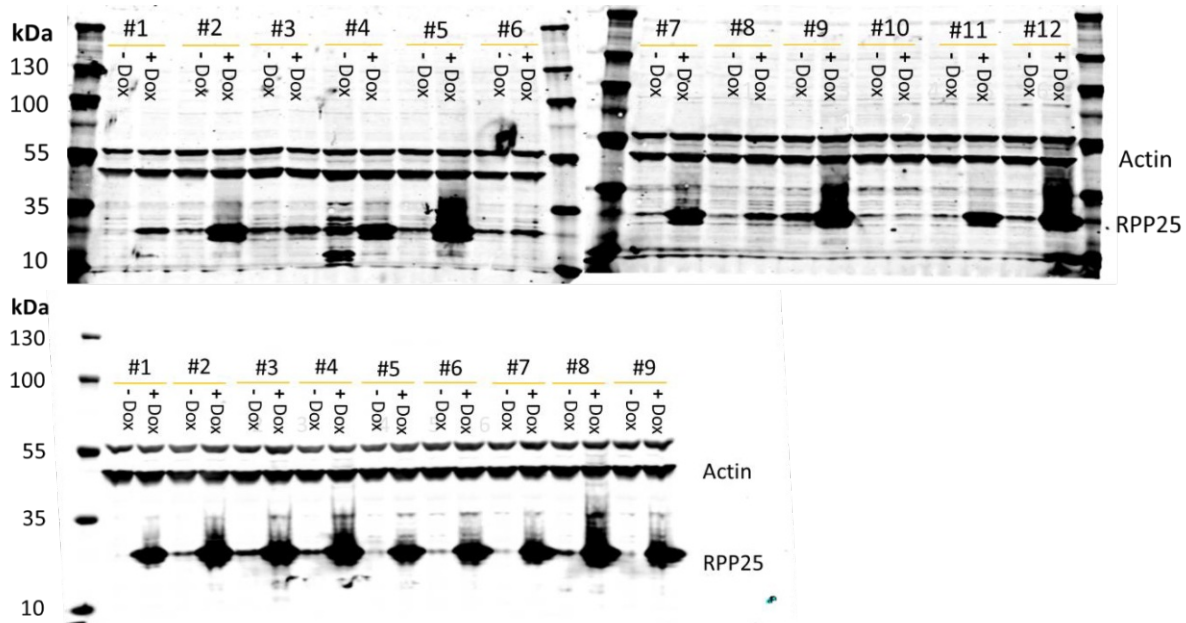
Figure 23: Single cell clones of cell line KYSE-150 (top) and U-2 OS (bottom) expressing RPP25 when doxycycline induced.
Cells were cultivated for 3 days +/- 0.5 µg/mL doxycycline. For both cell lines, clone #1 was selected for further experiments.

With these cell lines the CRISPR experiment was repeated again (Figure 24). These experiments demonstrate that the rescue is successful with clone U-2 OS_Cas9_Puro_RIEH_RN179_ #1. When no doxycycline is added to the medium, RPP25 is not expressed at basal levels and hence the cells are sensitive to RPP25L loss-of-function with gRNA_RN290. Conversely, when RPP25 is overexpressed by adding doxycycline, the RPP25L loss-of-function lethality phenotype can be rescued. Similar results were obtained for gRNA_RN291, albeit with a weaker effect. KYSE-150_Cas9_Puro_RIEH_RN179_#1, on the other hand, was not rescued by the induction of RPP25. It is unclear as to why the rescue does not function in this second cell line. Potential explanations are: insufficient expression levels of the paralog, a cell type specific off-target of the used gRNAs or a functionally non-redundant function of RPP25L in this particular cell line that cannot be compensated for by RPP25. The positive rescue data in one cell line clearly provide evidence for functional redundancy between RPP25L and RPP25.
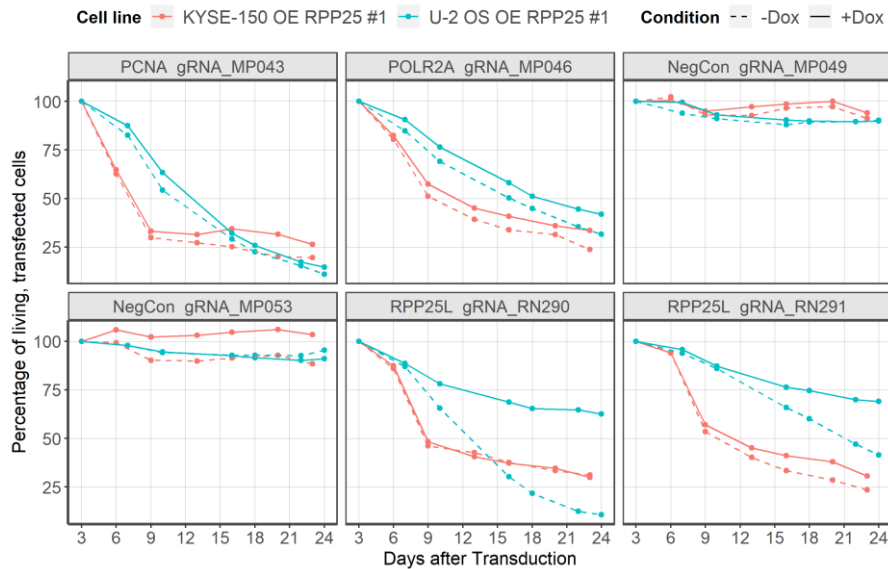
Figure 24: CRISPR depletion assay with inducible RPP25 overexpression using single cell clones
Cells have been transduced with gRNAs targeting RPP25L, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction. Overexpression of RPP25 was induced with 0.5μg/mL doxycycline. The U-2 OS clone #1 was rescued through overexpressing of RPP25L upon induction.

8.4 PAPSS2 and PAPSS1

According to the AVANA CERES results combined with gene expression data, low expression of the biomarker paralog PAPSS2 correlates with sensitivity to PAPSS1 loss-of-function (Figure 25). Again, the putative paralog dependency can only be seen in one direction, probably as the even the cell lines with lowest PAPSS1 expression have a sufficient PAPSS1 level to avoid sensitivity to PAPSS2.
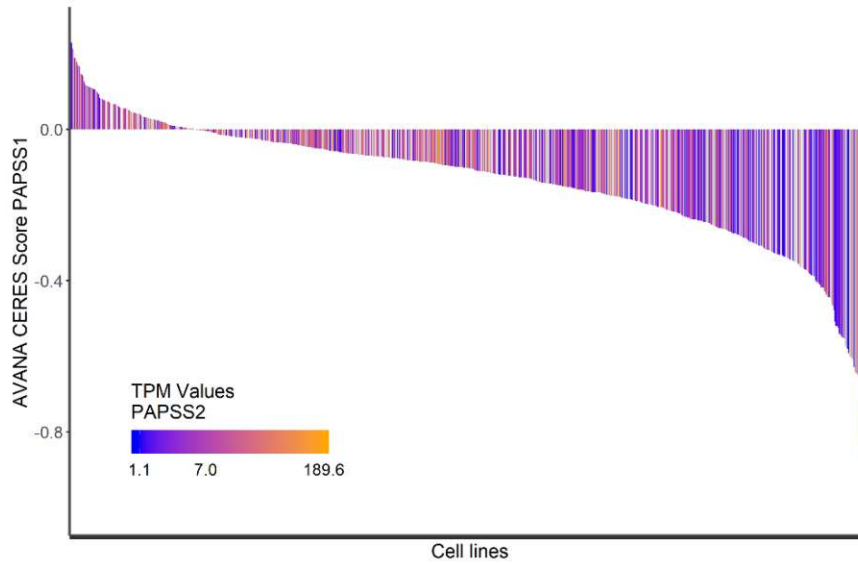
Figure 25: Sensitivity of cancer cell lines to PAPSS1 knockout (AVANA CERES screen results)

In multiple cancer patient samples from various tissues, PAPSS2 is not expressed (Figure 26a). Particularly, the low expression in bladder, eye, kidney, ovary, prostate, thyroid and uterus tumor tissues stands out.

PAPSS2 is deep deleted in 2.19% of all TCGA samples predicted by the GISTIC results (https://www.cbioportal.org/) In certain cancer types, the deep deletions are even more frequent, for example in 4.33% of glioblastoma (GBM), 3.99% of lung squamous cell carcinoma (LUSC), 11.59% of prostate cancer (PRAD), 3.91% of sarcoma (SARC). Furthermore, PAPSS2 is located in close proximity to the frequently deleted known tumor suppressor PTEN (Figure 26b). As there are no essential genes located between PTEN and PAPSS2 on chromosome 10 and assuming that PTEN as a tumor suppressor is actually deep deleted in many cancer cell lines, we conclude that PTEN deletions may cause bystander deletions of PAPSS2 and that those deletions might even be homozygous.
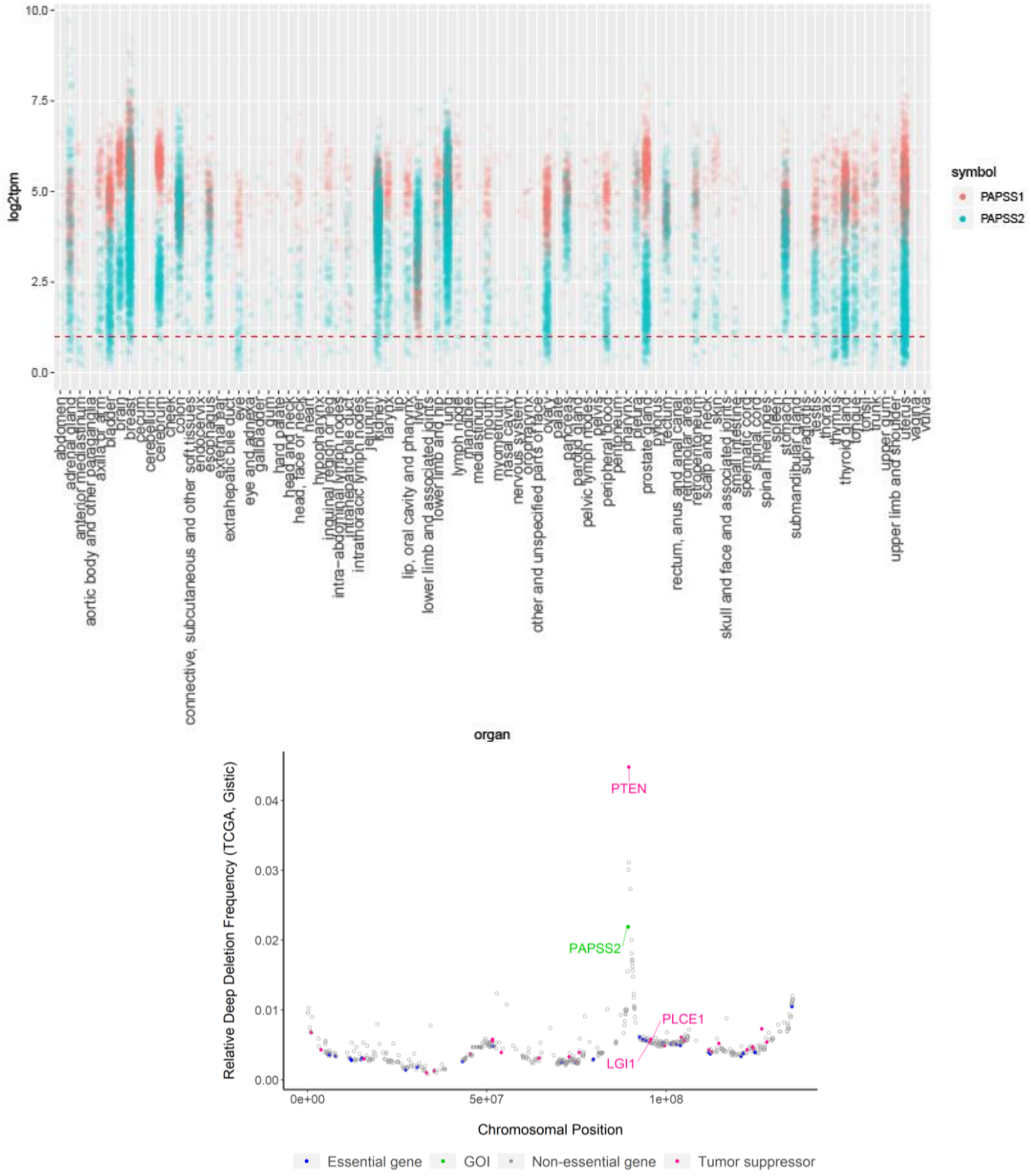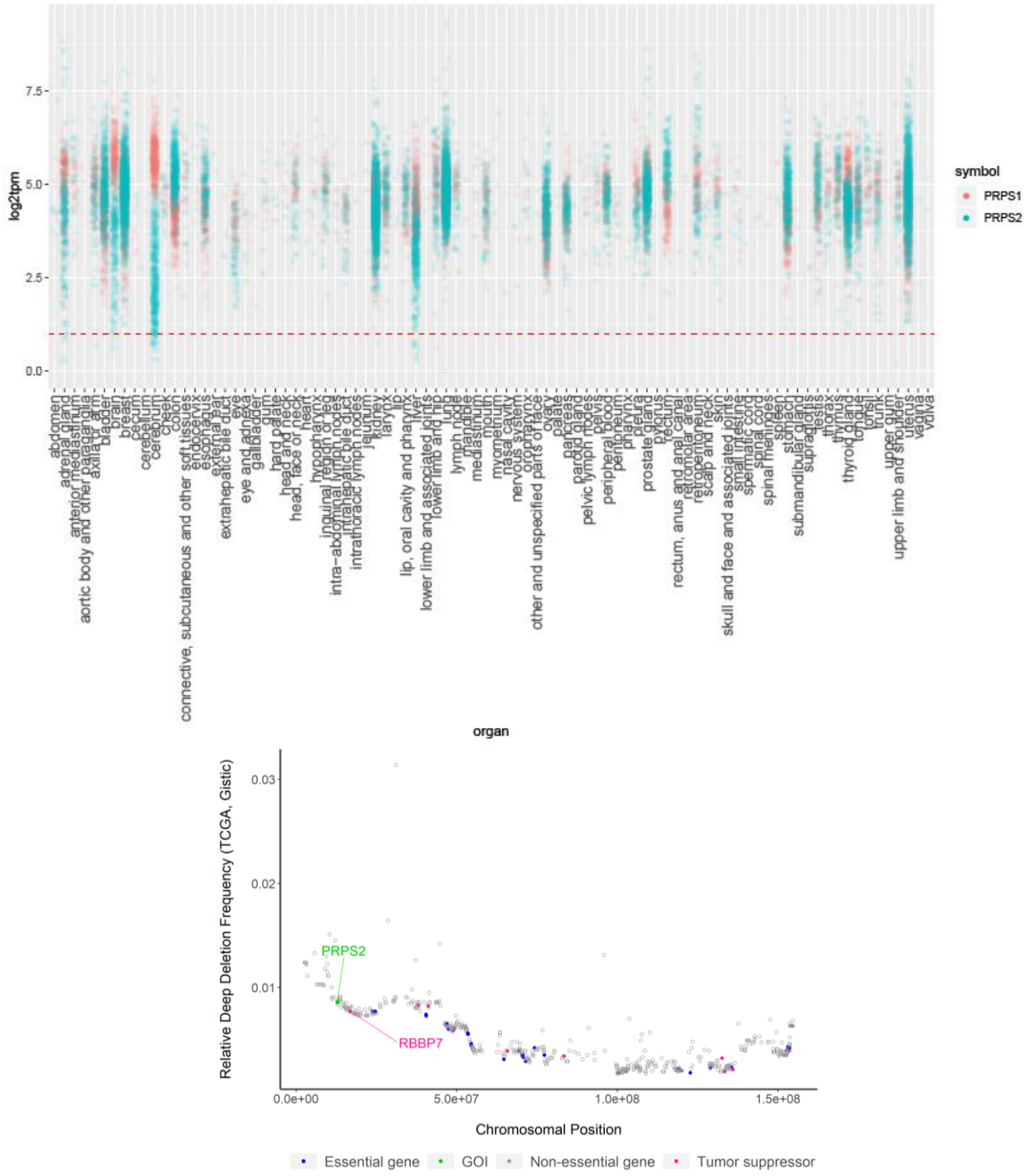
Figure 26: PAPSS2 Expression in cancer (TCGA).
Top: Expression in various cancer types, Bottom: Deep deletion frequency in all cancer types of genes encoded on chromosome 10

In normal tissues, low or absent PAPSS2 expression can be observed in the GTex data for blood, brain, heart, muscle and skin. Hence, unwanted damage to these organs is a potential risk if PAPSS1 targeted in the context of exploiting a putative paralog dependency between PAPSS1/2. In drug development, side effects to these tissues need to be considered and monitored accordingly in order to anticipate putative unwanted effects early on.

46
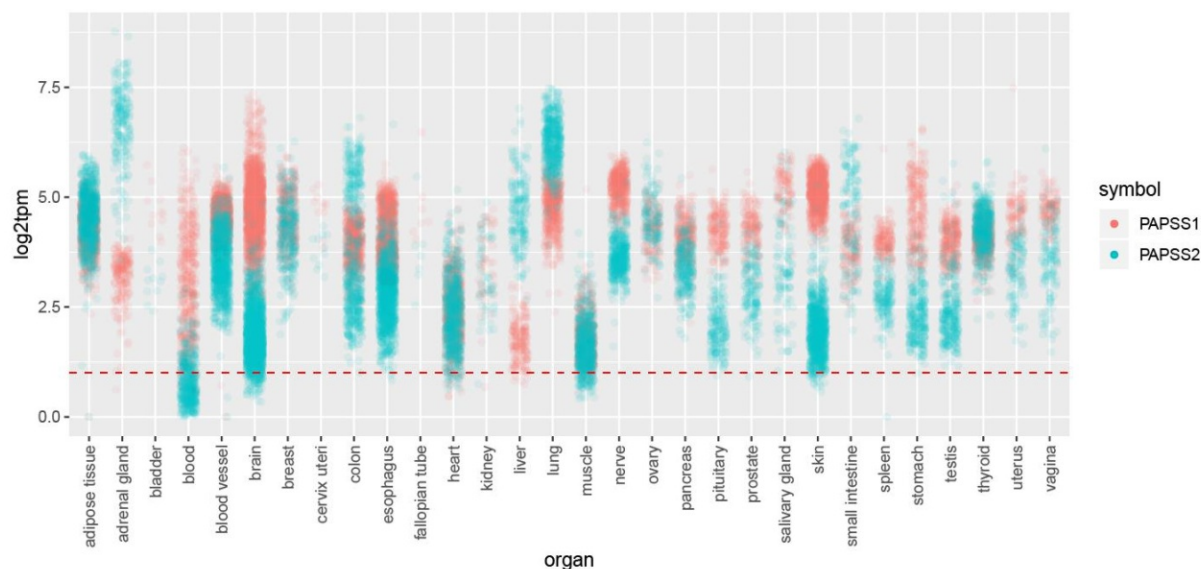
Figure 27: PAPSS1 and PAPSS2 expression in healthy tissues (GTex)

Contrary to the public domain data for PAPSS1, no depletion of GFP positive and hence Cas9/gRNA expressing cells was observed in H4_Cas9_Puro and Hep G2_Cas9_Puro cells (Table 7, Figure 28). As expected, transduced resistant cell lines AsPC-1_Cas9_Puro and HLF_Cas9_Puro did not deplete as well. There are two possible explanations for this behavior: H4 and Hep G2 are in fact not sensitive to PAPSS1 loss-of-function or none of the tested gRNAs was able to successfully knock out PAPSS1. As five different gRNAs targeting both the adenylyl-sulfate kinase domain (gRNA_RN328 and gRNA_RN329) as well as the sulfate adenylyltransferase domain (gRNA_RN330, gRNA_RN331 and gRNA_RN332) of PAPSS1 had been selected, it is unlikely that none of these gRNAs causes a dysfunction of PAPSS1. Hence, we reasoned that the knockout of PAPSS1 probably did not block proliferation or impacts viability of H4_Cas9_Puro and Hep G2_Cas9_Puro cells. As these cell lines do not express PAPSS2 according to the reported TPM-values, there seems to be no paralog dependency between PAPSS1 and PAPSS2.

Table 7: Selected Cell lines for investigating PAPSS2 and PAPSS1 paralog dependency

	Cell lines	Gene Expression PAPSS1 (TPM)	Gene Expression PAPSS2 (TPM)	PAPSS1 AVANA CERES Score
Sensitive cell lines	H4	28.1	1.0	-0.74
	Hep G2	38.6	1.2	-0.52
Resistant cell lines	AsPC-1	51.0	21.8	0.27
	HLF	45.9	52.0	0.15
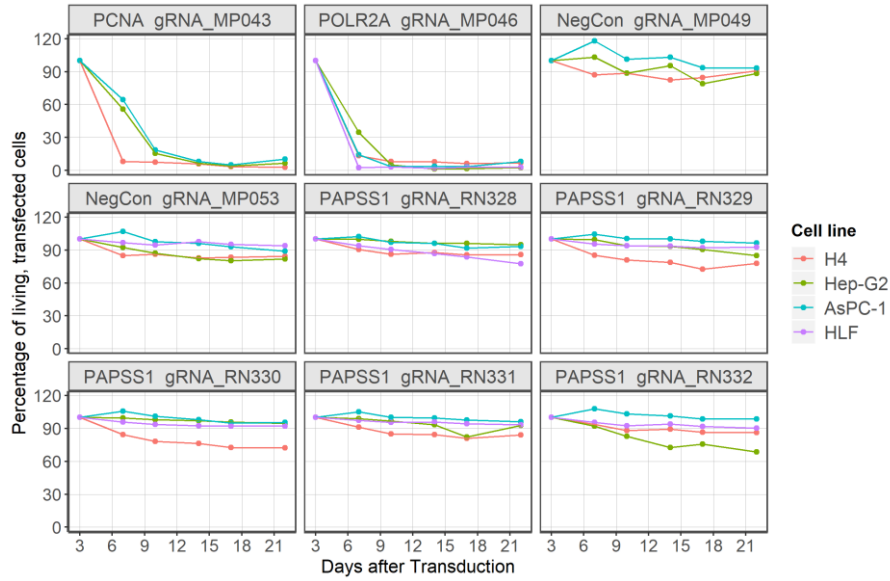
Figure 28: PAPSS1 CRISPR depletion assay shows that cell lines, which do not express PAPSS2 (H4 and Hep G2) are not sensitive to PAPSS1 KO.
Cells have been transduced with gRNAs targeting PAPSS1, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction.

8.5 PRPS2 and PRPS1

In the AVANA CERES screen, a subset of cell lines were identified which were sensitive to PRPS1 loss-of-function and had low expression of PRPS2 (Figure 29). Even though the number of sensitive cell lines was lower than in the other paralog dependencies we investigated the gene pair in this thesis as the correlation was clearly visible. Hence, PRPS2 was selected as the biomarker paralog for the target PRPS1.
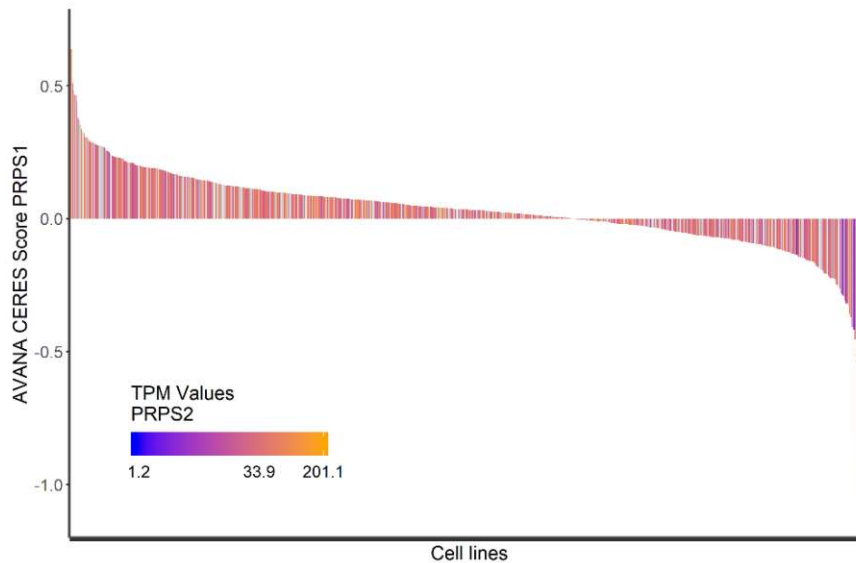
Figure 29: Sensitivity of cancer cell lines to PRPS1 knockout (AVANA CERES screen results)

Low level expression of PRPS2 occurs in a subset of abdomen, cerebrum and liver TCGA cancer samples (Figure 30a, https://www.cbioportal.org/). According to the GISTIC results, PRPS2 is deep deleted in 0.86% of all TCGA samples. Tumors with higher deep deletion frequencies include for instance esophageal carcinoma (ESCA) with 4.35%, lymphoma (DLBC) with 2.08%, lung squamous cell carcinoma (LUSC) with 1.8%, ovarian carcinoma (OV) with 1.73%, head and neck squamous cell carcinoma (HNSC) with 1.72% and stomach carcinoma (STAD) with 1.59%. PRPS2 is among the first 50 genes encoded on the chromosome X (Figure 30b). Notably, none of these genes is essential and the GISTIC deep deletion frequency increases the closer the genes are to the telomeric end of the chromosome (Figure 30). Therefore, we hypothesized that this arm of the chromosome might be lost in cancer cells. However, whether this loss is homozygous cannot be unambiguously inferred from this data.

Figure 30: PRPS2 Expression in cancer (TCGA).
Top: Expression in various cancer types, Bottom: Deep deletion frequency in all cancer types of genes encoded on the X chromosome

Unfortunately, PRPS2 is frequently expressed at low levels in healthy muscle, heart, brain and blood cells (Figure 31). Therefore, a therapy targeting PRPS1 may lead to side effects in these tissues, which need to be monitored closely. If PRPS1 is further considered for therapeutic purposes, tissue-specific drug delivery might be required.

50
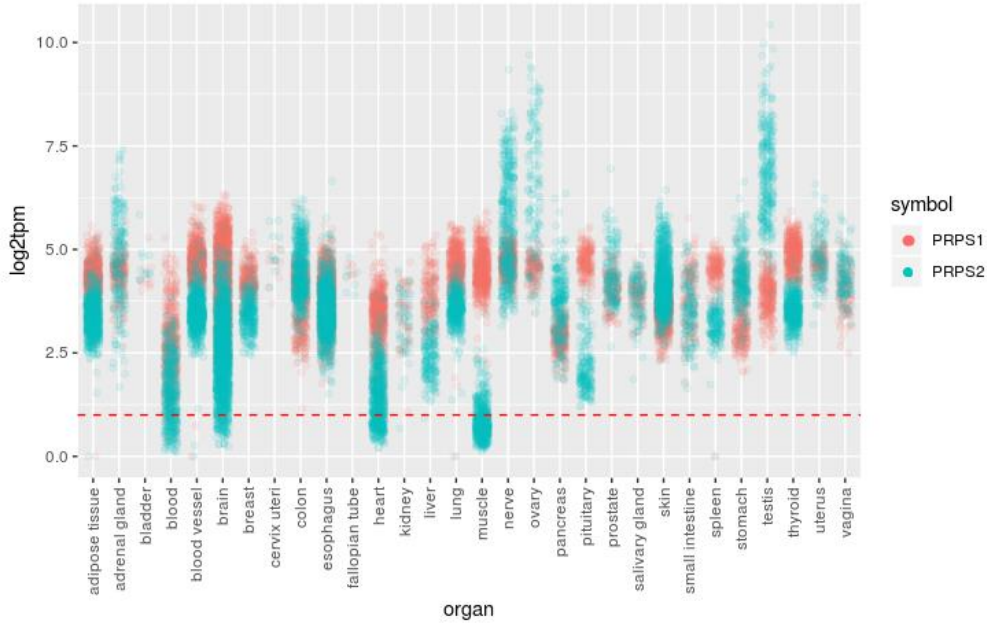
Figure 31: PRPS1 and PRPS2 expression in healthy tissues (GTex)

As a third, rarely expressed paralog exists (PRPS1L1), its expression was considered when selecting cell lines for practical experiments (Table 8). In the CRISPR depletion assay, SK-BR-3_Cas9_Puro was shown to be a sensitive cell line (Figure 32). However, introduction of a Cas9 expression construct into AU565 cells was not successful, as the positive controls did not deplete. Therefore, alternative sensitive cell lines were included. DK-MG_Cas9_Puro was not sensitive to PRPS1 KO despite the public domain data from the AVANA CERES screen (138). This cell line, however, also expressed PRPS2 at a higher level than the other sensitive cell lines, which might have masked the phenotype. Finally, the sensitivity of JOPACA-1_Cas9_Puro could be confirmed. Unfortunately, no gene expression data are presently available for JOPACA-1.

Table 8: Selected Cell lines for investigating PRPS2 and PRPS1 paralog dependency

	Cell lines	Gene expression PRPS1 (TPM)	Gene expression PRPS2 (TPM)	Gene expression PRPS1L1 (TPM)	PRPS1 Avana CERES Score
Sensitive cell lines	SK-BR-3	49.25	1.5	1.0	-1.04
	AU565	52.9	1.2	1.0	-1.00
	DK-MG	34.2	16.9	1.0	-0.45
	JOPACA-1	*NA*	*NA*	*NA*	-1.10
Resistant cell lines	SHP-77	78.4	31.7	1.0	0.47
	JHH-2	185.3	94.6	1.0	0.30
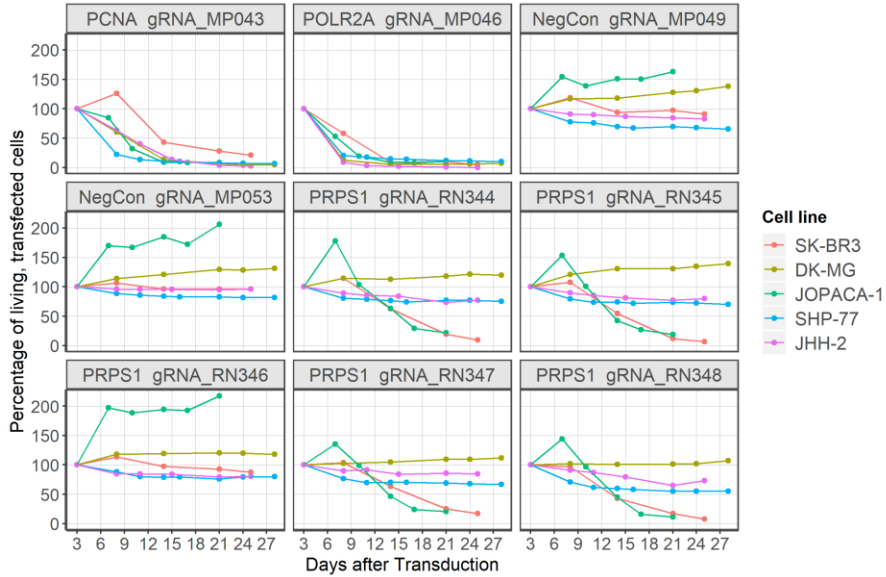
Figure 32: CRISPR depletion assay shows sensitivity of SK-BR-3 and JOPACA-1 to PRPS1 KO.
Cells have been transduced with gRNAs targeting PRPS1, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction.

For ecotropization, the RIEH-construct was introduced intro JOPACA-1_Cas9_Puro and SK-BR-3_Cas9_Puro. To confirm, that the sensitivity displayed in the CRISPR depletion assays was indeed caused by the loss of PRPS1, JOPACA-1_Cas9_Puro_RIEH_RN190 was produced, which overexpress codon-optimized PRPS1. Due to difficulties with ecotropization and the slow growth behavior of SK-BR-3 cells, the overexpression constructs could not be introduced into this cell line in the course of this thesis project. In addition, overexpression of PRPS1 could not be verified via western blot, as the PRPS1 antibody was not specific enough. A test western blot using cell lines which only express PRPS1 or PRPS2 but not both genes was used for this purpose (Figure 33). The used antibody seemed to recognize PRPS2 better than PRPS1.
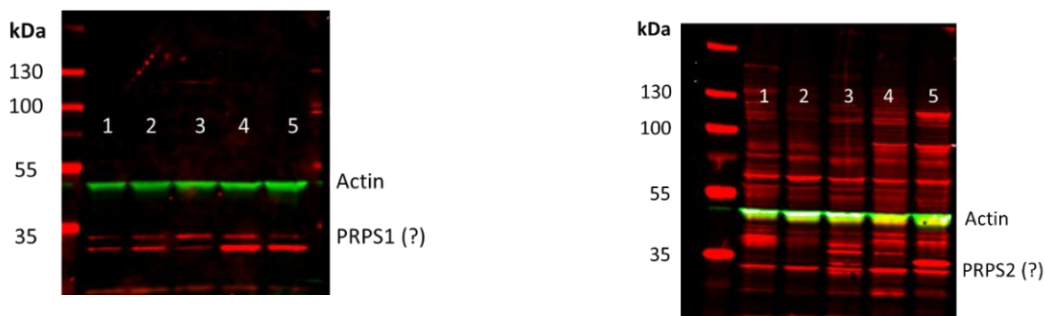
Figure 33: Validation of PRPS1 and PRPS2 antibodies.
Cells expressing only one paralog were cultivated for 5 days and results of the Western blot showed that the tested antibodies were not specific enough.
Left: PRPS1 antibody, Right: PRPS2 antibody. Lane 1 – 3: cell lines only expressing PRPS1 (NCI-H187, TT, KYM-1), lane 4 – 5: cell lines only expressing PRPS2 (HCC1500, PANC 08.13)

Despite not being able to assess the extent of overexpression on a Western Blot, the CRISPR depletion assay was repeated using 0.5μg/mL doxycycline, expecting a stable induction, similar to previous experiments. In JOPACA-1_Cas9_Puro_RIEH_RN190, the rescue through codon-optimized PRPS1 was

successful (Figure 34). This positive result both supports that PRPS1 induction was effective and that the gRNA mediated knockout indeed causes sensitivity due to the loss of PRPS1. gRNA_RN348 caused a sight depletion, presumably due to the high similarity between this sequence and the corresponding codon-optimized sequence (Figure 35). Unfortunately, there was not enough time to also test whether the overexpression of PRPS2 can rescue JOPACA-1, which would ultimately confirm this paralog dependency.
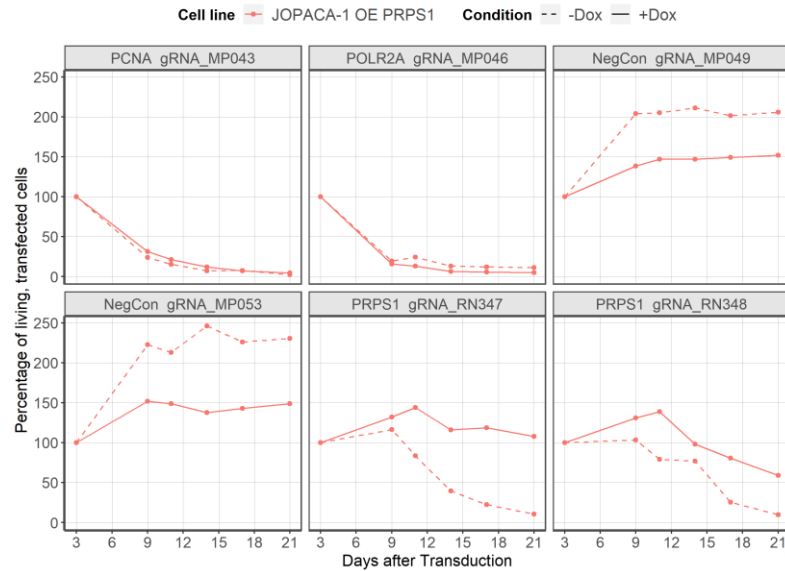
Figure 34: CRISPR depletion assay with inducible overexpression of codon-optimized PRPS1 confirms that the displayed sensitivity was indeed due to the knockout of PRPS1.
Cells have been transduced with gRNAs targeting PRPS1, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction. Overexpression of exogenous PRPS1 was induced with 0.5µg/mL doxycycline.
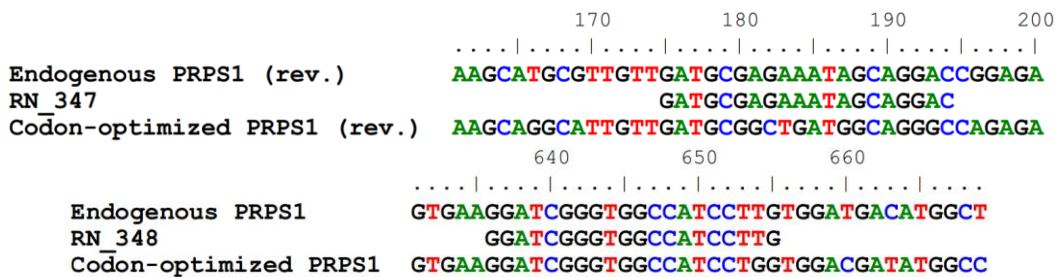
Figure 35: Sequence alignment of endogenous PRPS1, used PRPS1 gRNAs and codon-optimized PRPS1 overexpression construct

8.6   SLC25A37 and SLC25A28

The AVANA CERES data in in conjunction with the gene expression data of the paralogs show, that low expression of SLC25A37 correlates with sensitivity to SLC25A28 knockout (Figure 36). Different from previously described paralog dependencies, none of the cell lines in the AVANA Screen had completely absent expression (indicated by TPM-values of 1) of the biomarker SLC25A37. The correlation of sensitivity to SLC28A28 was established with cell lines harboring low-level SLC25A37 expression, not with cells in which SLC25A37 is completely absent.
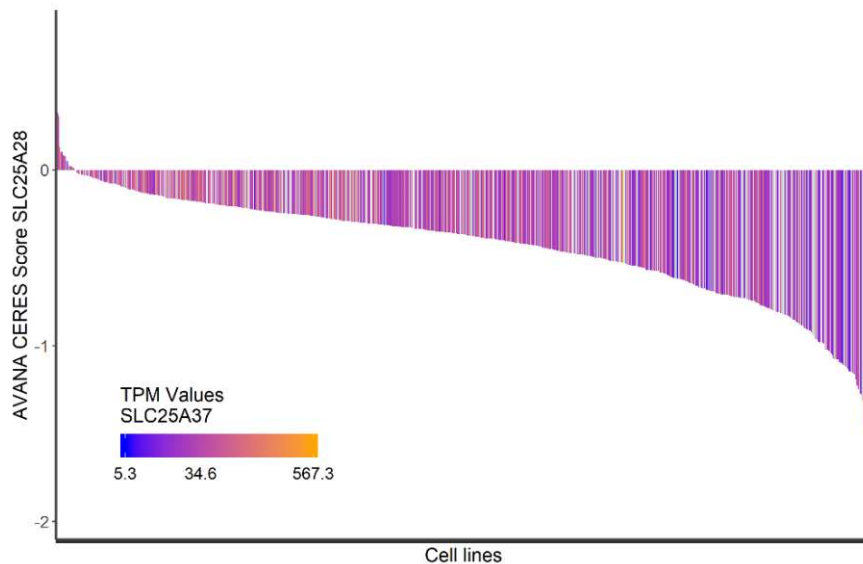
Figure 36: Sensitivity of cancer cell lines to SLC25A28 knockout (AVANA CERES screen results)

SLC25A37 is expressed at low levels in a subset of cancer tissues, including brain and liver samples (Figure 37a). Complete absence of SLC25A37 expression, however, does not seem to occur. In the AVANA screen, the SLC252A28 sensitive cell lines also had low-level expression of SLC25A37. Therefore it can be speculated, that partial loss of SLC25A37 might suffice to trigger sensitivity to SLC25A28.

The GISTIC results suggest a deep deletion frequency of 4.09% of all TCGA samples (https://www.cbioportal.org/). This number is surpassed in bladder (BLCA, 7.11%), breast (BRCA, 6.11%), colon (COAD, 5.78% and COADREAD 6.02%), lymphoma (DLBC, 6.25%), liver (LIHC, 8.65%), lung (LUSC 5.59%), ovary (OV, 7.77%), prostate (PRAD, 17.07%), rectum (READ, 7.88%) and uterine carcinoma (UCS, 8.93%). These values suggest that a large number of patients might benefit from a therapy targeting the putative paralog SLC25A28. Considering the location of SLC25A37 and the GISTIC deep deletion frequency of neighboring genes, it appears as if the end of the chromosome, where SLC25A37 is located, can break off in cancer cells (Figure 37b). Apart from SLC25A37, several tumor suppressor genes are located on this chromosomal arm. However, the essential gene ATP6V1B2 is also positioned between the start of chromosome 8 and SLC25A37. Hence, it is unlikely that this gene loss is homozygous. This interpretation is in line with the fact, that despite the high deep deletion frequency, at least low level expression of SLC25A37 occurs in all TCGA samples. It on the other hand leads to questioning whether the GISTIC deep deletions can be used as a proxy for homozygous deletions.
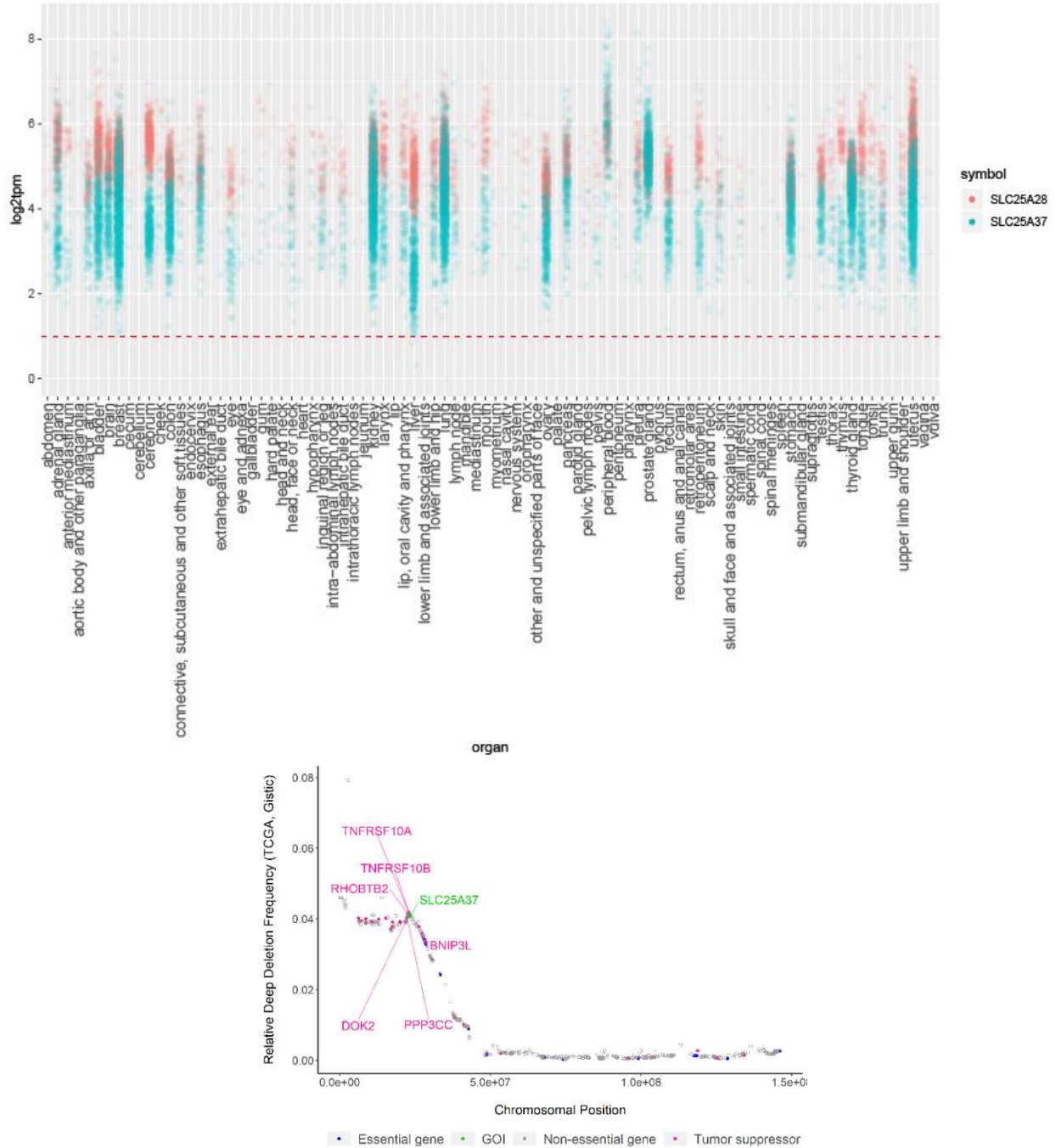
Figure 37: SLC25A37 Expression in cancer (TCGA).
Top: Expression in various cancer types, Bottom: Deep deletion frequency in all cancer types of genes encoded on chromosome 8

Looking at the GTex expression data, both genes are equally expressed in all normal tissues (Figure 38). Only in blood cells, SLC25A37 expression is especially pronounced. As SLC25A28 is the target paralog, no unwanted side effects are expected in any of the tissues.
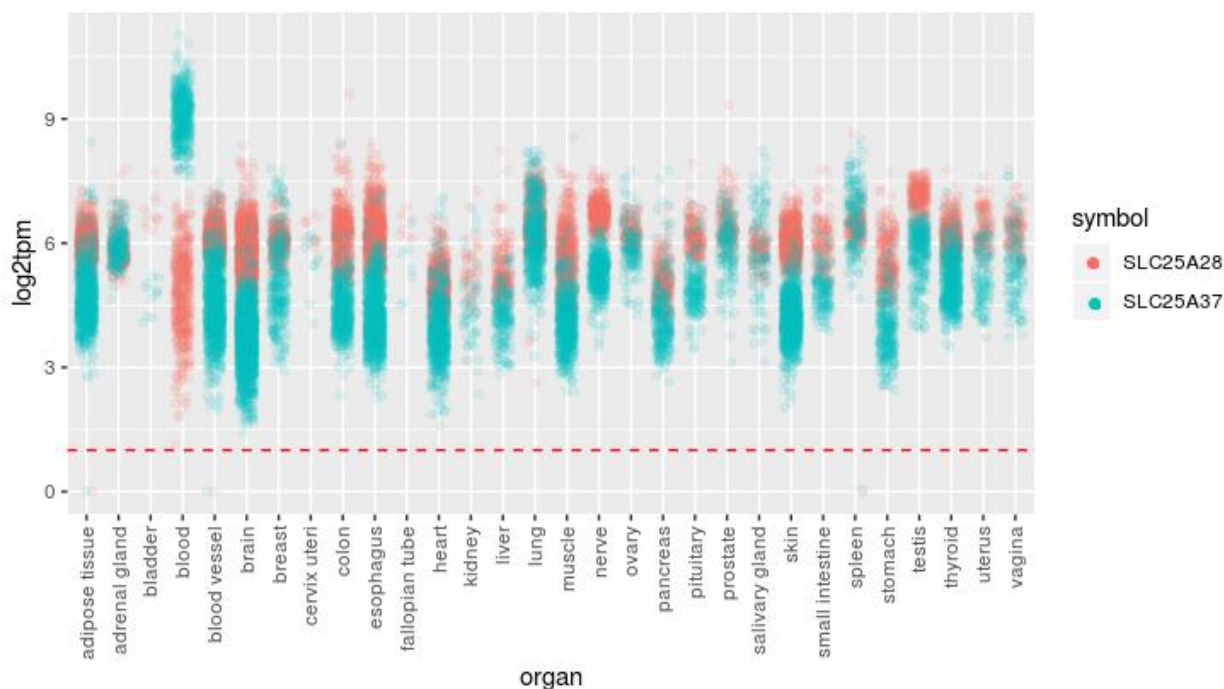
Figure 38: SLC25A28 and SLC25A37 expression in healthy tissues (GTex)

In the initial CRISPR depletion assay, the sensitivity of SNU-761_Cas9_Puro and Hep G2_Cas9_Puro to SLC25A28 knockout was established (Table 9, Figure 39). At first, HuH-6_Cas9_Puro had also been selected as a sensitive cell line, however, as the positive control did not deplete, this cell line was de-selected for further experiments. Probably, Cas9 introduction was not successful. From the tested gRNAs, gRNA_RN334 was not able to knock out SLC25A28.

Table 9: Selected Cell lines for investigating SLC25A37 and SLC25A28 paralog dependency

	Cell lines	Gene Expression SLC25A28 (TPM)	Gene Expression SLC25A37 (TPM)	SLC25A28 AVANA CERES Score
Sensitive cell lines	HuH-6	17.2	7.6	-1.10
	SNU-761	18.8	8.7	-1.10
	Hep G2	28.6	13.6	-1.45
Resistant cell lines	NCI-H716	20.5	22.1	0.111
	KYSE-270	20.8	29.5	-0.052
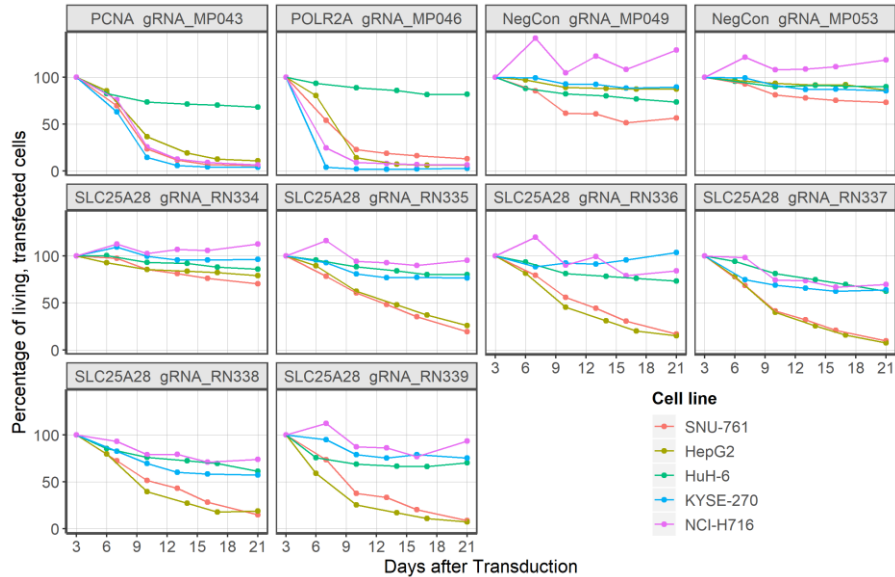
Figure 39: CRISPR depletion assay shows sensitivity of SNU-761 and Hep G2 to SLC25A28 KO
Cells have been transduced with gRNAs targeting SLC25A28, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction.

Next, SNU-761_Cas9_Puro_RIEH and Hep G2_Cas9_Puro_RIEH were produced in order to introduce the overexpression constructs. Codon-optimized overexpression constructs of SLC25A28 and SLC25A37 were introduced into SNU-761_Cas9_RIEH successfully – the transduced Hep G2_Cas9_Puro_RIEH cells died during geneticin selection and due to time constraints, the experiment could not be repeated. The overexpression of the biomarker SLC2A37 and the codon-optimized target gene SLC25A28 was confirmed via western blotting (Figure 40). Again, as SLC25A28 is already expressed in SNU-761, therefore a background signal was expected independent of doxycycline addition.
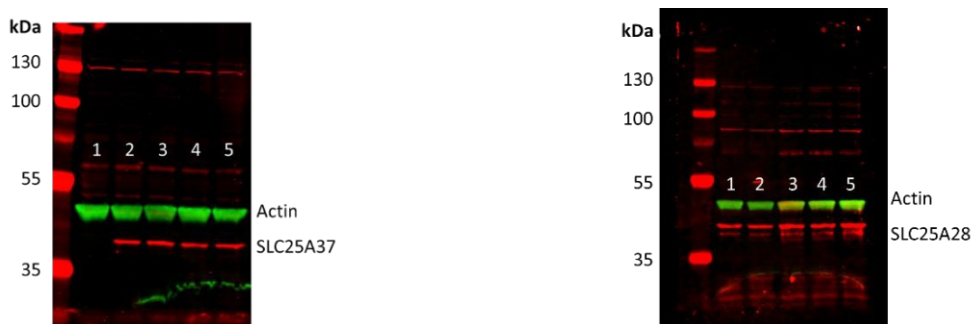
Figure 40: Western blot to confirm overexpression of SLC25A37 and SLC25A28 in SNU-761
Cells were cultivated for 4 days with different levels of doxycycline.
Lane 1 – 5 : 0 µg/mL doxycycline, 0.1 µg/mL doxycycline , 0.5 µg/mL doxycycline, 1 µg/mL doxycycline, 2 µg/mL doxycycline
Left: Overexpression of SLC25A37, Right: Overexpression of SLC25A28

After introduction of the overexpression constructs into SNU-761, the CRISPR depletion assay was repeated. Neither the overexpression of SLC25A28 nor SLC25A37 lead to a rescue of the phenotype (Figure 41 and Figure 42). However, the interpretation of these results is confounded by several experimental problems: In the CRISPR depletion assay, a mild depletion of cells transfected with the negative control was observed. This was not the case before the introduction of the RIEH construct and the overexpression constructs. Furthermore, cell numbers in doxycycline-induced wells were

significantly lower than when no doxycycline was added. This sensitivity was confirmed in via a doxycycline kill curve (Figure 43). At the used doxycycline concentration of 0.5 µg/mL only approximately 40% of cell viability remains when compared to untreated cells. If there is an additional difference in doxycycline sensitivity when the ells are exposed to the stress of transduction, the ratio of transduced to not-transduced cells would not change solely due to sensitivity to SLC25A28 loss-of-function. In conclusion, we cannot make a definitive statement, whether there is a paralog dependency between SLC25A28 and SLC25A37. Looking at the doxycycline kill curve, it is questionable, whether a rescue with an inducible overexpression construct can be achieved in this cell line. When treated with the lowest doxycycline concentration in this experiment, cell viability still dropped to about 60%. Therefore, even less doxycycline should be used in the experiments (< 0.03 µg/mL). Beforehand it needs to be verified that a concentration this low still causes a strong and stable induction.
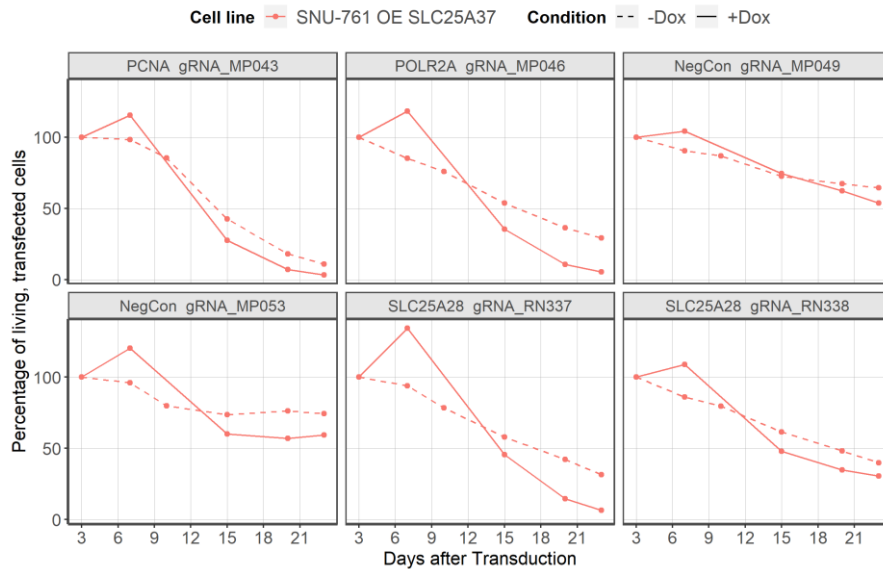
Figure 41: CRISPR depletion assay with inducible overexpression of SLC25A37.
Cells have been transduced with gRNAs targeting SLC25A28, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction. Overexpression of SLC25A37 was induced with 0.5µg/mL doxycycline.

58
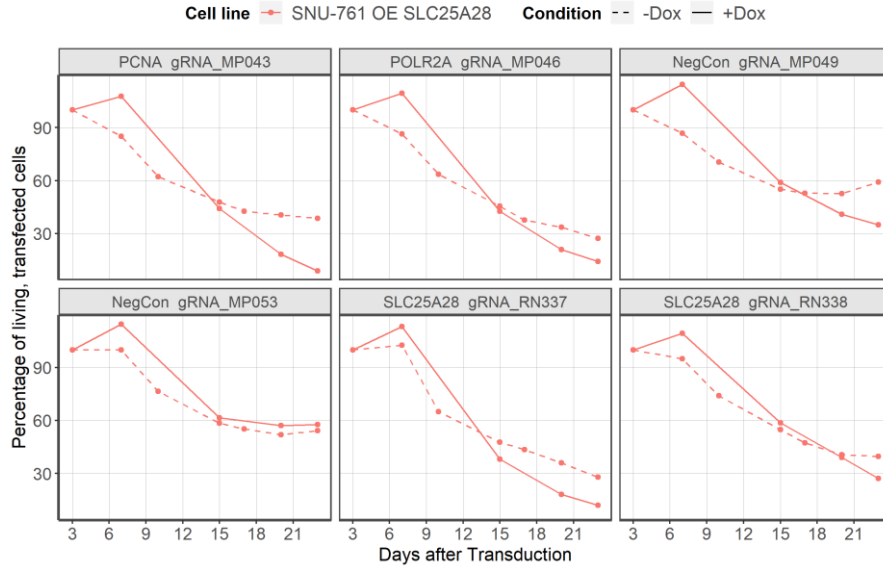
Figure 42: CRISPR depletion assay with inducible overexpression of codon-optimized SLC25A28.
Cells have been transduced with gRNAs targeting SLC25A28, essential genes (PCNA, POLR2A) and non-targeting negative
controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using
FACS and normalized to the ratio of transfected cells three days after transduction. Overexpression of exogenous SLC25A28
was induced with 0.5µg/mL doxycycline.
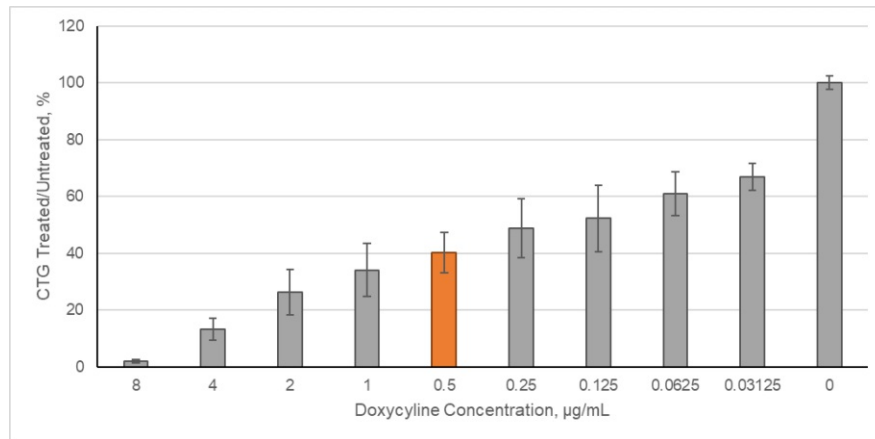
Figure 43: Doxycycline kill curve SNU-761_Cas9_Puro_RIEH_RN186.
Cell viability was measured with the CellTiter-Glo® assay after 7 days cultivation with the addition of doxycycline between
0.03 and 8 µg/mL. The highlighted concentration was employed for CRISPR depletion experiments.

8.7   VPS4B and VPS4A

The putative paralog dependency between VPS4A and VPS4B can be inferred from the AVANA CERES data as cells sensitive to VPS4A KO tend to express low levels of VPS4B (Figure 44). Again, similar to the case of the gene pair SLC25A28 and SLC25A37, in no cell line included in the screen was either paralog completely absent. Possibly, a dose-dependent effect is at play.
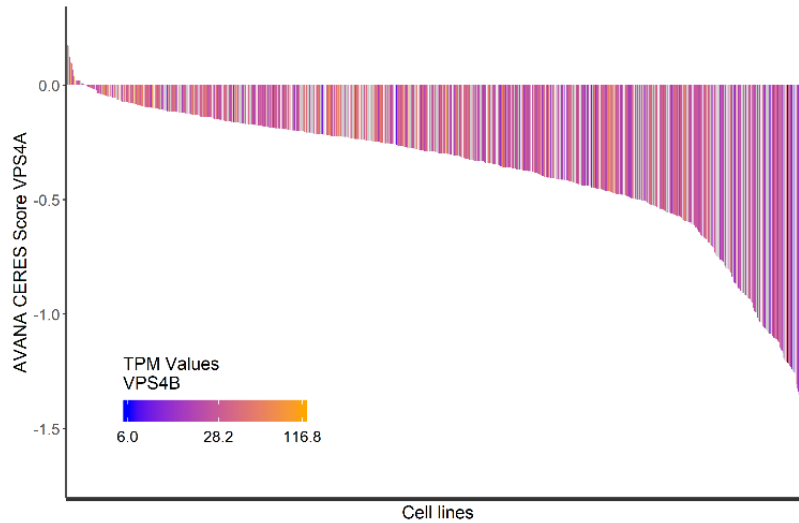
Figure 44: Sensitivity of cancer cell lines to VPS4A knockout (AVANA CERES screen results)

Just like in the case of SLC25A28 and SLC25A37, both paralogs are expressed in all TCGA samples (Figure 45a). Low expression of VPS4B occurs in abdomen, eye and liver tumor tissues. Possibly, this reduced expression suffices to trigger sensitivity to VPS4A knockout as hinted in the AVANA screen.

In all TCGA samples, VPS4B is deep deleted in 1.41% of the cases (https://www.cbioportal.org/). This number is exceeded for instance in colon and rectum carcinoma (2.44% of COAD, 2.60% of COADREAD and 2.42% of READ), in esophageal carcinoma (ESCA) with 3.26%, in head and neck squamous cell carcinoma (HNSC) with 3.64%, in pancreatic carcinoma (PAAD) with 3.26%, in prostate cancer (PRAD) with 3.86%, in stomach carcinoma (STAD) with 3.17% and in thyroid carcinoma (TGCT) with 2.67%. Similarly to PRPS2, VPS4B is located relatively close to the end of the chromosome 18 and the deep deletion frequency elucidated by the GISTIC algorithm increases the closer the gene is located to the end of the chromosome (Figure 45b). However, different to PRPS2 on the X chromosome, there are essential genes between VPS4B and the end of the chromosome. Therefore, it is likely that the loss of the gene is only heterozygous. This is also in accordance with the fact, that the cell lines in the AVANA screen as well as the TCGA samples show low-level but not absent expression.

In their paper, Szymanska et al. claimed that loss of VPS4B frequently occurs in various cancer types, highlighting colorectal cancer (137). Indeed, we also found that according to the GISTIC data, deep deletions of VPS4B occur relatively often in this cancer type, however, we found that these cannot be interpreted as homozygous deletions. When looking at gene expression, we could not find cancer samples or cancer lines which do not express VPS4B at least at a low level. They furthermore assumed that the downregulation of VPS4B might be caused by a passenger deletion of neighboring tumor suppressors such as DCC, SMAD2 and SMAD4. We do not think that this is the case, as the distance between these genes and VPS4B is quite large. The smallest difference, between VPS4B and DCC, is

nearly 10 Mbp (mega base pairs). Furthermore, at least two essential genes WDR7 and NARS are located in between. If VPS4B deficiency was due to a passenger deletion, it would more likely be caused by the tumor suppressors PHLPP1 (155) or SERPINB5 (156). The later gene is located just after VPS4B.

Figure 45: VPS4B Expression in cancer (TCGA).
Top: Expression in various cancer types, Bottom left: Deep deletion frequency in all cancer types of genes encoded on chromosome 18

Both paralogs are relatively equally expressed in all healthy tissues (Figure 46). Only in blood cells, VPS4A is the dominant form while VPS4B is not expressed in some samples. Again, should VPS4A emerge as a target for cancer therapy, potential negative consequences for blood cells are expected.
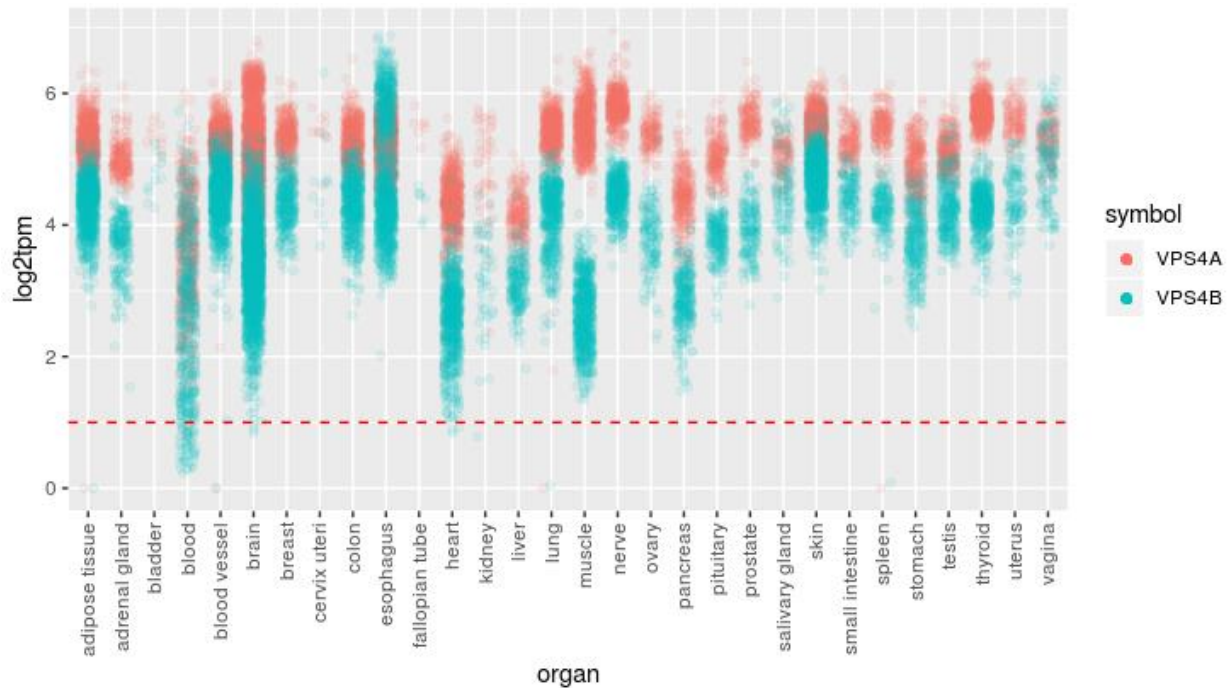
Figure 46: VPS4A and VPS4B expression in healthy tissues (GTex)

In order to investigate the sensitivity of the selected cell lines to the knockout of VPS4A, CRISPR depletion assays were performed (Table 10). The sensitive cell lines 59M_Cas9_Puro and NCI-H2110_Cas9_Puro did deplete upon transduction with gRNAs targeting VPS4A with a similar kinetic compared to the positive controls. As the cell line 59M_Cas9_Puro only grew very slowly, the number of cell splitting over the course of the CRISPR experiment was reduced. Resistant DMS 53_Cas9_Puro and SK-BR-3_Cas9_Puro cells transduced with gRNA_RN364 depleted as well, potentially due to off-target effects. When the sequence of this gRNA was investigated with the Human BLAT search (https://www.genome.ucsc.edu/cgi-bin/hgBlat), no perfect off-target matches were detected. However, eight hits were detected in which 16-19 bases match to the 20 base long gRNA and that probably caused the off-target depletion. For future CRISPR experiments, gRNA_RN362 and gRNA_363 were chosen. gRNA_RN365 was considered, but ultimately omitted as it potentially caused some off-target effect in SK-BR-3_Cas9_Puro. Unfortunately, the rescue experiments could not be completed within the timeframe of this thesis. 59M_Cas9_Puro_RIEH was produced, but the overexpression construct could not be introduced due to the slow cell growth. Similarly, NCI-H2110_Cas9_Puro_RIEH was produced, however, the cells repeatedly died after transduction with the ecotropic virus and geneticin selection. Therefore, conclusive experiments on this paralog dependency are still outstanding.

Table 10: Selected Cell lines for investigating VPS4B and VPS4A paralog dependency

	Cell lines	Gene Expression VPS4A (TPM)	Gene Expression VPS4B (TPM)	VPS4A AVANA CERES Score
Sensitive cell lines	NCI-H2110	13.8	7.4	-1.26
	59M	21.1	9.2	-1.24
Resistant cell lines	SK-BR-3	27.3	21.2	0.12
	DMS 53	35.7	90.3	0.10
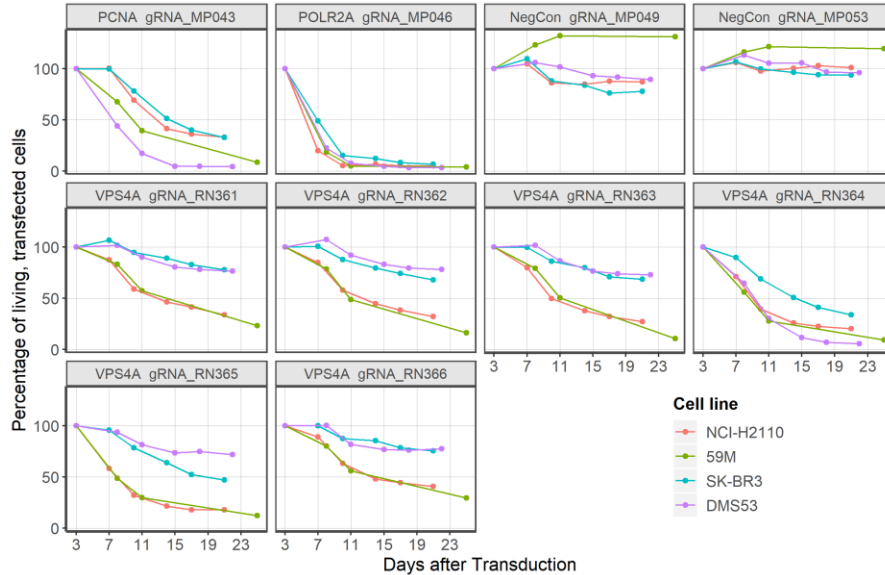
Figure 47: CRISPR depletion assay shows sensitivity to VPS4A KO in cell lines which do not express VPS4B
Cells have been transduced with gRNAs targeting VPS4A, essential genes (PCNA, POLR2A) and non-targeting negative controls. As the gRNAs are co-expressed with GFP, the percentage of living, transfected cells were measured over time using FACS and normalized to the ratio of transfected cells three days after transduction.

9 Discussion

Paralog dependencies are a subgroup of synthetic lethal genetic dependencies. Paralog dependency is defined such that the simultaneous loss of two or more genes, which have emerged from gene duplication events, leads to a detectable phenotype, while the loss of only one of these genes does not cause this phenotype. In the context of cancer, therapeutically exploitable paralog dependencies elicit cell death or strong impairments in proliferative capacity as their combined loss-of-function phenotypes. If in a cancer cell one of these genes is lost by means of mutation, promoter methylation or other means, the other paralog becomes a putative target for targeted cancer therapy. Healthy cells, which should express both paralogs should not be effected this way. In this Master thesis project, 6 putative paralog dependencies were investigated in order to determine if these gene pairs (a) show evidence for functional molecular redundancy and (b) constitute an interesting opportunity for targeted cancer therapy.

The experiments conducted in the course of this thesis validate the paralog dependency between DNAJC15 and DNAJC19 and suggest that these two proteins are functionally redundant in the assay conditions that we chose. The proliferative capacity of NCI-H1975 cells, sensitive to DNAJC19 loss-of-function could be rescued via the overexpression of the biomarker paralog DNAJC15. In the second sensitive cell line, CAL-12T, the rescue was not successful. However, as the CRISPR-mediated DNAJC19 loss-of-function phenotype could not be rescued by providing exogenously expressed DNAJC19, unspecific effects might dominate in this experiment. As overexpression of codon-optimized DNAJC19 should compensate for the CRISPR/Cas9 mediated loss of wildtype DNAJC19, we reasoned that the selected gRNA depleted CAL-12T cells due to an off-target effect. Unfortunately, only one gRNA targeting DNAJC19 was available, as with the other two tested gRNAs no effects in DNAJC19 sensitive cells were observed. As DNAJC19 is a relatively short gene with only 6066 bases, designing gRNAs without off-target matches imposes a challenge. As an alternative method, siRNA mediated knockdown experiments are planned for the future. Furthermore, additional gRNAs could be used to further corroborate the findings.

As demonstrated in the bioinformatics analysis, the biomarker paralog DNAJC15 is located on chromosome 13 in proximity to the known tumor suppressor RB1. However, between these two genes lies the core-essential gene TPT1. Therefore, it is unlikely, that homozygous passenger deletions of DNAJC15 occur in association with the loss of RB1 in cancer cells. An alternative explanation for the loss of DNAJC15 expression in cancer cells is DNA methylation. We could demonstrate, that there is a correlation between the degree of DNA methylation on CpG islands and gene expression of DNAJC15 in TCGA tumor samples which is in accordance with previous publications (80; 89; 92; 94; 95). We conclude that DNA methylation is the prime mechanism for expression of DNAJC15 in tumors, as absent expression is not observed without concurrent DNA hypermethylation. A potential caveat should however be stressed when considering DNAJC19 as a target in DNAJC15 hypermethylated tumors: The DNA methylation might be reversible (157), providing the cancer cells a direct escape mechanism from sensitivity to DNAJC19 by means of re-expression of DNAJC15. However, this epigenetic inactivation of DNAJC15 has been connected with chemotherapeutic resistance (91; 93; 96). Hence, we suggest that targeting DNAJC19 could synergize with standard of care chemotherapeutics: As soon as the cancer cells develop chemotherapeutic resistance through silencing DNAJC15, they become susceptible to the loss of DNAJC19. In future experiments, it is planned to investigate whether the silencing of DNAJC15 is reversible in NCI-H1975 as described by Lindesy et al. (91). Treatment with de-methylating agents

should reverse the sensitivity to DNAJC19 loss-of-function similar to the induction of codon-optimized DNAJC19 expression.

In accordance with the strong correlation in the AVANA CERES screening data, our experiments document a paralog dependency between RPP25 and RPP25L. This could be shown with the cell line U-2 OS, which is sensitive to RPP25L CRISPR-mediated knockout but can be rescued via the overexpression of RPP25. In the second investigated cell line, KYSE-150, no successful rescue was achieved. For both cell lines single cell cloning was necessary to avoid leaky expression of RPP25. Hence, it is possible that the reduction of genetic variability caused off-target sensitivity in the KYSE-150 clone. In addition, improper expression levels of the biomarker paralog might underlie the insufficient rescue. Conceivably, testing new KYSE-150 clones might yield a cell line in which the paralog dependency could be demonstrated. Future experiments could determine the ability of codon-optimized RPP25L to rescue the effects observed in KYSE-150_Cas9_Puro_RIEH and U-2 OS_Cas9_Puro_RIEH cells. This way, it could be confirmed, if the depletion of transduced cells in the CRISPR assays is indeed caused by the knockout of RPP25L and is not due to an off-target effect. However, so far there is no reason to question the specificity of the used gRNAs as no exact sequence matches were found at other sites in the human genome. When considering RPP25L for targeted cancer therapy, it needs to be taken into account that deep deletion of the biomarker RPP25 rarely occur. Nevertheless, RPP25 expression is lost in multiple cancer tissues such as in brain, liver and kidney tumors. One reason is silencing of RPP25 through DNA hypermethylation, as there is correlation between CpG methylation and RPP25 expression in TCGA samples. However, this explanation is insufficient as not all RPP25 deficient samples are methylated. For instance epigenetic histone modifications could also account for the lack of expression. Unfortunately, RPP25 is not expressed ubiquitously. Therefore, unwanted off-target effects might occur in blood, brain, heart, muscle and liver cells. Only a tumor directed delivery of a putative RPP25L inhibitor, sparing tissues in which RPP25 is not expressed, would be a therapeutic option.

The paralog dependency between PAPSS2 and PAPSS1 could not be confirmed by our experiments. This finding raises concerns about the effects observed in the AVANA CRISPR screens (138). Already in the initial CRISPR depletion assay, the proliferation of expected sensitive cell lines with no expression of PAPSS2 were not affected by CRISPR mediated loss-of-function of PAPSS1. This was surprising, as in literature PAPSS1 and PAPSS2 are described as the only two enzymes synthesizing PAPS, which in turn is the central substrate for all sulfation pathways (112; 116). There are multiple explanations for the experimental results: First, it is possible that none of the tested gRNAs efficiently cut in the PAPSS1 locus. This could for example be verified via cell sorting after transduction with the gRNA co-expressing GFP and subsequent western blotting or sequencing of the PAPSS1 locus. PAPSS1 should no longer be expressed in the transduced cells. Second, the experimental time points for our measurement might have been sub-optimal. They might have been too soon after the start of the CRISPR-mediated mutagenesis such that PAPS stores were not depleted and thus, the cells were not affected by the lack of PAPSS1 and PAPSS2. This explanation seems unlikely, as the cells were monitored for 21 days after transduction, just like in the AVANA CERES screen (138). Another explanation would be that the assumption, that only PAPSS1 and PAPSS2 produce PAPS is wrong and that in fact other proteins can fulfill this function. In order to examine this hypothesis, sulfation of proteins, metabolites or xenobiotics cells before and after knockout of PAPSS1 could be investigated. For instance, protein tyrosine sulfation could be detected via sulfotyrosine specific antibodies or mass spectrometry of sulfated peptides (158). Finally, the cell proliferation was potentially not influenced by the knockout of PAPSS1 because PAPS or

sulfation is just not necessary for cellular survival. As sulfation is a metabolic process that participates in a plethora of biological cellular functions, this again seems unlikely. Leung et al. reported increased sensitivity of cancer cells to chemotherapeutics like cisplatin or topotecan when PAPSS1 is silenced (126; 127). The extent of the sensitization is also dose-dependent on PAPSS1. The cell lines used in their study all expressed both PAPSS1 and PAPSS2. Possibly, when both genes are missing the sensitivity to DNA damaging agent is even more pronounced. As the bioinformatics analysis clearly demonstrated that PAPSS2 is frequently deleted as a passenger deletion of PTEN, it represents a high-potential biomarker. Therefore, further investigation of this gene pair is warranted, although results of this thesis suggesting that these genes do probably not lend themselves for a monotherapy.

Unfortunately, there was not enough time for conclusive experiments to unambiguously validate the paralog dependencies PRPS1/2, SLC25A28/37 and VPS4A/B.

For the validation of the paralog dependency between PRPS1 and PRPS2 only one final experiment is required: demonstrating the rescue of JOPACA-1 cells from sensitivity to PRPS1 CRISPR-mediated loss-of-function though the overexpression of PRPS2. All preliminary experiments with this cell line were successful. Furthermore, the second sensitive cell line SK-BR-3_Cas9_Puro_RIEH is already available and can be tested in CRISPR depletion assays after the introduction of overexpression constructs. In cancer samples originating from the cerebrum, liver or adrenal gland, a loss of PRPS2 expression occurs according to TCGA data. This might be explainable by the loss of the telomeric end of chromosome X, where PRPS2 is located. This deletion might even be homozygous, as no essential genes are located between the telomeric end of the chromosome and PRPS2. Differences in the regulation of PRPS1 and PRPS2 may offer an alternative explanation: PRPS2 but not PRPS1 expression is heavily regulated on a translational level (70). Cunningham et al. demonstrated an upregulation of exclusively PRPS2 in Myc-overexpressing cancer cells. Conceivable, this regulation could also be reversed. Unfortunately, PRPS2 is frequently not expressed in multiple healthy tissues including muscle, heart, brain and blood cells. Presumably, these tissues rely on the expression of PRPS1 and may therefore be sensitive to a potential cancer therapy targeting this protein. Different to the other paralog dependencies discussed in this Master thesis, there is a third paralog gene PRPS1L1. This may lead to complex adaptive mechanisms to a putative PRPS1 inhibitor, as it is conceivable that cancer cell lines might upregulate the expression of this gene to avoid sensitivity to PRPS1 despite the loss of PRPS2.

As the cell line SNU-761, used in the CRIPSR depletion assays to show the rescue from SLC25A28 sensitivity through SLC25A37 overexpression, was sensitive to doxycycline, not definitive conclusion on this paralog dependency can be drawn. The sensitivity was too severe to be ignored, hence it was not possible to distinguish between reduced viability due to the knockout of SLC25A28 or the induction reagent. In further experiments with SNU-761, the doxycycline level needs to be reduced drastically to avoid sensitivity, however, this way the concentration might fall below the threshold for stable induction. Due to time constraints, the overexpression constructs was not introduced into the other investigated sensitive cell line Hep G2. On the basis of gene expression in TCGA samples, it seems as if both paralogs are present in all cancer cells. However, the biomarker SLC25A37 expression level is frequently reduced in tumor tissues. This could be caused by a heterozygous loss of an arm of chromosome 8, where SLC25A37 together with numerous tumor suppressors is located. This reduced expression may convey sensitivity to SLC25A28 exploitable for cancer therapy. However, the level of SLC25A37 at which cancer cells escape this sensitivity is crucial to estimate the potential of this approach. In healthy tissues, both

SLC25A28 and SLC25A37 are expressed equally at relatively high levels. Therefore, no significant side effects are expected when SLC25A28 is targeted.

Despite multiple attempts, introduction of the overexpression constructs of VPS4A and VPS4B into NCI-2110_Cas9_Puro_RIEH was not successful, most likely due to high stress upon transduction. Unfortunately, ectotropization of the other sensitive cell line 59M was not achieved mostly due to slow growth. Nevertheless, Szymanska et al. already provided strong indication, that these two genes are in a paralog dependency (137). Despite their assertion that VPS4B is frequently lost in cancer cells, especially colorectal cancer, we found that at least low level expression is present in all TCGA samples. Furthermore, they assumed that lack of expression of VPS4B in cancer cells is due to passenger alterations caused by the deletion of neighboring tumor suppressors such as DCC, SMAD2, and SMAD4. However, our analysis of the deep deletion frequency in cancer cells depending on chromosomal position showed that, if a passenger deletion is the cause for downregulation of VPS4B, the respective tumor suppressor gene is more likely PHLPP1 or SERPINB5. Still, low expression, possibly due to heterozygous deletion, of VPS4B could suffice to make the tumor cells more sensitive to a therapy targeting VPS4A than healthy cells. For this purpose, it would be interesting, at which level VPS4B expression rescues the cells from VPS4A sensitivity. As VPS4B is frequently absent in healthy blood cells, these should be carefully observed when targeting VPS4A.

To conclude, paralog pairs offer an attractive method to identify synthetic lethal interactions. Interesting targets can be found using screening data like the AVANA CRISPR screen in conjunction with gene expression data. Furthermore, screens employing both Cas9 and Cas12a enable simultaneous knockout of two genes and therefore enable the direct observation of synthetic lethality (159). In this thesis, we verified putative paralog dependencies through overexpression of the biomarker paralog in cell lines sensitive to knockout of the target paralog. If the two genes are functionally redundant, loss of this sensitivity is expected to occur. In a next step, the impact on cells when both paralogous genes are lost can be analyzed. Possible consequences include reduced proliferation, senescence or apoptosis. Moreover, a targeting strategy for the potential development of specific inhibitory agents needs to be developed. This raises the biggest issue when it comes to paralog dependencies. In order to employ this concept for target cancer therapy, it is of critical importance that a potential drug only inhibits the intended paralog and not multiple members of a given paralog family (20). As these paralog proteins are very similar due to their evolutionary relatedness, this presents a major challenge. It is plausible that cells, which rely on only one instead of two paralogs, are more sensitive to a drug even if it does not discriminate between the paralogs. This was demonstrated for ENO1 and ENO2, where the lethal dose of a compound targeting enolases was significantly lower for ENO1 deficient cells (51). However, this approach is closer to classical chemotherapies than targeted cancer therapies. Another factor, which needs close observation, are healthy tissues, which do not express the biomarker paralog. Here, side effects may occur. However, these could be prevented through tissue specific drug delivery (160). The next crucial step for the two confirmed targets, DNAJC19 and RPP25L, will therefore be structural biology studies for drug development.

10 Acknowledgements

11 References

1. **World Health Organization.** *WHO report on cancer: setting priorities, investing wisely and providing care for all.* Geneva : Licence: CC BY-NC-SA 3.0 IGO, 2020.

2. *The cancer genome.* **Stratton, M. R., Campbell, P. J. and Futreal, P. A.** doi:10.1038/nature07943, 2009, Nature, Vol. 458, pp. 719 - 724.

3. *Searching for synthetic lethality in cancer.* **Brough, R., et al., et al.** 2011, Current Opinion in Genetics & Development, p. DOI 10.1016/j.gde.2010.10.009.

4. *Synthetic Lethality in Lung Cancer—From the Perspective of Cancer Genomics.* **Shimomura, I., Yamamoto, Y. and Ochiya, T.** 2019, Medicines, p. doi:10.3390/medicines6010038.

5. *Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction?* **Baylin, S. B. and Ohm, J. E.** doi:10.1038/nrc1799, 2006, Nature Review, Vol. 6.

6. *Human Cancer Syndromes: Clues to the Origin and Nature of Cancer.* **Fearon, E. R.** DOI: 10.1126/science.278.5340.1043, 1997, Science, Vol. 278, pp. 1043 - 1050.

7. *Oncogenes and Tumor Suppressor Genes.* **Lee, E. Y.H.P and Muller, W. J.** doi: 10.1101/cshperspect.a003236, 2010, Cold Spring Harb Perspect Biol.

8. *An overview of targeted cancer therapy.* **Padma, V. V.** DOI 10.7603/s40681-015-0019-4, 2015, BioMedicine, Vol. 5, pp. 1 - 6.

9. *Clinical Cancer Advances 2020: Annual Report on Progress Against Cancer From the American Society of Clinical Oncology.* **Markham et al., M. J.** https://doi.org/10.1200/JCO.19.03141, 2020, Journal of Clinical Oncology, Vol. 38, pp. 1081-1101.

10. *Principles of cytotoxic chemotherapy.* **Lind, M. J.** https://doi.org/10.1016/j.mpmed.2007.10.003, 2007, Medicine, Vol. 39, pp. 19 - 23.

11. *The concept of synthetic lethality in the context of anticancer therapy.* **Kaelin, W. G. Jr.** 2005, Nature Reviews (Cancer), p. doi:10.1038/nrc1691.

12. *Molecular-targeted therapy for cancer.* **Dong, B. and Zhu, Y.-M.** DOI: 10.5732/cjc.009.10313, 2010, Chinese Journal of Cancer, Vol. 29, pp. 340 - 345.

13. *Systematic identification of genomic markers of drug sensitivity in cancer cells.* **Garnett et al., M. J.** doi:10.1038/nature11005, 2012, Nature, Vol. 483, pp. 570 - 575.

14. *Synthetic Lethality in Cancer Therapeutics.* **Beijersbergen, R. L., Wessels, L. F.A and Bernards, R.** 2017, Annu. Rev. Cancer Biol., pp. doi:10.1146/annurev-cancerbio-042016-073434.

15. *Synthetic dosage lethality in the human metabolic network is highly predictive of tumor growth and cancer patient survival.* **Megchelenbrink, W., et al., et al.** 2015, PNAS, p. www.pnas.org/cgi/doi/10.1073/pnas.1508573112.

16. *Synthetic lethality in lung cancer and translation to clinical therapies.* **Leung, A. W. Y., et al., et al.** DOI 10.1186/s12943-016-0546-y, 2016, Molecular Cancer.

17. *Development of Synthetic Lethality Anticancer Therapeutics.* **Fang, B.** 2014, Journal of Medicinal Chemistry, pp. dx.doi.org/10.1021/jm500415t | J. Med. Chem. 2014, 57, 7859–7873.

18. *Genome-scale analysis identifies paralog lethality as a vulnerability of chromosome 1p loss in cancer.* **Viswanathan et al., S. R.** 2018, Nature Genetics, pp. https://doi.org/10.1038/s41588-018-0155-3.

19. *Paralog buffering contributes to the variable essentiality of genes in cancer cell lines.* **De Kegel, B. and Ryan, C. J.** 2019, PLoS Genet, p. https://doi.org/10.1371/journal.pgen.1008466.

20. *Consequences of Nonadaptive Alterations in Cancer.* **Kamb, A.** 2003, Molecular Biology of the Cell, pp. doi/10.1091/mbc.E02–11–0732.

21. *Molecular Mechanisms of Paralogous Compensation and the Robustness of Cellular Networks.* **Diss, G., et al., et al.** 2014, J. Exp. Zool. (Mol. Dev. Evol.), p. DOI: 10.1002/jez.b.22555.

22. *Paralog dependency indirectly affects the.* **Dandage, R. and Landry, C. R.** 2019, Mol Syst Biol., p. DOI 10.15252/msb.20198871.

23. *Synthetic lethality of PARP inhibition in cancers lacking BRCA1 and BRCA2 mutations.* **Dedes, K. J., et al., et al.** 2011, Cell Cycle, p. DOI: 10.4161/cc.10.8.15273.

24. *The Cancer Genome Atlas Pan-Cancer Analysis Project.* **Weinstein, J. N., et al., et al.** 2013, Nat Genet., p. doi:10.1038/ng.2764.

25. *Recombination and variability in populations of Drosophila pseudoobscura.* **Dobzhansky, T.** 1945, Genetics of Natural Populations XIII.

26. *Synthetic lethality and semi-lethality among functionally related mutants of Drosophila melanogaster.* **Lucchesi, J. C.** 1967, Genetics 59.

27. *Synthetic Lethality and Cancer: Cohesin and PARP at the Replication Fork.* **O'Neil, N. J., van Pel, D. M. and Hieter, P.** doi:10.1016/j.tig.2012.12.004., 2013, Trends Genet.

28. *Synthetic lethality: emerging targets and opportunities in melanoma.* **Thompson, N., Adams, D. J. and Ranzani, M.** doi: 10.1111/pcmr.12573, 2017, Pigment Cell & Melanoma Research.

29. *RNA Interference and MicroRNA-Mediated Silencing.* **Fischer, S. E.J.** 2015, Curr. Protoc. Mol. Biol., p. doi: 10.1002/0471142727.mb2601s112.

30. *A Perspective on the Future of High-Throughput RNAi Screening: Will CRISPR Cut Out the Competition or Can RNAi Help Guide the Way?* **Taylor, J. and Woodcock, S.** 2015, Journal of Biomolecular Screening, p. DOI: 10.1177/1087057115590069.

31. *Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening.* **McDonald III, E. R., et al., et al.** http://dx.doi.org/10.1016/j.cell.2017.07.005, 2017, Cell, Vol. 170, pp. 577-592.

32. *Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy.* **Farmer, H., et al., et al.** 2005, Nature, p. doi:10.1038/nature03445.

33. *Simultaneous Targeting of PARP1 and RAD52 Triggers Dual Synthetic Lethality in BRCA-Deficient Tumor Cells.* **Sullivan-Reed et al., K.** 2018, Cell Reports, p. https://doi.org/10.1016/j.celrep.2018.05.034.

34. *PARP inhibitors in pancreatic cancer: molecular mechanisms and clinical applications.* **Zhu, H., et al., et al.** 2020, Molecular Cancer, pp. https://doi.org/10.1186/s12943-020-01167-9.

35. *Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase.* **Bryant, H. E., et al., et al.** 2005, Nature, p. https://doi.org/10.1038/nature03443.

36. *Synthetic lethal targeting of PTEN mutant cells with PARP inhibitors.* **Mendes-Pereira, A. M., et al., et al.** DOI 10.1002/emmm.200900041, 2009, EMBO Mol Med, Vol. 1.

37. *ENO1 (Enolase 1, (alpha)).* **Trojanowicz, B., Hoang-Vu, C. and Sekulla, C.** DOI: 10.4267/2042/44795, 2010, Atlas Genet Cytogenet Oncol Haematol.

38. *Characterization of alpha-alpha, beta-beta, gamma-gamma and alpha-gamma human enolase isoenzymes, and preparation of hybrid enolases (alpha-gamma, beta-gamma and alpha-beta) from homodimeric forms.* **Shimizu, A., Suzuki, F. and Kato, K.** 1983, Biochimica et Biophysica Acta.

39. *ENO1 gene product binds to the c-myc promoter and acts as a transcriptional repressor: relationship with Myc promoter-binding protein 1 (MBP-1).* **Feo, S., et al., et al.** doi:10.1016/S0014-5793(00)01494-0, 2000, FEBS Letters, Vol. 473, pp. 47-52.

40. *The Biological Significance and Regulatory Mechanism of c-Myc Binding Protein 1 (MBP-1).* **Liu, Z., et al., et al.** doi:10.3390/ijms19123868, 2018, Int. J. Mol. Sci.

41. *Silencing of ENO1 by shRNA Inhibits the Proliferation of Gastric Cancer Cells.* **Qiao, H., et al., et al.** DOI: 10.1177/1533033818784411, 2018, Technology in Cancer Research & Treatment, Vol. 17, pp. 1-9.

42. *Alpha-enolase promotes cell glycolysis, growth, migration, and invasion in non-small cell lung cancer through FAK-mediated PI3K/AKT pathway.* **Fu, Q.-F., et al., et al.** DOI 10.1186/s13045-015-0117-5, 2015, Journal of Hematology & Oncology.

43. *Diagnostic value of α-enolase expression and serum α-enolase autoantibody levels in lung cancer.* **Zhang, L., Wang, H. and Dong, X.** http://dx.doi.org/10.1590/S1806-37562016000000241, 2018, J Bras Pneumol., Vol. 44, pp. 18-23.

44. *Enolase-1 is a therapeutic target in endometrial carcinoma.* **Zhao, M., et al., et al.** doi:10.18632/oncotarget.3639, 2015, Oncotarget, Vol. 6, pp. 15610-15627.

45. *Impact of genomic stability on protein expression in endometrioid endometrial cancer.* **Lomnytska, M., et al., et al.** doi:10.1038/bjc.2012.67, 2012, British Journal of Cancer, Vol. 106, pp. 1297 – 1305.

46. *Enolase-a Is Frequently Down-Regulated in Non-Small Cell Lung Cancer and Predicts Aggressive Biological Behavior.* **Chang, Y. S., et al., et al.** 2003, Clinical Cancer Research, Vol. 9, pp. 3641–3644.

47. *Passenger Deletions Generate Therapeutic Vulnerabilities in Cancer.* **Muller, F. L., et al., et al.** doi:10.1038/nature11331, 2012, Nature, Vol. 488, pp. 337–342.

48. *Evidence that MIG-6 is a tumor-suppressor gene.* **Zhang, Y-W, et al., et al.** https://doi.org/10.1038/sj.onc.1209790, 2007, Oncogene , Vol. 26, pp. 269–276.

49. *The Quest for the 1p36 Tumor Suppressor.* **Bagachi, A. and Mills, A. A.** DOI: 10.1158/0008-5472.CAN-07-2095, 2008, Cancer Res, Vol. 68, pp. 2551-2556.

50. *CAMTA1, a 1p36 Tumor Suppressor Candidate, Inhibits Growth and Activates Differentiation Programs in Neuroblastoma Cells.* **Henrich, K.-O., et al., et al.** DOI: 10.1158/0008-5472.CAN-10-3014, 2011, Cancer Res, Vol. 71, pp. 3142-3151.

51. *Passenger deletion of ENO1 as a collateral lethality target in cancer. .* **Lin, Y.-H., et al., et al.** DOI: 10.1158/1538-7445.AM2016-3837, 2016, Proceedings of the 107th Annual Meeting of the American Association for Cancer Research, pp. 16-20.

52. *Roles of RNase P and Its Subunits.* **Jarrous, N.** http://dx.doi.org/10.1016/j.tig.2017.06.006, 2017, Trends in Genetics, Vol. 33, pp. 594 - 603.

53. *Human RNase P ribonucleoprotein is required for formation of initiation complexes of RNA polymerase III.* **Serruya, R., et al., et al.** doi: 10.1093/nar/gkv447, 2015, Nucleic Acids Research, Vol. 43, pp. 5442–5450.

54. *Archaeal/Eukaryal RNase P: subunits, functions and RNA diversification.* **Jarrous, N. and Gopalan, V.** doi:10.1093/nar/gkq701, 2010, Nucleic Acids Research, Vol. 38, pp. 7885–7894.

55. *tRNA processing defects induce replication stress and Chk2-dependent disruption of piRNA transcription.* **Molla-Herman, A., et al., et al.** DOI 10.15252/embj.201591006, 2015, The EMBO Journal, Vol. 34, pp. 3009 - 3027.

56. *RNase MRP Cleaves the CLB2 mRNA To Promote Cell Cycle Progression: Novel Method of mRNA Degradation.* **Gill, T., et al., et al.** DOI: 10.1128/MCB.24.3.945–953.2004, 2004, Mol. Cell. Biol, Vol. 24, pp. 945–953.

57. *Crystal Structure of Human Rpp20/Rpp25 Reveals Quaternary Level Adaptation of the Alba Scaffold as Structural Basis for Single-stranded RNA Binding.* **Chan, C. W., Kiesel, B. R. and Mondragón, A.** https://doi.org/10.1016/j.jmb.2018.03.029, 2018, J Mol Biol, Vol. 430, pp. 1403–1416.

58. *Heterodimerization regulates RNase MRP/RNase P association, localization, and expression of Rpp20 and Rpp25.* **Welting, T. J. M., et al., et al.** http://www.rnajournal.org/cgi/doi/10.1261/rna.237807., 2007, RNA, Vol. 13, pp. 65 – 75.

59. *Differential association of protein subunits with the human RNase MRP and RNase P complexes.* **Welting, T. M. J., et al., et al.** http://www.rnajournal.org/cgi/doi/10.1261/rna.2293906., 2006, RNA, Vol. 12, pp. 1373–1382.

60. *Purification and characterization of Rpp25, an RNA-binding protein subunit of human ribonuclease P.* **Guerrier-Takada, C., et al., et al.** DOI: 10.1017.S1355838202027954, 2002, RNA, Vol. 8, pp. 290-295.

61. *Heterodimerization of the human RNase P/MRP subunits Rpp20 and Rpp25 is a prerequisite for interaction with the P3 arm of RNase MRP RNA.* **Hands-Taylor, K. L. D., et al., et al.** doi:10.1093/nar/gkq141, 2010, Nucleic Acids Research, Vol. 38, pp. 4052–4066.

62. *Next-generation characterization of the Cancer Cell Line Encyclopedia.* **Ghandi, M. and Huang et al., F. W.** https://doi.org/10.1038/s41586-019-1186-3, 2019, Nature, Vol. 569, pp. 503–508.

63. *PRPS1 Mutations: Four Distinct Syndromes and Potential Treatment.* **de Brouwer, A. P. M., et al., et al.** DOI 10.1016/j.ajhg.2010.02.024., 2010, The American Journal of Human Genetics, Vol. 86, pp. 506 - 518.

64. *Overexpression, Purification, and Characterization of Recombinant Human 5-Phosphoribosyl- 1-pyrophosphate Synthetase Isozymes I and II.* **Nosal, J. M. and Switzer, R. L.** 1993, Journal of Biological Chemistry, Vol. 268, pp. 10168 - 10175.

65. *Mammalian phosphoribosyl pyrophosphate synthetase.* **Tatibana, M., et al., et al.** https://doi.org/10.1016/0065-2571(94)00017-W, 1995, Advan. Enzyme Regul., Vol. 35, pp. 229 - 249.

66. *Cloning of cDNAs for Human Phosphoribosylpyrophosphate Synthetases 1 and 2 and X Chromosome Localization of PRPS1 and PRPS2 Genes.* **Becker, M. A., et al., et al.** DOI: 10.1016/0888-7543(90)90043-t, 1990, Genomics, Vol. 8, pp. 555 - 561.

67. *Tissue-differential expression of two distinct genes for phosphoribosyl pyrophosphate synthetase and existence of the testis-specific transcript.* **Taira, M., et al., et al.** https://doi.org/10.1016/0167-4781(89)90040-7, 1988, Biochimica et Biophysica Acta, Vol. 1007, pp. 203 - 208.

68. *Promoter regions of the human X-linked housekeeping genes PRPS1 and PRPS2 encoding phosphoribosylpyrophosphate synthetase subunit I and II isoforms.* **Ishizuka, T., et al., et al.** DOI: 10.1016/0167-4781(92)90521-z, 1992, Biochimica et Biophysica Acta, Vol. 1130, pp. 139 - 148.

69. *Prps1l1, a testis-specific gene, is dispensable for mouse spermatogenesis.* **Wang, Z., et al., et al.** DOI: 10.1002/mrd.23053, 2018, Mol Reprod Dev., Vol. 85, pp. 802 - 804.

70. *Protein and nucleotide biosynthesis are coupled through a single rate limiting enzyme, PRPS2, to drive cancer.* **Cunningham, J. T., et al., et al.** doi:10.1016/j.cell.2014.03.052., 2014, Cell, Vol. 157, pp. 1088 - 1103.

71. *Association of PRPS1 Mutations with Disease Phenotypes.* **Mittal, R., et al., et al.** http://dx.doi.org/10.1155/2015/127013, 2015, Disease Markers.

72. *Inhibition of Human 5-Phosphoribosyl-1-pyrophosphate Synthetase by 4-Amino-8-(beta-D-ribofuranosylamino)-pyrimido[5,4-d]pyrimidine-5 '-monophosphate: Evidence for Interaction at the ADP Allosteric Site.* **Fry, D. W., Becker, M. A. and Switzer, R. L.** 1995, Molecular Pharmacology, Vol. 47, pp. 810 - 815.

73. *Targeted Quantitative Kinome Analysis Identifies PRPS2 as a Promoter for Colorectal Cancer Metastasis.* **Miao, W. and Wang, Y.** DOI: 10.1021/acs.jproteome.9b00119, 2019, J. Proteome Res., Vol. 18, pp. 2279 - 2286.

74. *Direct role of nucleotide metabolism in C-MYCdependent proliferation of melanoma cells.* **Mannava, S., et al., et al.** https://doi.org/10.4161/cc.6390, 2008, Cell Cycle, Vol. 7, pp. 2392-2400.

75. *Phosphoribosyl pyrophosphate synthetases 2 knockdown inhibits prostate cancer progression by suppressing cell cycle and inducing cell apoptosis.* **Qiao, H., et al., et al.** doi: 10.7150/jca.37401, 2020, Journal of Cancer, Vol. 11, pp. 1027 - 1037.

76. *Down-Regulation of Phosphoribosyl Pyrophosphate Synthetase 1 Inhibits Neuroblastoma Cell Proliferation.* **Li, J., et al., et al.** doi:10.3390/cells8090955, 2019, Cells, Vol. 8.

77. *Phosphoribosylpyrophosphate Synthetase 1 Knockdown Suppresses Tumor Formation of Glioma CD133+ Cells Through Upregulating Cell Apoptosis.* **Li, C., et al., et al.** DOI 10.1007/s12031-016-0783-y, 2016, J Mol Neurosci, Vol. 60 , pp. 145–156.

78. *Mutant PRPS1: a new therapeutic target in relapsed acute lymphoblastic leukemia .* **Mullighan, C. G.** DOI: 10.1038/nm.3876, 2015, Nature Medicine, Vol. 21, pp. 553 - 554.

79. *Phosphoribosyl-pyrophosphate synthetase 2 (PRPS2) depletion regulates spermatogenic cell apoptosis and is correlated with hypospermatogenesis.* **Lei, B., et al., et al.** DOI: 10.4103/aja.aja_122_19, 2019, Asian Journal of Andrology, Vol. 21, pp. 1 - 7.

80. *Methylation-controlled J-protein MCJ acts in the import of proteins into human mitochondria.* **Schusdziarra, C., et al., et al.** doi:10.1093/hmg/dds541, 2013, Human Molecular Genetics, Vol. 22, pp. 1348–1357.

81. *Multi-faceted role of HSP40 in cancer.* **Mitra, A., Shevde, L. A. and Samant, R. S.** DOI 10.1007/s10585-009-9255-x, 2009, Clin Exp Metastasis, Vol. 26, pp. 559 - 567.

82. *J protein mutations and resulting proteostasis collapse.* **Koutras, C. and Braun, J. E. A.** doi: 10.3389/fncel.2014.00191, 2014, Frontiers in Cellular Neuroscience, Vol. 8.

83. *DNAJ Proteins and Protein Aggregation Diseases.* **Kakkar, V., Prins, L. C. B. and Kampinga, H. H.** DOI: 10.2174/1568026611212220004, 2012, Current Topics in Medicinal Chemistry, Vol. 12, pp. 2479-2490.

84. *Heat Shock Proteins: Agents of Cancer Development and Therapeutic Targets in Anti-Cancer Therapy.* **Yun, C. W., et al., et al.** doi:10.3390/cells9010060, 2020, Cells, Vol. 9.

85. *Heat Shock Proteins and Cancer.* **Wu, J., et al., et al.** http://dx.doi.org/10.1016/j.tips.2016.11.009, 2017, Trends in Pharmacological Sciences, Vols. 226 - 256, p. 38.

86. *An update on the discovery and development of selective heat shock protein inhibitors as anti-cancer therapy.* **Olotua, F., et al., et al.** https://doi.org/10.1080/17460441.2018.1516035, 2018, Expert Opinion on Drug Discovery, Vol. 13, pp. 903–918.

87. *Biological Roles of Neural J Proteins.* **Zhao, X., Braun, A. P. and Braun, J. E. A.** DOI 10.1007/s00018-008-8089-z, 2008, Cell. Mol. Life Sci., Vol. 65, pp. 2385 – 2396.

88. *The diversity of the DnaJ/Hsp40 family, the crucial partners for Hsp70 chaperones.* **Qiu, X.-B., et al., et al.** DOI 10.1007/s00018-006-6192-6, 2006, Cell. Mol. Life Sci., Vol. 63, pp. 2560–2570.

89. *Methylation-Controlled J Protein Promotes c-Jun Degradation To Prevent ABCB1 Transporter Expression.* **Hatle, K. M., et al., et al.** doi:10.1128/MCB.01804-06, 2007, Molecular and Cellular Biology, Vol. 27, pp. 2952–2966.

90. *Loss of Expression of a New Member of the DNAJ Protein Family Confers Resistance to Chemotherapeutic Agents Used in the Treatment of Ovarian Cancer.* **Shridhar, V., et al., et al.** 2001, Cancer Research, Vol. 61, pp. 4258 - 4265.

91. *Epigenetic inactivation of MCJ (DNAJD1) in malignant paediatric brain tumours.* **Lindsey, J. C., et al., et al.** DOI 10.1002/ijc.21353, 2006, Int. J. Cancer, Vol. 118, pp. 346 - 352.

92. *Cell type-specific methylation of an intronic CpG island controls expression of the MCJ gene.* **Strathdee, G., et al., et al.** DOI: 10.1093/carcin/bgh066, 2004, Carcinogenesis, Vol. 25, pp. 693 - 701.

93. *ETV7-Mediated DNAJC15 Repression Leads to Doxorubicin Resistance in Breast Cancer Cells.* **Alessandrini, F., et al., et al.** https://doi.org/10.1016/j.neo.2018.06.008, 2018, Neoplasia, Vol. 20, pp. 857 - 870.

94. *Hypomethylation and hypermethylation of DNA in Wilms tumors.* **Ehrlich, M., et al., et al.** DOI: 10.1038/sj.onc.1205890, 2002, Oncogene, Vol. 21, pp. 6697 - 6702.

95. *Prognostic Significance of Promoter DNA Methylation in Patients with Childhood Neuroblastoma.* **Lau, D. T., et al., et al.** doi:10.1158/1078-0432.CCR-12-0294, 2012, Clin Cancer Res, Vol. 18, pp. 5690-5700.

96. *Deficiency of mitochondrial modulator MCJ promotes chemoresistance in breast cancer.* **Fernández-Cabezudo, M. J.** doi:10.1172/jci.insight.86873, 2016, JCI Insight.

97. *Role of Magmas in protein transport and human mitochondria biogenesis.* **Sinha, D., et al., et al.** doi:10.1093/hmg/ddq002, 2010, Human Molecular Genetics, Vol. 19, pp. 1248 – 1262.

98. *Mutation of DNAJC19, a human homologue of yeast inner mitochondrial membrane co-chaperones, causes DCMA syndrome, a novel autosomal recessive Barth syndrome-like condition.* **Davey, K. M., et al., et al.** doi: 10.1136/jmg.2005.036657, 2006, J Med Genet , Vol. 43, pp. 385–393.

99. *Novel inhibitors of heat shock protein Hsp70-mediated luciferase refolding that bind to DnaJ.* **Cassel, J. A., et al., et al.** http://dx.doi.org/10.1016/j.bmc.2012.03.067, 2012, Bioorganic & Medicinal Chemistry, Vol. 20, pp. 3609–3614.

100. *Intracellular iron and heme trafficking and metabolism in developing erythroblasts .* **Kafina, M. D. and Paw, B. H.** doi:10.1039/c7mt00103g., 2017, Metallomics, Vol. 9, pp. 1193-1203.

101. *Regulation of Mitochondrial Iron Import through Differential Turnover of Mitoferrin 1 and Mitoferrin 2.* **Paradkar, P. N., et al., et al.** doi:10.1128/MCB.01685-08, 2009, Moluecular and Cellular Biology, Vol. 29, pp. 1007 - 1016.

102. *Abcb10 physically interacts with mitoferrin-1 (Slc25a37) to enhance its stability and function in the erythroid mitochondria.* **Chen, W., et al., et al.** https://doi.org/10.1073/pnas.0904519106, 2009, PNAS, Vol. 106, pp. 16263-16268.

103. *Identification of Distal cis-Regulatory Elements at Mouse Mitoferrin Loci Using Zebrafish Transgenesis.* **Amigo, J. D., et al., et al.** doi:10.1128/MCB.01010-10, 2011, Molecular and Cellular Biology, Vol. 31, pp. 1344-1356.

104. *A dynamic intron retention program enriched in RNA processing genes regulates gene expression during terminal erythropoiesis.* **Pimentel, H., et al., et al.** doi: 10.1093/nar/gkv1168, 2016, Nucleic Acids Research, Vol. 44.

105. *Mitoferrin is essential for erythroid iron assimilation.* **Shaw, G.C., et al., et al.** doi:10.1038/nature04512, 2006, Nature, Vol. 440, pp. 96-100.

106. *Identification of SLC25A37 as a major depressive disorder risk gene.* **Huo, Y.-X., et al., et al.** http://dx.doi.org/10.1016/j.jpsychires.2016.09.011, 2016, Journal of Psychiatric Research, Vol. 83, pp. 168-175.

107. *Mitoferrin-1 is required for brain energy metabolism and hippocampus-dependent memory.* **Baldauf, L., et al., et al.** https://doi.org/10.1016/j.neulet.2019.134521, 2019, Neuroscience Letters, Vol. 713.

108. *Drosophila mitoferrin is essential for male fertility: evidence for a role of mitochondrial iron metabolism during spermatogenesis.* **Metzendorf, C. and Lind, M. I.** doi: 10.1186/1471-213X-10-68, 2010, BMC Developmental Biology, Vol. 10.

109. *Reduction of Mitoferrin Results in Abnormal Development and Extended Lifespan in Caenorhabditis elegans.* **Ren, Y., et al., et al.** https://doi.org/10.1371/journal.pone.0029666, 2012, PLoS, Vol. 7.

110. *PINK1 and PARK2 Suppress Pancreatic Tumorigenesis through Control of Mitochondrial Iron-Mediated Immunometabolism.* **Li, C., et al., et al.** https://doi.org/10.1016/j.devcel.2018.07.012, 2018, Developmental Cell, Vol. 46, pp. 441-455.

111. *Molecular Cloning, Expression, and Characterization of Human Bifunctional 3'-Phosphoadenosine 5'-Phosphosulfate Synthase and Its Functional Domains.* **Venkatachalam, K. V., Akita, H. and Strott, C. A.** DOI: 10.1074/jbc.273.30.19311, 1998, Journal of Biological Chemistry, Vol. 273, pp. 19311-19320.

112. *Human DHEA sulfation requires direct interaction between PAPS synthase 2 and DHEA sulfotransferase SULT2A1.* **Mueller, J. W., et al., et al.** DOI 10.1074/jbc.RA118.002248, 2018, J. Biol. Chem, Vol. 293, pp. 9724–9735.

113. *Human PAPS Synthase Isoforms Are Dynamically Regulated Enzymes with Access to Nucleus and Cytoplasm.* **Schröder, E., et al., et al.** doi:10.1371/journal.pone.0029559, 2012, PLoS ONE, Vol. 7.

114. *Structural Mechanism for Substrate Inhibition of the Adenosine 5'-Phosphosulfate Kinase Domain of Human 3'-Phosphoadenosine 5'-Phosphosulfate Synthetase 1 and Its Ramifications for Enzyme Regulation.* **Sekulic, N., Konrad, M. and Lavie, A.** DOI 10.1074/jbc.M701713200, 2007, Journal of Biological Chemistry, Vol. 282, pp. 22112 - 22121.

115. *The Crystal Structure of Human PAPS Synthetase 1 Reveals Asymmetry in Substrate Binding.* **Harjes, S., Bayer, P. and Scheidig, A. J.** doi:10.1016/j.jmb.2005.01.005, 2005, J. Mol. Biol., Vol. 347, pp. 623-635.

116. *Human 3'-phosphoadenosine 5'-phosphosulfate (PAPS) Synthase: Biochemistry, Molecular Biology and Genetic Deficiency.* **Venkatachalam, K. V.** DOI: 10.1080/1521654031000072148, 2003, Life, Vol. 55, pp. 1 - 11.

117. *Characterization and expression of human bifunctional 3'-phosphoadenosine 5'-phosphosulphate synthase isoforms.* **Fuda, H., et al., et al.** doi: 10.1042/BJ20020044, 2002, Biochem. J., Vol. 365, pp. 497 - 504.

118. *PTEN: multiple functions in human malignant tumors.* **Milella, M., et al., et al.** doi: 10.3389/fonc.2015.00024, 2015, Frontiers in Oncology, Vol. 5.

119. *Enhanced PAPSS2/VCAN sulfation axis is essential for Snail-mediated breast cancer cell migration and metastasis.* **Zhang, Y., et al., et al.** https://doi.org/10.1038/s41418-018-0147-y, 2019, Cell Death & Differentiation, Vol. 26, pp. 565 - 579.

120. *Altered responsiveness to TGF-b results in reduced Papss2 expression and alterations in the biomechanical properties of mouse articular cartilage.* **Ramaswamy, G., et al., et al.** doi:10.1186/ar3762, 2012, Arthritis Research & Therapy, Vol. 14.

121. *TGFβ in Cancer.* **Massagué, J.** DOI 10.1016/j.cell.2008.07.001, 2008, Cell, Vol. 134, pp. 215 - 230.

122. *Clinical and Radiographic Features of the Autosomal Recessive form of Brachyolmia Caused by PAPSS2 Mutations.* **Iida, A., et al., et al.** DOI: 10.1002/humu.22377, 2013, Human Mutation, Vol. 34, pp. 1381 - 1386.

123. *Inactivating PAPSS2 Mutations in a Patient with Premature Pubarche.* **Noordam, C., et al., et al.** DOI: 10.1056/NEJMoa0810489, 2009, N. Engl. J. Med., Vol. 360, pp. 2310 - 2318.

124. *PAPSS2 Deficiency Causes Androgen Excess via Impaired DHEA Sulfation—In Vitro and in Vivo Studies in a Family Harboring Two Novel PAPSS2 Mutations.* **Oostdijk, W., et al., et al.** doi: 10.1210/jc.2014-3556, 2015, J Clin Endocrinol Metab, Vol. 100, pp. E672 - E680.

125. *A heterodimer of human 30-phospho-adenosine-50-phosphosulphate (PAPS) synthases is a new sulphate activating complex.* **Grum, Daniel, et al., et al.** doi:10.1016/j.bbrc.2010.04.039, 2010, Biochemical and Biophysical Research Communications, Vol. 395, pp. 420-425.

126. *3'-Phosphoadenosine 5'-phosphosulfate synthase 1 (PAPSS1) knockdown sensitizes non-small cell lung cancer cells to DNA damaging agents.* **Leung, A. W. Y., et al., et al.** DOI: 10.18632/oncotarget.3635, 2015, Oncotarget, Vol. 6, pp. 17161 - 17177.

127. *In Vivo Validation of PAPSS1(3'-phosphoadenosine 5'-phosphosulfate synthase 1) as a Cisplatin-sensitizing Therapeutic Target.* **Leung, A. W.Y., et al., et al.** DOI: 10.1158/1078-0432.CCR-17-0700, 2017, Clin Cancer Res.

128. *Comparative sequence and expression analyses of four mammalian VPS4 genes.* **Beyer, A., et al., et al.** doi:10.1016/S0378-1119(02)01205-2, 2003, Gene, Vol. 303, pp. 47-59.

129. *Mammalian Cells Express Two VPS4 Proteins Both of Which are Involved in Intracellular Protein Trafficking .* **Scheuring, S., et al., et al.** doi:10.1006/jmbi.2001.4917, 2001, J. Mol. Biol., Vol. 321, pp. 469-480.

130. *The Expression Changes of Vacuolar Protein Sorting 4B (VPS4B) Following Middle Cerebral Artery Occlusion (MCAO) in Adult Rats Brain Hippocampus.* **Cui, G., et al., et al.** DOI 10.1007/s10571-013-9989-5, 2014, Cell Mol Neurobiol, pp. 83-94.

131. *Dynamic subunit turnover in ESCRT-III assemblies is regulated by Vps4 to mediate membrane remodelling during cytokinesis.* **Mierzwa, B. E., et al., et al.** doi:10.1038/ncb3559., 2017, Nat Cell Biol, Vol. 19, pp. 787-798.

132. *Vacuolar Protein Sorting 4B (VPS4B) Regulates Apoptosis of Chondrocytes via p38 Mitogen-Activated Protein Kinases (MAPK) in Osteoarthritis.* **Xu, L., et al., et al.** DOI: 10.1007/s10753-017-0633-2, 2017, Inflammation, Vol. 40, pp. 1924-1932.

133. *Vacuolar protein sorting 4B regulates apoptosis of intestinal epithelial cells via p38 MAPK in Crohn's disease.* **Zhang, D., et al., et al.** http://dx.doi.org/10.1016/j.yexmp.2014.12.007, 2015, Experimental and Molecular Pathology, Vol. 98, pp. 55-64.

134. *Identification of an AAA ATPase VPS4B-Dependent Pathway That Modulates Epidermal Growth Factor Receptor Abundance and Signaling during Hypoxia.* **Lin, H. H., et al., et al.** DOI: 10.1128/MCB.06053-11, 2012, Molecular and Cellular Biology, Vol. 32, pp. 1124–1138.

135. *An Internal Standard-Assisted Synthesis and Degradation Proteomic Approach Reveals the Potential Linkage between VPS4B Depletion and Activation of Fatty Acid β-Oxidation in Breast Cancer Cells.* **Liao, Z., et al., et al.** http://dx.doi.org/10.1155/2013/291415, 2013, International Journal of Proteomics, Vol. 2013.

136. *Vacuolar protein sorting 4B, an ATPase protein positively regulates the progression of NSCLC via promoting cell division.* **Liu, Y., et al., et al.** DOI 10.1007/s11010-013-1699-2, 2013, Mol Cell Biochem, Vol. 381, pp. 163-171.

137. *Synthetic lethality between VPS4A and VPS4B triggers an inflammatory response in colorectal cancer.* **Szymanska, E., et al., et al.** doi: 10.15252/emmm.201910812, 2020, EMBO Molecular Medicine, Vol. 12.

138. *Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells.* **Meyers, R. M. and al., et.** doi:10.1038/ng.3984, 2017, Nature Genetics, Vol. 49, pp. 1779-1784.

139. *An Optimized microRNA Backbone for Effective Single-Copy RNAi.* **Fellmann, C., et al., et al.** http://dx.doi.org/10.1016/j.celrep.2013.11.020, 2013, Cell Reports, Vol. 5, pp. 1704–1713.

140. *Plat-E: an efficient and stable system for transient packaging of retroviruses.* **Morita, S., .Kojima, T and Kitamura, T.** DOI: 10.1038/sj.gt.3301206, 2000, Gene Therapy, Vol. 7, pp. 1063–1066.

141. *Evaluation of an Automated Instrument for Viability and Concentration Measurements of Cryopreserved Hematopoietic Cells.* **Szabo, S. E., et al., et al.** doi: 10.1532/LH96.04020, 2004, Laboratory Hematology, Vol. 10, pp. 109-111.

142. *Trypan Blue Exclusion Test of Cell Viability.* **Strober, W.** doi: 10.1002/0471142735.ima03bs111, 2015, Current Protocols in Immunology, pp. A.3.B.1-A.3.B.3.

74

143. *UniProt: a worldwide hub of protein knowledge.* **The UniProt Consortium.** doi: 10.1093/nar/gky1049, 2019, Nucleic Acids Research, Vol. 47, pp. D505-D151.

144. *NCBI BLAST: a better web interface.* **Johnson, M., et al., et al.** doi:10.1093/nar/gkn201, 2008, Nucleic Acids Research, Vol. 36.

145. *Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples.* **Wagner, G. P., Kin, K. and Lynch, V. J.** DOI 10.1007/s12064-012-0162-3, 2012, Theory Biosci., Vol. 131, pp. 281-285.

146. *The Genotype-Tissue Expression (GTEx) project.* **The GTEx Consortium.** https://doi.org/10.1038/ng.2653, 2013, Nature Genetics, Vol. 45, pp. 500-585.

147. *The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data.* **Cerami, E., et al., et al.** doi:10.1158/2159-8290.CD-12-0095, 2012, Cancer Discov., Vol. 2, pp. 401-404.

148. *Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal.* **Gao, J., et al., et al.** doi:10.1126/scisignal.2004088, 2013, Sci Signal, Vol. 6.

149. *Assessing the significance of chromosomal aberrations in cancer: Methodology and application to glioma.* **Beroukhim, R., Getz, G. and al, et.** https://doi.org/10.1073/pnas.0710052104, 2007, PNAS, Vol. 104, pp. 20007–20012.

150. *Identification and characterization of essential genes in the human genome.* **Wang, T., et al., et al.** doi: 10.1126/science.aac7041, 2015, Science, Vol. 350, pp. 1069 - 1101.

151. *Gene essentiality and synthetic lethality in haploid human cells.* **Blomen, V. A., et al., et al.** DOI: 10.1126/science.aac7557, 2015, Science , Vol. 350, pp. 1092-1096.

152. *High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities.* **Hart, T., et al., et al.** https://doi.org/10.1016/j.cell.2015.11.015, 2015, Cell, Vol. 163, pp. 1515-1526.

153. *TSGene: a web resource for tumor suppressor genes.* **Zhao, M., Sun, J. and Z.Zhao.** doi: 10.1093/nar/gks937, 2013, Nucleic Acids Res, Vol. 41, pp. D970-6.

154. *The SMART App: an interactive web application for comprehensive DNA methylation analysis and visualization.* **Li, Y., Ge, D. and Lu, C.** https://doi.org/10.1186/s13072-019-0316-3, 2019, Epigenetics & Chromatin, Vol. 12.

155. *PTEN, NHERF1 and PHLPP form a tumor suppressor network that is disabled in glioblastoma.* **Molina, J. R., et al., et al.** doi: 10.1038/onc.2011.324, 2012, Oncogene, Vol. 10, pp. 1264-1274.

156. *Serpin peptidase inhibitor (SERPINB5) haplotypes are associated with susceptibility to hepatocellular carcinoma.* **Yang, S., et al., et al.** DOI: 10.1038/srep26605, 2016, Sci Rep., Vol. 6.

157. *The Role of DNA Methylation in Cancer.* **Lakshminarasimhan, R. and Liang, G.** doi: 10.1007/978-3-319-43624-1_7, 2016, Adv Exp Med Biol., Vol. 945, pp. 151 - 172.

158. *Tyrosine Sulfation as a Protein Post-Translational Modification.* **Yang, Y., et al., et al.** doi:10.3390/molecules20022138, 2015, Molecules, Vol. 20, pp. 2138 - 2164.

159. *Genetic interaction mapping and exon-resolution functional genomics with a hybrid Cas9–Cas12a platform.* **Gonatopoulos-Pournatzis, T., et al., et al.**

160. *Targeting Strategies for Tissue-Specific Drug Delivery .* **Zhao, Z., et al., et al.** https://doi.org/10.1016/j.cell.2020.02.001, 2020, Cell, Vol. 181, pp. 151-167.

161. *Cohesin in cancer: chromosome segregation and beyond.* **Losada, A.** doi:10.1038/nrc3743, 2014, Nature Reviews Cancer, Vol. 14, pp. 389 - 393.

162. *Cohesin gene mutations in tumorigenesis: from discovery to clinical significance.* **Solomon, D. A., Kim, J.-S. and Waldman, T.** http://dx.doi.org/10.5483/BMBRep.2014.47.6.092, 2014, BMB Rep., Vol. 47, pp. 299 - 310.

163. *Synthetic lethal interaction between the tumour suppressor STAG2 and its paralog STAG1.* **Benedetti, L., et al., et al.** doi: 10.18632/oncotarget.16838., 2017, Oncotarget, Vol. 8, pp. 37619-37632.

164. *Synthetic lethality between the cohesin subunits STAG1 and STAG2 in diverse cancer contexts.* **van der Lelij, P., et al., et al.** DOI: 10.7554/eLife.26980, 2017, eLife.

165. *Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres.* **Remeseiro, S., et al., et al.** DOI: 10.1038/emboj.2012.11, 2012, EMBO Journal, Vol. 31, pp. 2076 - 2089.

166. *A unique role of cohesin-SA1 in gene regulation and development.* **Remeseiro, S., et al., et al.** doi:10.1038/emboj.2012.60, 2012, The EMBO Journal, Vol. 31, pp. 2090–2102.

167. *Differential regulation of telomere and centromere cohesion by the Scc3 homologues SA1 and SA2, respectively, in human cells.* **Canudas, S. and Smith, S.** www.jcb.org/cgi/doi/10.1083/jcb.200903096, 2009, J. Cell Biol., Vol. 187, pp. 165–173.

168. *Distinct Functions of Human Cohesin-SA1 and Cohesin-SA2 in Double-Strand Break Repair.* **Kong, X., et al., et al.** http://dx.doi.org/10.1128/MCB.01503-13., 2014, Molecular and Cellular Biology, Vol. 34, pp. 685–698.

169. *Gene regulation and chromatin organization: relevance of cohesin mutations to human disease.* **Watrin, E., Kaiser, F. J. and Wendt, K. S.** http://dx.doi.org/10.1016/j.gde.2015.12.004, 2016, Current Opinion in Genetics & Development, Vol. 37, pp. 59–66.

12 List of Figures
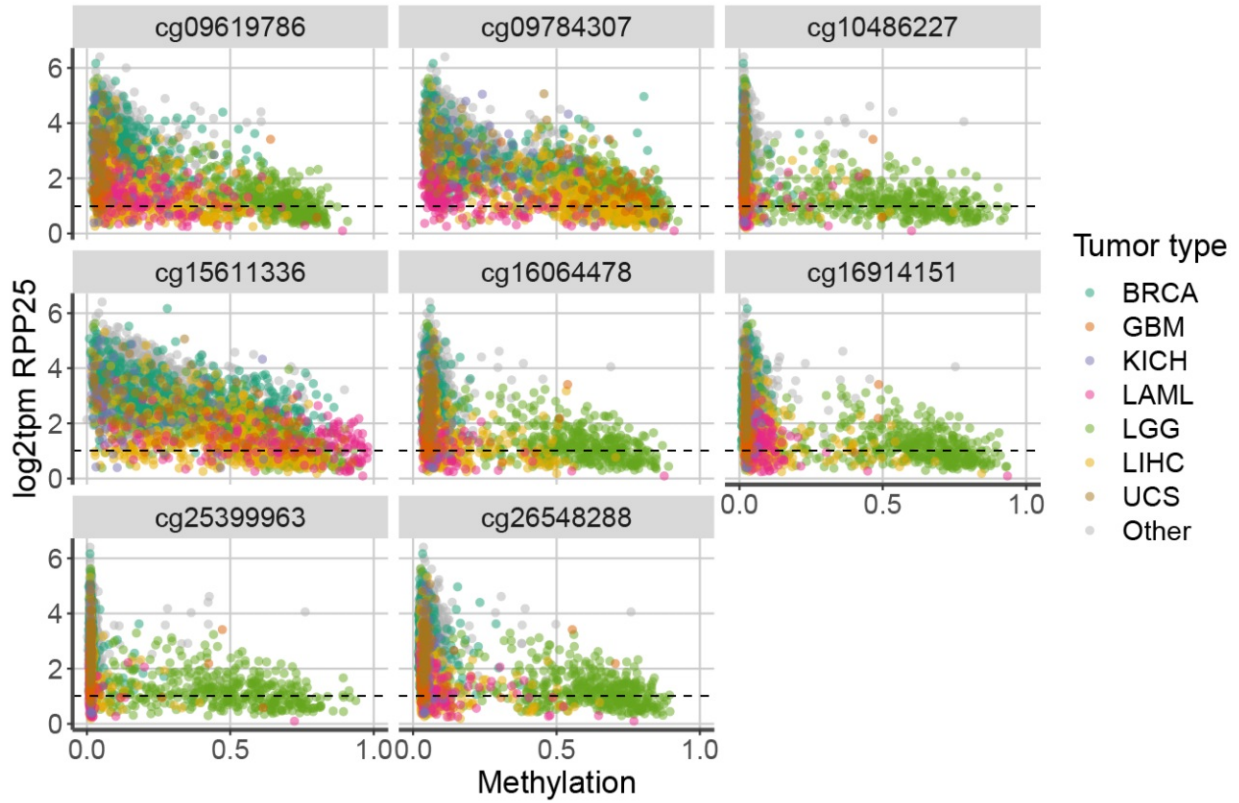
13 List of Tables

14 Appendix

Figure 48: Correlation between methylation of RPP25 CpG islands and gene expression in TCGA samples