Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

Sebastian Weingart, M.Sc.

Born in Henstedt-Ulzburg, Germany

Oral examination: 20.06.2022

A pooled CRISPR/Cas9 screen to identify genes synthetically lethal with epigenetic drugs in non-small cell lung cancer cells

Referees: Prof. Dr. Ursula Klingmüller

Prof. Dr. Holger Sültmann

In deepest love to
Agnes, Sofia, and Antonia

and

in deep gratitude
to my parents

Abstract

Lung cancer is the most common cause of cancer-related death worldwide, resulting in 1.3 million deaths per year. Most lung tumors are non-small cell lung cancers (NSCLC), and approximately 11 % of NSCLC harbor a mutation in the oncogenic *KRAS* gene. The clinical success of epigenetic drugs has been minimal in NSCLC, with no exception for the *KRAS* mutant NSCLC subtype. This project aimed to identify genes that act synthetically lethal with epigenetic drugs to better understand the resistance mechanisms to these compounds in NSCLC cells, focusing on *KRAS* mutant cells.

Target genes were identified using a pooled CRISPR/Cas9 screening approach, which was established in this project. Four genes synthetically lethal with the histone methyltransferase inhibitor DZNep were identified in *KRAS* mutant H2030 cells, and two genes were identified as synthetically lethal with the histone deacetylase inhibitor Entinostat in the same cell line. The epigenetic regulators *CHD8* and *EP300* were selected among the target genes for further characterization.

In silico analysis of *CHD8* and *EP300* in the TCGA NSCLC dataset revealed a strong positive correlation of the expression of both genes and indicated functional overlaps. The synthetically lethal effect of *CHD8* or *EP300* depletion with DZNep treatment was validated in the *KRAS*-mutant cell line H1944 but lacking in *EGFR*-mutant H1975 cells. Flow cytometry analysis confirmed elevated apoptosis in H2030, but not H1975 cells, after DZNep treatment in conjunction with *CHD8* or *EP300* knockdown. Gene expression profiling and functional annotation of deregulated genes after DZNep treatment in combination with *CHD8* or *EP300* knockdown verified a high number of affected cell-cycle and cell viability-related genes. The cytokine *CCL20* was identified as upregulated by DZNep but downregulated upon knockdown of *CHD8* or *EP300*, suggesting a role in conferring *CHD8* or *EP300*-mediated resistance to DZNep treatment in H2030 cells. Proof of concept experiments showed that *CHD8* is critical for the upregulation of *CCL20* upon DZNep treatment in mutant H2030 cells and that a knockdown of *CCL20* mimics the synthetically lethal effect of *CHD8* depletion with DZNep treatment.

Zusammenfassung

Mit 1,3 Millionen Todesfällen pro Jahr ist Lungenkrebs weltweit die häufigste krebsbedingte Todesursache. Die Mehrzahl der Lungentumoren sind nicht-kleinzellige Bronchialkarzinome (non-small cell lung cancer, NSCLC). Etwa 11 % aller NSCLC weisen eine Mutation im onkogenen *KRAS*-Gen auf. Der klinische Erfolg von epigenetischen Inhibitoren war in NSCLC bisher minimal, ohne Ausnahme für den *KRAS*-mutierten NSCLC-Subtyp. Dieses Projekt hatte zum Ziel, in *KRAS*-mutierten NSCLC Zellen solche Gene zu identifizieren, deren Entfernung aus der Zelle zusammen mit der Gabe von epigenetischen Inhibitoren synthetisch letal wirkt. Die Erkenntnisse dieser Studie sollen helfen, die Resistenzmechanismen gegen epigenetische Inhibitoren besser zu verstehen. Die Zielgene wurden mit Hilfe eines CRISPR/Cas9-Screens identifiziert, der im Rahmen dieses Projekts etabliert wurde. Vier Gene, die mit dem Histonmethyltransferase-Inhibitor DZNep synthetisch letal sind, wurden in *KRAS*-mutierten H2030-Zellen identifiziert. Zwei weitere Gene wurden mit dem Histondeacetylase-Inhibitor Entinostat in derselben Zelllinie als synthetisch letal identifiziert. Die epigenetischen Regulatoren *CHD8* und *EP300* wurden aus den Zielgenen für die weitere Charakterisierung ausgewählt. Die Analyse von *CHD8* und *EP300* im TCGA-NSCLC-Datensatz zeigte eine positive Korrelation der Expression beider Gene und deutete auf funktionelle Überschneidungen hin. Ein *knockdown* von *CHD8* oder *EP300* unter DZNep-Behandlung in der *KRAS*-mutanten Zelllinie H1944 zeigte -identisch zu H2030 Zellen- eine synthetische Letalität und erhöhte Apoptose, in *EGFR*-mutierten H1975-Zellen zeigte sich diese jedoch nicht. Microarray-Genexpressionsanalysen verifizierten eine hohe Anzahl deregulierter Zellzyklus- und Zellviabilitätsgene. Das Zytokin *CCL20* wurde als möglicher Mediator einer *CHD8*- oder *EP300*-vermittelten Resistenz gegen DZNep in H2030-Zellen identifiziert. Initiale Experimente zeigten, dass *CHD8* eine zentrale Rolle für die erhöhte Expression von *CCL20* nach einer DZNep-Behandlung in *KRAS*-mutierten H2030-Zellen einnimmt und dass ein *knockdown* von *CCL20* ebenfalls eine synthetische Letalität mit einer DZNep Behandlung der Zellen aufweist.

Table of contents

List of figures

List of tables

Abbreviations

°C	Degree Celsius
µL	Microliter
µM	Micromolar
ATCC	American type culture collection
ATP	Adenosine triphosphate
Bp	Base pair
BRAF	Murine sarcoma viral oncogene homolog B1
Cas	CRISPR-associated Protein
CCL20	C-C Motif Chemokine Ligand 20
cDNA	Complementary DNA
CHD8	Chromodomain-helicase-DNA-binding protein 8
CO_2	Carbon dioxide
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DAVID	Database for Annotation, Visualization and Integrated Discovery
DKFZ	German Cancer Research Center
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethylsulfoxid
DNA	desoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
DPBS	Dulbecco's phosphate buffered saline
DZNep or DZN	3-Deazaneplanocin A
E. coli	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EP300	and Histone acetyltransferase p300
ERK	Extracellular signal-regulated kinase
EZH2	enhancer of zeste homologue 2
FACS	Fluorescence-activated Cell Sorting
FBS	fetal bovine serum
FC	Fold change
FDA	Food and Drug Administration
Fwd	Forward
GAP	GTPase-activating protein
GDP	guanosine diphosphate
GEF	guanine nucleotide exchange factor
GFP	green fluorescent protein
GO	Gene Ontology
GPX4	Glutathione peroxidase 4
GTP	guanosine triphosphate
h	hours
HDACi	Histone deacetylase inhibitor
HVR	hypervariable region

IC	Inhibitory concentration
IPA	Ingenuity Pathway Analysis
KDM5C	Lysine-specific demethylase 5C
KEGG	Kyoto Encyclopedia of Genes and Genomes
KRAS	V-KI-RAS Kirsten rat sarcoma viral oncogene homolog
LC	Lung cancer
MAP	mitogen-activated protein
MAPK	mitogen-activated protein kinase
MEK	Mitogen-activated kinase kinase
mL	Milliliter
mM	Millimolar
mRNA	Messenger RNA
MTOR	mechanistic target of rapamycin
mTOR	Mammalian target of rampamycin
n.s.	Not significant
N/A	Not available
NCT	National Center of Tumor Diseases
NF-κB	polypeptide gene enhancer in B-cells
NSCLC	Non-small-cell lung carcinoma
NT	Non-targeting
PAGE	polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	polymerase chain reaction
PE	Phycoerythrin
PRC1	Polycomb repressive complex member 1
PRC2	Polycomb repressive complex member 2
PRMT1	Protein arginine N-methyltransferase 1
RAS	Rat sarcoma protein
Rev	Reverse
RFP	red fluorescent protein
RNA	ribonucleotide acid
RNAi	RNA interference
RT	room temperature
RTK	receptor tyrosine kinase
RT-qPCR	Reverse transcriptase quantitative polymerase chain reaction
SCLC	small-cell lung cancer
sec	second(s)
sgRNA	single-guide RNA
siRNA	small interfering RNA
TCGA	The Cancer Genome Atlas
TKI	Tyrosine-kinase inhibitor
TP53	Tumor protein p53
UPL	Universal Probe Library
UV	ultraviolet
WHO	World Health Organization
WT	Wild-type

1 Introduction

1.1 Lung cancer

Lung cancer is the most frequently diagnosed cancer worldwide, accounting for 2.1 million newly diagnosed cases in 2018. With 1.8 million deaths per year, lung cancer accounts for most cancer-related deaths[1]. Notably, lung cancer causes as many deaths per year as the next four most common cancers combined, i.e., breast, colon, prostate, and pancreatic cancers[2]. It is the 2nd most common cancer in both genders, behind prostate cancer in men and breast cancer in women[3]. In 2017, lung cancer mortality (14.6 deaths per 100,000 cases) exceeded breast cancer mortality (14.0 deaths per 100,000 cases) in European females. Lung cancer typically affects elderly patients. In 2015, 86 % of newly diagnosed cases occurred in patients 60 years of age or older. In contrast, children and young adults are rarely affected[4]. Lung cancer symptoms typically present only in advanced stages of the disease and may include a long-standing cough, loss of appetite, breathlessness, fatigue, and pain while breathing or coughing[5].

The most common cause of lung cancer is tobacco smoking, accounting for approximately 80 % of cases in women and 90 % in men[6]. The likelihood of developing lung cancer is enhanced 23 times in men who smoke compared to non-smokers and 13 times in women who smoke[7]. The risk of developing lung cancer is increased by 20 - 30 % for non-smokers if they are exposed to passive smoking. The variation of lung cancer incidences around the world may be explained by geographical differences in tobacco use[8]. Due to smoking habits, the frequency of lung cancer is higher for men in developed countries compared to less-developed ones. Tobacco control policies led to decreasing rates of lung cancer in men in developed countries. However, presumably linked to the increasing prevalence of tobacco consumption of women, the incidence of lung cancer in women is rising worldwide[8]. Due to exposure to a complex mix of carcinogens in tobacco smoke, lung cancer is one of the most heavily mutated and genomically altered cancers[9]. Other than active or passive smoking, further risk factors for developing lung cancer include exposure to air pollution, silica dust, and asbestos.

More than 50 % of patients with lung cancer die within one year after diagnosis[10]. The overall 5-year survival rate for patients diagnosed with lung cancer is 18.6 %, compared

to 65 % in colorectal and 90 % in breast cancer[10]. Only 16 % of all lung cancer patients are diagnosed with early-stage cancer due to the late onset of symptoms[10]. Moreover, distant metastases are found in about 65 % of lung cancer cases at the time of diagnosis. For these, the 5-year survival rate further drops to 5 %[10]. In contrast, if diagnosed with early-stage lung cancer, the 5-year survival rate is 56 %[10]. The gap between early-stage vs. advanced-stage lung cancer survival rates illustrates the impact of the tumor stage at the time of diagnosis on the patient's prognosis. After diagnosis, the therapeutic regimen depends on the stage and the subtype of the disease and may include resection, chemotherapy, or targeted therapy aiming at the inhibition of receptor tyrosine kinases. Especially for early-stage tumors, surgical resection remains the most successful therapy for a cure if a tumor is entirely resectable[11].

1.1.1 Classification of lung cancer

Lung cancers are historically classified into Non-small-cell lung carcinoma (NSCLC) and small-cell lung carcinoma (SCLC) according to their histological features[9]. Approximately 85 % of diagnosed lung cancer cases are characterized as NSCLC, and 15 % are found to be SCLC[2] (Figure 1).

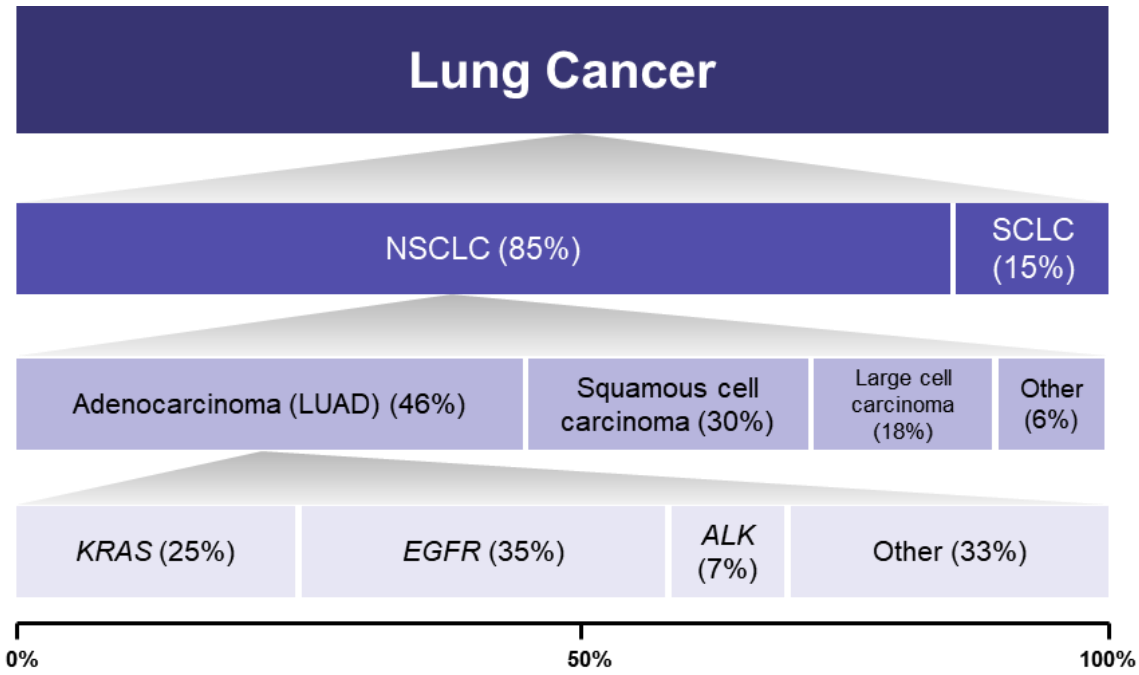


Figure 1 | Illustration of the Lung cancer landscape

Lung cancer is histologically divided into Non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC is further divided into three subclasses: Adenocarcinoma (LUAD), squamous cell carcinoma, and large cell carcinoma. LUAD can be subdivided based on molecular alterations, according to mutations in the *KRAS* or *EGFR* gene or translocations of the *EML4-ALK* gene.

Although NSCLC and SCLC are both primary tumors of the lung, share the same risk factors, and lead to comparable symptoms, their location of origin in the lung is different. NSCLC usually arises from the peripheral lung tissue, whereas SCLC mostly arises from primary and secondary bronchi. Histologically, SCLC is composed of small, round to spindled cells with dark nuclei, scant cytoplasm, and granular nuclear chromatin with indistinct nucleoli[12]. Almost 50 % of NSCLC cases are lung adenocarcinomas (LUAD), followed by squamous cell carcinomas (SCCs, ~30 %) and large cell carcinomas (LCC, ~ 20 %)[9]. Historically, LUAD was defined by the WHO as carcinoma with an acinar/tubular structure or mucin production, whereas SCC was defined according to keratinization or intercellular bridges[13]. In 2015, the WHO classified LUAD into five distinct histologic subtypes: lepidic, papillary, acinar, micropapillary, and solid[14]. Importantly, the outcomes of surgical resection of early-stage LUAD correlates with these subtypes, with micropapillary and solid subtypes being significantly worse[15]. Although smoking is the leading cause for all subtypes, LUAD is the most common form among never-smokers, defined by a threshold of not having smoked more than 100 cigarettes in their lifetimes[16].

A variety of genes have been identified as oncogenic drivers for LUAD, including *Kirsten Rat Sarcoma virus* (*KRAS*) and *Epidermal Growth Factor Receptor* (*EGFR*), who contribute to ~25 % (*KRAS*) and ~35 % (*EGFR*) of LUAD cases. Notably, mutations in the *KRAS* and *EGFR* genes are mutually exclusive in almost all cases upon the time of diagnosis[17]. In 2021, the first specific KRAS inhibitor, Sotorasib, was approved for *KRAS* G12C mutated NSCLC by the US Food and Drug Administration (FDA). In January 2022, approval was granted in the EU for the same indication. *EGFR* mutant LUAD can be treated with highly specific tyrosine kinase inhibitors (TKi), such as Gefitinib, Erlotinib, Afatinib, or Osimertinib. A minor share of LUAD cases (~7 %) harbors a fusion of the *anaplastic lymphoma kinase* (*ALK*) gene with the *echinoderm microtubule-associated protein-like 4* (*EML4*) gene. Patients with *EML4-ALK* fusions are eligible to be treated with the tyrosine kinase inhibitor Crizotinib, which provides clinical benefit to this patient group[9]. With Tepotinip, an inhibitor of the receptor tyrosine kinase MET, another evolving option for targeted therapy of LUAD has recently been given a “breakthrough therapy” status by the FDA. Approximately 1.4 % of newly diagnosed LUAD patients harbor skipping alterations in exon 14 of the *MET* proto-oncogene[18]. Taken together, LUAD represents the most prominent subtype among all lung cancers, and targeted therapies providing clinical benefits are available to

LUAD patients with alterations in *EGFR*, *EML4-ALK*, and *MET*, and recently also to a subset of *KRAS* mutant tumors.

1.1.2 *KRAS*-mutant lung adenocarcinoma

HRAS and *KRAS* genes were identified in 1964 from studies that investigated the cancer-causing Harvey sarcoma virus and the Kirsten sarcoma virus[19]. These viruses were initially discovered in rats, hence the name *RAS*, which is short for rat sarcoma. In 1982, human oncogenes were found to be related to *RAS* oncogenes. With *NRAS*, a third *RAS* gene was identified only one year later in neuroblastoma cells. Shortly after this finding, *RAS* oncogenes were identified in carcinogen-induced animal tumors[19]. In 1984, a mutated *KRAS* oncogene was found in a biopsy sample from a lung cancer patient, but not in normal adjacent cells[20].

1.1.2.1 The mechanism of RAS activation

RAS proteins have been found to be important mediators of signaling pathways that contribute to cell cycle progression and cell proliferation, such as the *mitogen-activated protein kinase* (MAPK/ERK) pathway[21]. RAS proteins are molecular switches of low molecular weight (KRAS: 21.7 kDa, HRAS and NRAS: 21.3 kDa; Source: Genecards database), that alternate between active guanosine triphosphate (GTP)-bound and inactive guanosine diphosphate (GDP)-bound states. The conversion from the inactive GDP-bound form to the active GTP-bound form is triggered by guanine nucleotide exchange factors (GEFs, e.g., SOS1, SOS2, and RASGRF). In contrast, GTPase-activating proteins (GAPs) mediate the conversion back to the inactive form[22]. RAS proteins undergo posttranslational modifications, which are required to attach them to cellular membranes[23].

RAS proteins are turned on or off by GEFs and GAPs when they are translocated to the plasma membrane. Through their recruitment to the plasma membrane, GEFs and GAPs achieve high local concentrations in direct proximity to RAS. High local concentrations enable interactions with RAS despite its considerably higher affinity for RAF-kinases, which is in the low nanomolar range compared to the micromolar range for GEFs and GAPs[21].

Structural studies of RAS in complex with GAP and GEF revealed two switch regions, which undergo conformational changes between the GDP and GTP states[24]. The very

slow off-rate for GDP allows RAS proteins to remain in their inactive states until signals trigger a GDP to GTP exchange, which requires GEF activity to accelerate the exchange reaction by several orders of magnitude[25]. GEF binding to RAS results in conformational changes in the switch regions, resulting in GDP release and replacement with GTP[26]. Once GTP is bound to RAS, this leads to the dissociation of the GEF and subsequently to the formation of active GTP-bound RAS. Efficient GTP hydrolysis is slow and requires interaction with a GAP that accelerates the cleavage step by several orders of magnitude[25]. Subsequently, *RAS* is inactivated and again present in the inactive GDP-bound form[21].

The binding of nucleotides and subsequent hydrolysis has been mapped to a catalytic domain, which is highly homologous between KRAS, HRAS, and NRAS[27,28]. Furthermore, x-ray structures of the three proteins confirmed a high similarity. Differences between the three forms of RAS proteins are embedded in the C-terminal hypervariable region (HVR), where RAS proteins share only 15 % homology. The HVR consists of an unstructured linker region and a membrane-interacting lipid anchor. Both regions are involved in membrane interactions of RAS proteins, which is relevant for signal transduction[29,30].

RAS signaling can be activated by various upstream signals, which are mainly mitogenic factors binding to receptor tyrosine kinases (RTK)[19]. Once a ligand binds to an RTK, this triggers autophosphorylation of the receptor, which in turn allows the binding of adapter proteins, such as GRB2. The adapter proteins are capable of linking an RTK to GEFs like SOS1/2, thereby triggering the activation of RAS. It has been shown that GRB2 binds the autophosphorylated tail of the EGF receptor through a single SH2 domain while simultaneously binding to the GEF SOS1 through two SH3 domains, thereby facilitating the signal transduction from the receptor to the RAS protein[31].

RAS proteins, in turn, activate effectors by recruitment to the plasma membrane[21], and they have a central role in transducing mitogenic signals. To date, more than ten downstream effectors of RAS have been reported. The RAS effector interactions can be highly diverse, while the mechanism of RAS activation is highly conserved[21]. RAS effectors are involved in the activation of mitogenic pathways, cumulating in cell survival, proliferation, and migration. In particular, RAS-GTP complexes activate several downstream signaling effectors such as the canonical RAF-MEK-ERK, the PI3K-AKT-mTOR and RalGDS-RalA/B pathways[19,32] (Figure 2).

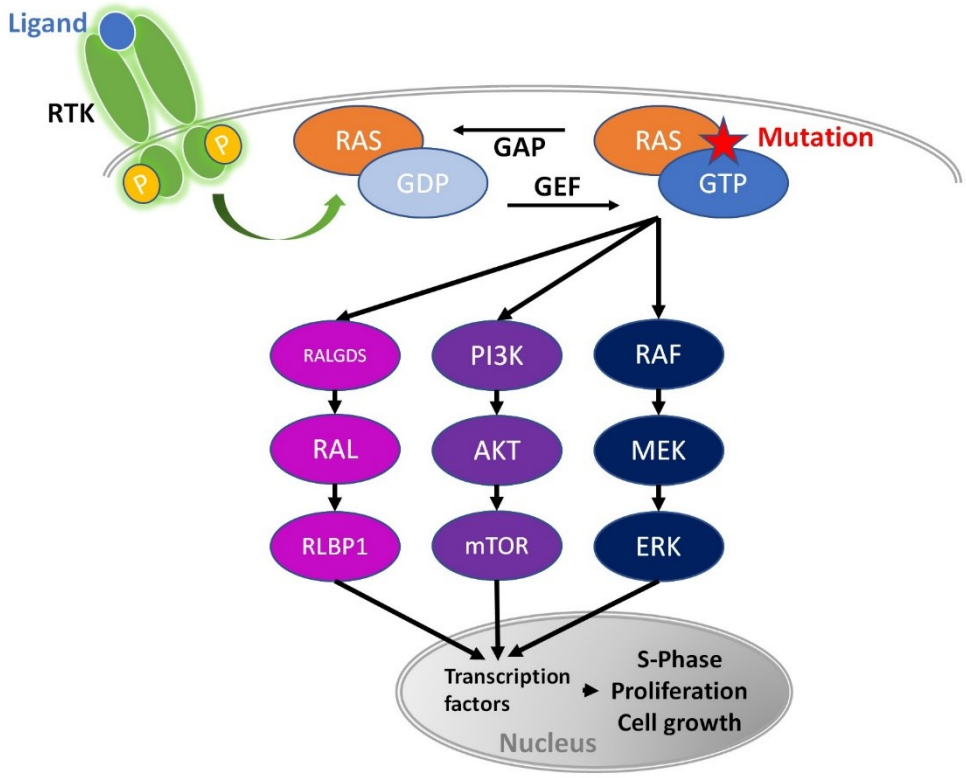


Figure 2 | Rat sarcoma viral oncogene homolog (RAS) signaling pathways

ERK, extracellular signal-regulated kinase; GAP, GTPase activating protein; GEF, guanine nucleotide exchange factor; MEK, MAPK/ERK kinase; mTOR, mammalian target of rapamycin; PI3K, phosphatidylinositol-4,5-bisphosphate 3-kinase; RAF, rapidly growing fibrosarcoma; RALGDS, RAL guanine nucleotide dissociation stimulator; RAS, rat sarcoma virus; RTK, receptor tyrosine kinase. (Adapted from Ferrer et al.[32])

In the *mitogen-activated protein kinase* (MAPK) pathway, the *rapidly growing fibrosarcoma* (RAF) protein is activated by RAS and phosphorylates the *MAPK/ERK kinase* (MEK). MEK then activates the *extracellular signal-regulated kinase* (ERK). ERK translocates to the nucleus to stimulate the expression of a set of genes involved in cell proliferation, survival, differentiation, and cell-cycle regulation[33,34].

Once activated by RAS, *Phosphatidylinositol-4,5-bisphosphate 3-kinase* (PI3K) phosphorylates *phosphatidylinositol 4,5-biphosphate* (PIP2) to *phosphatidylinositol 3,4,5-triphosphate* (PIP3). PIP3 activates the serine/threonine-specific protein kinase AKT via intermediates. AKT activation leads to stimulation of the cell-cycle progression, survival, and resistance to apoptosis through phosphorylation of the *mammalian target of rapamycin* (mTOR) and other physiological substrates[35,36].

1.1.2.2 *KRAS* as an oncogenic driver

KRAS is the most frequent mutated *RAS* isoform, with 91 % of cases harboring a missense mutation in pancreatic cancer, and 42 % in colon carcinoma cases. In lung cancers, a *KRAS* missense mutation is found in 25 - 33 % of cases[21]. In particular, oncogenic alterations in *KRAS* are frequently detected in NSCLC, but not SCLC[9], and are rarely found in Squamous-cell cancers[32]. Mutated KRAS differs functionally from the wild-type variant. In particular, the oncogenic forms prevent GAPs from increasing the catalytic rate of GTPase, thereby keeping KRAS in its constitutively GTP-bound active state. This, in turn, activates oncogenic pathways and cellular signal transduction without the existence of extracellular signals[37–40], resulting in uncontrolled cell proliferation[32].

KRAS mutant NSCLC frequently harbors co-mutations. Three expression-based clusters have been identified to date and are based on the presence of co-mutations in the *Serine/Threonine Kinase 11* gene (*STK11*) (KL subgroup), the *TP53* gene (KP subgroup), and *CDKN2A/B* inactivation plus low *thyroid transcription factor-1* (*TTF-1*) (KC subgroup)[18]. Different biological processes are related to these three clusters, and there is growing evidence that co-occurring *STK11* mutations are associated with reduced overall survival in *KRAS* mutant NSCLC[18,41].

KRAS mutations are typically found in NSCLC tumors from patients who heavily smoked, and only 5 to 10 % of *KRAS*-mutant lung cancers arise in never or light smokers[42]. In concordance, the frequencies of *KRAS* mutations range from 6.7 to 40.0 % for heavy smokers and from 2.9 to 11.4 % for never/light smokers[43]. Furthermore, smokers and never smokers have a different spectrum of mutations and codon variants in *KRAS*[32]. *KRAS* mutant tumors from smokers also exhibit a higher frequency of co-mutations in the *Tumor protein p53* (*TP53*) or *Serine/Threonine Kinase 11* (*STK11*)[44]. From a clinical perspective, *KRAS* mutant lung cancers have been associated with poorer overall survival than *KRAS* wild-type tumors[45,46]. However, other studies were not consistent with this finding, especially when taking the stage of the disease into consideration[47,48]. Moreover, in contrast to colorectal cancer, where *KRAS* mutations correlate with a poor response to targeted therapy for EGFR, no significant correlation of the *KRAS* mutation status and clinical outcomes were found for anti-EGFR therapy in advanced-stage *EGFR*-mutant NSCLC[49,50].

1.1.2.3 *KRAS* mutation subtypes

At least nine different *KRAS* mutations have been identified at codons 12, 13, and 61 in NSCLC[32]. The majority of them are missense mutations and occur in codon 12 of the *KRAS* gene of NSCLC tumors[51] (Figure 3).

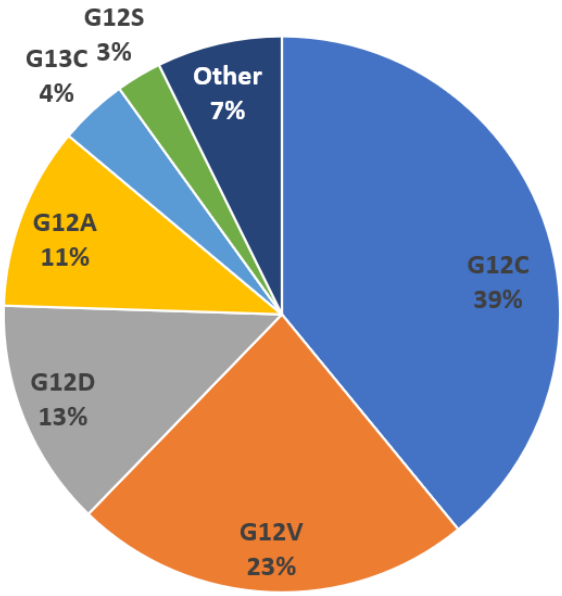


Figure 3 | Frequency of *KRAS* missense mutations in NSCLC

The frequency of *KRAS* missense mutations in NSCLC has been assessed in The Cancer Genome Atlas (TCGA) primary lung adenocarcinoma (n = 489) cohort. G12C mutations are most common (39 %), followed by G12V (23 %) and G12D (13 %).

According to *The Cancer Genome Atlas* (TCGA) database, the most frequently occurring *KRAS* mutation in NSCLC is G12C (39 %), followed by G12V (23 %), and G12D (13 %). Different amino acid substitutions in the oncogenic *KRAS* gene result in heterogeneity of downstream signaling mechanisms. For instance, *in vitro* analysis of cell lines showed that mutant *KRAS*-G12C or *KRAS*-G12V cell lines exhibit distinct downstream signaling properties. Compared to cells with other mutations or wild-type cells, both variants had decreased levels of phosphorylated AKT. In contrast, *KRAS*-G12D cell lines showed a higher affinity for PI3K-AKT signaling. Noteworthy, the affinity for the MAPK/ERK kinases was not altered with different types of *KRAS* mutations, but another study showed that tumors carrying *KRAS*-G12C exhibited higher ERK1/2 phosphorylation levels than tumors carrying *KRAS*-G12D[52]. In line with this observation, a study with genetically engineered mice showed that *KRAS*-G12C tumors were more sensitive to inhibition of MEK than tumors harboring a *KRAS*-G12D mutation[53].

Different amino acid substitutions in the KRAS protein may also implicate different clinical outcomes for NSCLC patients. For instance, a study with 179 surgically treated LUAD patients reported shorter overall survival (OS) of patients with *KRAS*-G12C mutations compared to those with G12D or G12A- but not with G12V mutations[54]. Although these results were confirmed by another study[52], further in-depth studies for both the adjuvant and advanced-stage settings are needed to fully understand the prognostic significance of different *KRAS* mutation subtypes in NSCLC[32].

1.1.2.4 Therapeutic strategies

Since its discovery of mutated KRAS as an oncogenic driver in cancer, many attempts have been made to leverage it as a therapeutic target. For decades, mutated KRAS has been considered undruggable due to its high affinity to GTP and the lack of a binding pocket that could be exploited by small molecules[51]. Several candidates for direct inhibition of KRAS have been reported, but almost all of them failed to proceed beyond initial characterization. In 2013, a new regulatory pocket adjacent to the cysteine residue of G12C mutated KRAS was identified[55] and led to the development of small molecule compounds that irreversibly bind to this pocket. Of note, these compounds bind to KRAS preferentially when GDP is bound and block the conversion to the GTP-bound state. Because the G12C mutation variant has a high intrinsic GTPase activity, the levels of KRAS-GTP would decrease over time[21]. However, follow-up studies indicated only limited potency of these compounds. Further research efforts in this direction resulted in the development of ARS853, which showed a reduction of KRAS-GTP levels by more than 90 % and potently suppressed MAPK and PI3K-AKT downstream signaling[51]. *In vivo* studies with the successor-compound ARS-1620 showed durable tumor regression in mice[56]. In 2018, the first clinical phase 1/2 study for solid tumors was launched for Sotorasib, a compound that led to the regression of KRAS-G12C mutant tumors and improved anti-tumor efficacy of chemotherapy and targeted agents in pre-clinical models[57]. Following successful clinical trials, thereby being the first of its kind, the compound was FDA-approved for treatment of *KRAS* G12C mutated NSCLC in 2021, followed by approval in the EU in January 2022. In addition, other compounds that target KRAS-G12C mutant tumors are being investigated in clinical trials[58].

Besides the direct approach to inhibiting KRAS, downstream effectors have been extensively investigated as therapeutic targets. Despite clinical efficacy in Renal-Cell-

Carcinoma, the RAF kinase inhibitor Sorafenib showed no clinical benefit and low response rates in *KRAS* mutant NSCLC and other solid tumors in clinical studies[21]. BRAF inhibitors have shown clinical activity in *BRAF* mutant NSCLC, but information on a therapeutic value in *KRAS* mutant NSCLC is lacking. Pre-clinical research, however, indicated that the serine/threonine-protein kinase CRAF, but not BRAF, might be a promising target for KRAS-G12V driven NSCLC.[59] Trametinib, an inhibitor of the MAPK/ERK kinase (MEK), failed to improve outcomes compared to standard docetaxel chemotherapy and showed an unfavorable toxicity profile[60]. Another MEK inhibitor, Selumetinib, failed to improve outcomes compared with standard second-line Pemetrexed-based chemotherapy and showed no clinical activity with a 0 % overall response rate[61]. Further clinical trials involved inhibitors for the *mammalian target of rapamycin* (mTOR), *focal adhesion kinase* (FAK, aka PTK2), and other downstream effectors of KRAS, but were also of little to no clinical benefit[32,62,63].

To leverage their potential, several combinations of downstream inhibitors with chemotherapy were tested for pre-clinical and clinical activity against *KRAS* mutant NSCLC. For instance, the addition of the MEK inhibitor Selumetinib increased the median overall survival and median progression-free survival of Docetaxel-treated patients with *KRAS* mutant NSCLC vs. Docetaxel alone, but this result could not be confirmed in a larger phase III study[64,65]. Pre-clinical studies suggested the need to inhibit multiple KRAS downstream effectors simultaneously to block downstream signaling effectively[66,67]. In murine models, combined inhibition of MEK and PI3K significantly improved the outcome compared to the single-compound treatment. However, clinical trials indicated that too high concentrations of the inhibitors would be required to block RAS signaling, which might restrict their therapeutic efficacy[32].

In recent years, monoclonal antibodies targeting the *Programmed cell death protein* 1 (PD-1) and the *Programmed cell death 1 ligand 1* (PD-L1) have influenced the treatment landscape of many cancers, including NSCLC. *KRAS* mutant NSCLC are attractive targets for PD-1/PD-L1 inhibition because they are associated with a high mutational burden[68] and are often characterized by a high expression levels of *PD-L1*[69] and a high number of tumor-infiltrating T-cells. Indeed, clinical studies reported significantly increased survival in advance-stage NSCLC patients, and patients with *KRAS* mutant tumors were among those who achieved the highest overall survival benefits[70].

Because KRAS and its downstream effectors showed to be challenging targets, several studies investigated synthetically lethal interactions with mutated *KRAS*. Depletion of such genes would result in decreased viability of *KRAS*-mutant tumor cells, but not of *KRAS* wild-type cells. For instance, it was shown that the fibroblast growth factor receptor 1 (FGFR1) is involved in resistance to MEK inhibition with Trametinib. Accordingly, the combined inhibition of FGFR1 and MEK showed elevated tumor regression in pre-clinical *KRAS*-G12V mutated tumor models[71]. Other reported targets for synthetically lethal interactions involve Polo-like-Kinase 1 (PLK1)[72], Serine/threonine-protein kinase 33 (STK33)[73], TANK Binding Kinase 1 (TBK1)[74]. However, no successful clinical trials have been reported so far[32,75].

An alternative strategy for the treatment of NSCLC has been focusing on exploiting the epigenetic alterations in tumor cells vs. non-malignant cells, which will be covered in the next chapter.

1.2 Epigenetics

Epigenetics describes chromatin modifications that influence gene expression and other DNA-dependent processes without altering the DNA sequence itself. In particular, epigenetic processes involve DNA methylation, histone modifications, and non-coding RNA-mediated (ncRNA) modifications[76,77] (Figure 4).

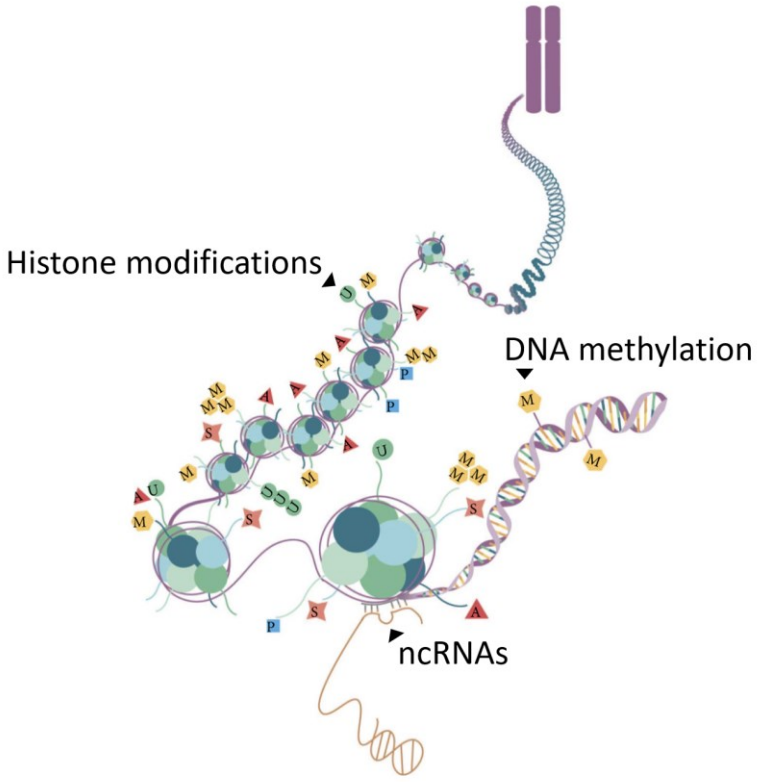


Figure 4 | An overview of major epigenetic mechanisms

Epigenetic mechanisms are reversible and do not affect the DNA coding sequence. Mechanisms include DNA methylation, histone/nucleosome modifications, and non-coding RNA-mediated (ncRNA) modifications and are involved in the regulation of gene expression and other DNA-dependent processes. (Adapted from Xiang Shi et al.[78])

DNA methylation is a process in which a methyl group is covalently added to cytosine, resulting in 5-methylcytosine (5mC). This modification is catalyzed by DNA methyltransferases (DNMTs)[79]. Five types of DNMTs have been identified. DNMT1 is required for copying the DNA methylation pattern to the newly generated DNA strand during DNA replication, whereas the *de novo* methyltransferases DNMT3A and DNMT3B act on unmethylated DNA[80].

Nucleosomes provide a compaction to DNA, which wraps around each nucleosome core approximately 1.8 times[81]. Nucleosomes are built of histone proteins. Five main families of histone proteins have been identified, comprised of the core histone families H2A, H2B, H3, and H4, and the linker histone family H1[82]. Histone proteins are prone to various modifications, including ubiquitylation, sumoylation, methylation, phosphorylation, and acetylation[83]. Unlike DNA methylation, histone modifications can promote not only the silencing of gene expression but also play an essential role in activated transcription[78]. In particular, whether a gene is activated or repressed depends

on the site of the histone methylation or acetylation[84]. Histone methylation is exerted by Histone methyltransferases (HMTs) and Histone demethylases (HDMs). HMTs facilitate the transfer of up to three methyl groups from S-Adenosyl methionine (SAM) to defined lysine and arginine residues of histone proteins, e.g., H3K4, H3K9, H3K27, H4R3, and others. Histone acetylation is catalyzed by Histone acetyltransferases (HATs) and Histone deacetylases (HDACs)[84,85].

Enzymes that are involved in epigenetic DNA and histone modifications can be classified into writers, editors, and readers. Writers establish a mark on either DNA or histone proteins, whereas editors modify or remove these marks. Readers mediate the interaction of a mark with proteins and thereby facilitate effects on transcription[86].

1.2.1 Epigenetic alterations in cancer

Several tumor entities have high frequencies of mutations in epigenetic writer-, editor-, and reader enzymes involved in DNA and histone modifications. In particular, cancers of the Ovar, Uterus, and Lung are among the tumor entities with the highest mutation frequencies of these mediators[86]. Mutations in genes coding for epigenetic enzymes contribute to several epigenetic alterations that have been identified in cancer cells.

In human cells, DNA methylation occurs in CpG dinucleotide islands, which account for roughly 1 % of the human genome and are enriched in gene promoter sequences[87]. In affected cancer cells, there is a reduced level of cytosine methylation in CpG islands, causing increased mitotic recombination and chromosomal instability[88,89]. Moreover, CpG islands of tumor suppressor gene promoters were found to be highly methylated in cancer cells, leading to their transcriptional repression. Other genes involved in DNA repair, apoptosis, or the epithelial-mesenchymal transition (EMT), were shown to be dysregulated in cancer cells by altered cytosine methylation[90,91].

Alterated histone methylation and acetylation modification patterns have been found in many tumor entities[84]. The global reduction of H4K20 trimethylation is a hallmark of many primary tumors[92]. HDACs are overexpressed in various cancers, and HDAC overexpression can result in silencing of tumor suppressor genes[93,94]. Moreover, mutations affecting HAT or HDAC enzymes have shown to be linked to aberrant gene expression signatures and are linked to carcinogenesis[95].

1.2.1.1 DNA methylation in NSCLC

In lung cancer, different tumor suppressor genes are silenced by promoter hypermethylation[96]. Many of these genes are involved in physiological cellular function, such as cell cycle regulation (*p16*), DNA repair (*MGMT*), and apoptosis (*DAPK*, *caspase-8*). In particular, a study found that methylation of *p16* and subsequent downregulation of the gene correlated with reduced survival after tumor resection in early-stage NSCLC[97]. Similarly, it has been shown that altered methylation of several genes in lung cancer correlated with fast recurrence following surgical resection of early-stage tumors in NSCLC patients. Importantly, this effect has been demonstrated regardless of other clinical factors such as histology, stage, gender, or smoking history[98]. Remarkably, alterations in DNA methylation in NSCLC patients might be linked to their smoking history. For instance, elevated expression of the DNA-methyltransferase *DNMT1* has been found in lung cancer patients with a smoking history, likely due to tobacco smoke ingredients that reduce the degradation of this enzyme[96,99]. Likewise, increased DNA methylation was shown to be caused by tobacco-induced chronic inflammation[100]. Recently, a study showed that *KRAS*-G12V overexpression in an isogenic lung model caused a strong global increase in differently methylated CpG sites[101].

1.2.1.2 Histone modifications in NSCLC

Altered epigenetic modifications of histones have been linked to lung cancer. For example, histone H4 was found to exhibit aberrant modifications compared to normal lung cells, and has been suggested to play an important role in lung carcinogenesis[102]. Based on these findings, H4K20me3 has been suggested as a potential biomarker for the early detection of and therapeutic approaches to lung cancer[102]. Other studies showed that elevated expression of *HDAC1* and *HDAC3* genes correlated with a poor prognosis in LUAD patients[103–105]. As for DNA methylation, the smoking history of patients might be connected to histone modifications in lung cancer patients, as studies showed that tobacco smoke ingredients induce H3K4 methylation[106–108].

1.2.2 Epigenetic alterations as treatment targets in NSCLC

Unlike DNA mutations, epigenetic dysregulation is transient and can be modulated by different pharmacologic approaches, with hypomethylating agents and histone

deacetylase inhibitors being the most commonly studied approaches[109]. Epigenetic therapy may help to overcome the problem of lung cancer heterogeneity by affecting the expression of multiple tumor suppressor genes simultaneously. Moreover, it may be an effective therapeutic approach for tumors that are driven by currently undruggable mutations, such as *KRAS* mutant NSCLC[109].

1.2.2.1 Targeting DNA methylation alterations

Inhibitors of DNA methyltransferases (DNMTi) can reverse DNA hypermethylation. Reversing the hypermethylation status of tumor suppressor genes has thus become a priority for the epigenetic treatment of tumors. 5-Azacytidine (AZA) and decitabine (DCN) are the most extensively used DNMTi in experimental and clinical studies[110,111]. Both drugs are incorporated into the DNA of a tumor cell, resulting in covalent non-reversible binding of DNMTs to the DNA, leading to a global reduction of DNA methylation. However, toxicity remains a significant problem with targeting DNA methylation. In a clinical study in which 103 NSCLC patients were treated with AZA, the overall response rate was only 8 %, and high toxicity was frequently observed. Similar results were observed in a study with DCN[112].

1.2.2.2 Targeting histone modification alterations

HDAC inhibitors (HDACi) have been shown to induce apoptosis, necrosis, and cell cycle arrest in tumor cells[94]. Furthermore, the overall tumor cell apoptotic threshold was observed to be lower after HDACi treatment in pre-clinical experiments due to the activation of death receptor-mediated and intrinsic mitochondrial pathways[109]. In a study where 16 NSCLC cell lines were treated with the HDAC inhibitors Trichostatin A (TSA) and vorinostat (VST), both compounds displayed anti-tumor activities in half of all cell lines[113]. Another study demonstrated that treatment with TSA led to increased apoptosis in H157 lung cancer cells and an increase in the sub-G0/G1 fraction[114]. However, a phase II clinical study with VST in NSCLC was terminated with a 0 % response rate[115]. Novel HDAC inhibitors have been developed since then, including N-Hydroxy-4-(4-phenylbutyryl-amino) benzamide (HTPB). HTPB suppressed the growth of lung cancer models by inducing cell cycle arrest and mitochondrial-mediated apoptosis *in vivo*[116], but follow-up clinical studies have not been reported to date.

The histone deacetylase inhibitor Entinostat inhibits HDAC1 and HDAC3 but is not sensitive to other HDACs[117]. The compound has demonstrated activity against a broad panel of cell lines from different tumor origins *in vitro*. A clinical combination study of Entinostat with low-dose AZA in heavily pre-treated lung cancer patients resulted in a low response rate but also showed one case of durable response as well as a high fraction of stable disease for at least 12 weeks among the responders[117]. However, another clinical trial using a combination of Entinostat and the EGFR-inhibitor Erlotinib in NSCLC patients harboring an *EGFR* mutation did not increase outcomes compared to treatment with Erlotinib alone[118].

The histone-lysine N-methyltransferase *Enhancer of zeste homolog 2* (*EZH2*) catalyzes the transfer of methyl groups to lysine residue 27 (K27) of histone H3. EZH2 is integral to the polycomb repressive complex 2 (PRC2) that is involved in the repression of gene transcription. *EZH2* is frequently upregulated in lung cancer; moreover, *EZH2* expression is induced by oncogenic *KRAS* during tumor initiation in autochthonous models[119]. Another study demonstrated that *EZH2* mRNA levels are associated with a poor prognosis in lung cancer patients and that downregulation of *EZH2* impairs cell viability in lung cancer cell lines[120]. Taken together, these findings make *EZH2* an attractive target for therapeutic intervention in NSCLC. The small molecule compound 3-Deazaneplanocin A (DZNep) acts as an inhibitor of *S-adenosylhomocysteine hydrolase* (AHCY) and modulates PRC2 through downregulation of *EZH2*[121]. In 2012, a study showed that DZNep inhibits the growth of a subset of NSCLC cell lines *in vitro*[122]. It was also found that DZNep treatment or *EZH2* depletion decreased H3K27 trimethylation[123]. Importantly, inhibition through DZNep can sensitize *BRG-1 (SMARCA4)* and *EGFR*-mutant cells to topoisomerase inhibitors[124]. Clinical studies with DZNep have not been reported to date[125,126], presumably due to the high toxicity of the compund[127]. Therefore, the identification of genes that act synthetically lethal with DZNep might allow for using lower concentrations by increasing the sensitivity of the target cells for the drug.

1.3 CRISPR/Cas9 gene editing

1.3.1 The CRISPR/Cas9 technology

The *Clustered Regularly Interspaced Short Palindromic Repeats* (CRISPR) and *CRISPR-associated Protein* (Cas) system has been discovered as a prokaryotic adaptive immunity mechanism used to cleave viral nucleic acids[128]. A variety of CRISPR/Cas systems has been identified in diverse species of bacteria and archaea. To date, six CRISPR/Cas types and at least 29 subtypes have been described[129–131].

A universal principle of all subtypes is the reliance on CRISPR RNA (crRNA) for target specificity. This principle has been exploited for laboratory applications by the design of guide RNAs (gRNA), which enable target specific DNA cleavage under experimental conditions. In particular, any locus containing an *adjacent protospacer motif* (PAM) can be targeted for cleavage by CRISPR/Cas systems with appropriate spacer sequences (Figure 5)[132].

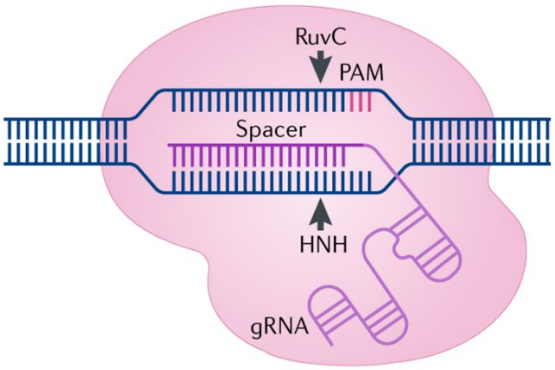


Figure 5 | Target specificity of Cas9 proteins

Cas9 interacts with the backbone of the guide RNA (gRNA). The spacer fragment of the gRNA binds to a DNA target sequence next to a 5′ protospacer adjacent motif (PAM). Two Cas9 nuclease domains, RuvC and HNH, catalyze the formation of a blunt DNA double-strand break. (Adapted from Pickar-Oliver and Gersbach[132]).

The Cas9 nuclease originated from *Streptococcus pyogenes* (SpCas9), which was the first Cas nuclease adapted to genome editing outside of prokaryotic cells in 2013[133–135]. It remains the most widely used Cas nuclease for genome editing[132]. SpCas9 specifically targets DNA by recognizing a 5' PAM-NGG sequence (N=any nucleotide, G=Guanin) and subsequent base-pairing of a 20-nucleotide long spacer element[133,136]. Since Cas9 systems require two crRNAs for the activation of nuclease activity, single-

guide RNAs (sgRNA) have been engineered to combine the functions of both crRNAs into one RNA molecule[133].

Once a target site has been recognized, Cas9 generates DNA breaks that induce the DNA damage response and thereby trigger the error-prone non-homologous end joining (NHEJ) pathway, resulting in the accumulation of small insertions or deletions (indels)[137]. When a coding exon is targeted, indel-mediated frameshift mutations, which in turn introduce stop codons downstream of the target site, will disrupt gene expression[134,138]. This mechanism can be utilized to specifically abrogate the expression of genes to study their biological implications[132]. Compared to previous methods for gene editing, the target specificity of CRISPR/Cas9 relies solely on base pairing of nucleic acids rather than on protein–DNA recognition, as it is the case in zinc finger nucleases. This paradigm shift has transformed genome editing by removing the need for extensive expertise in engineering custom-targeted DNA-binding proteins[132].

1.3.2 Pooled CRISPR/Cas9 knockout screens

In contrast to single gene depletion workflows, pooled screens facilitate thousands of genetic perturbations in a single experiment. Until the availability of CRISPR/Cas-mediated gene knockout screens, RNA interference (RNAi) has been the most commonly used system for such large-scale screens. However, the RNAi technology comes with limitations such as incomplete suppression of target genes and frequent off-target effects. Cas9-based high-throughput screens facilitate a knockout of genes compared to a knockdown in RNAi-based screens, which, combined with higher target specificity, results in less false-negative hits and higher rates of target validation[139,140]. However, to achieve robust statistical results, CRISPR/Cas9-based pooled screens demand a high number of sgRNA-transduced cells, which can impair the facilitation of screening experiments[132].

Because sgRNAs can be efficiently assessed with high-throughput sequencing in the treated vs. the non-treated study arm of a screen, the CRISPR/Cas9 technology provides a convenient and robust platform to study thousands of genes simultaneously[141]. A common readout strategy for pooled CRISPR/Cas9 screens is to identify gene depletions that decrease the survival or proliferation of cells, either by gene knockdown alone, or combined with drug treatment. The latter identifies synthetically lethal genes

with a drug, thereby contributing to understanding the biological mechanisms of drug resistance.

1.4 Aim of this Ph.D. thesis

Approximately 10 % of all lung cancer cases harbor a mutation in the oncogenic *KRAS* gene. Various studies have been performed to identify novel targets for the treatment of NSCLC, including the development of compounds that exploit epigenetic alterations in the tumor cells but leave non-malignant cells unaffected. However, the clinical success of these compounds has been minimal, with no exception for the *KRAS* mutant non-small cell lung cancer (NSCLC) subtype.

This thesis aimed to identify genes that act synthetically lethal with the epigenetic drugs 3-Deazaneplanocin A (DZNep) and Entinostat in *KRAS*- and *EGFR* mutant NSCLC cell lines. Therefore, a pooled CRISPR/Cas9 screen targeting 6,500 genes was established in two NSCLC cell lines, one of which harbors a *KRAS* mutation, and the other harbors an *EGFR* mutation, and the cells were screened for synthetically lethal genes with epigenetic drug treatment. In parallel, the study aimed to identify novel essential viability genes in both cell lines.

Building on the screening results, this study further aimed to characterize cell viability and apoptosis after epigenetic drug treatment in conjunction with siRNA-mediated knockdown of the identified target genes. Finally, gene expression profiling was to be utilized to elucidate possible mechanisms of synthetic lethality between gene knockdown and epigenetic drug treatment, pointing to potential effectors of synthetic lethality.

2 Materials and Methods

2.1 Materials

2.1.1 Consumables

Table 1 | Consumables

Description	Supplier
Bacterial culture tube, 12mL	Greiner
BT Barrier Pipet Tips, Pre-Sterile (10 µL, 20 µL, 100 µL, 200 µL, 1000 µL)	Neptune Scientific
Cellstar® TC Cell culture flask (T25 = 25 cm^2, T75 = 75 cm^2, T175 = 25^2)	Greiner Bio One International
Costar® reagent reservoir (50 mL)	Corning
Costar® serological pipettes (5 mL, 10 mL, 25 mL, 50 mL)	Corning
Countess™ Cell counting chamber slides	Invitrogen Corporation
CulturePlate-96 Black	Perkin Elmer
FACS tube, 5mL Polystyrene Round Bottom Tube	BD
Falcon™ Conical Centrifugation Tube (15 mL, 50 mL)	Thermo Fisher Scientific
Falcon™ Polystyrene 6-well plate, Flat bottom, Clear	Thermo Fisher Scientific
Falcon™ Polystyrene 12-well plate, Flat bottom, Clear	Thermo Fisher Scientific
Falcon™ Polystyrene 24-well plate, Flat bottom, Clear	Thermo Fisher Scientific
Falcon™ Polystyrene 96-well plate, Flat bottom, Clear	Thermo Fisher Scientific
LightCycler® 480 Multiwell Plate 384, white	Roche
LightCycler® 480 Sealing Foil	Roche
Millex-GP syringe filter unit (0.22 µm)	Sigma-Aldrich
Nunc® CryoTubes® (2.0 mL)	Sigma-Aldrich
Nunc™ Cell Culture/Petri Dishes	Sigma-Aldrich
PCR microtube strips (0.2 mL)	Sigma-Aldrich
Petri dish, 6cm	Greiner
PIPETMAN® Classic - 10 µL, 20 µL, 100 µL, 200 µL, 1000 µL	Gilson
Safe-Lock microcentrifuge tubes (0.2 mL, 0.5 mL, 1.5 mL, 2.0 mL, 5.0 mL)	Eppendorf Corporation
Serological pipettes (2.5 mL, 5 mL, 10 mL, 25 mL)	BD Falcon
Syringe PP/PE without needle, Luer Lock tip (50 mL)	Sigma-Aldrich
XCEED Soft-Nitril-Glove L Powder Free	Microflex

2.1.2 Chemicals

Table 2 | Chemicals

Description	Supplier
2-Propanol (Isopropanol)	Sigma-Aldrich
3M sodium acetate	Sigma-Aldrich
Ethanol	Sigma-Aldrich
Methanol	Sigma-Aldrich
Sodium Chloride	Sigma-Aldrich
Sodium Hydroxide	Sigma-Aldrich

2.1.3 Instruments

Table 3 | Instruments

Description	Supplier
Agilent 2100 BioAnalyzer	Agilent Technologies
Avanti J-30I centrifuge	Beckmann-Coulter
Axiovert 40 CFL (Inverse Microscope)	Carl Zeiss
BBD 6220 CO_2-Incubator	Thermo Fisher Scientific
BD FACSCanto II	BD Biosciences
Bio Dancer Shaker	New Brunswick Scientific
Clean PCR workstation	Peqlab
Countess II FL Automated Cell Counter	Invitrogen
Ecotron (Incubation shaker)	Infors
Electrophoresis chamber	Renner
Electrophoresis Power Supply - EPS 301	Amersham pharmacia biotech
FS-300N Ultrasonic Homogenizer Sonicator	Vevor
GeneTouch Thermal Cycler	Bioer Technology
Heraeus™ Fresco™ Microcentrifuge	Thermo Fisher Scientific
Heraeus™ Megafuge™ 16 (Centrifuge)	Thermo Fisher Scientific
HiSeq2000 High throuput sequencing device	Illumina
Infinite® M200 (Multiplate Reader)	Tecan Group
JS-24 rotor, fixed angle	Beckmann-Coulter
LightCycler®480 II Real Time PCR System	Roche
LKB - GPS 200/400 (Electrophoresis Power Supply)	Pharmacia
Maxisafe 2020 Class II Biological Safety Cabinets	Thermo Fisher Scientific
Micropipette manual, Pipetman	Gilson
Microwave	Sharp
Nalgene polycarbonate centrifuge tubes, 30 mL	Thermo Fisher Scientific
Nalgene Mr. Frosty (Freezing container)	Thermo Fisher Scientific
NanoDrop 1000 Spectrophotometer	Peqlab Biotechnologie
Picus® NxT - 120 µL, 300 µL, 1000 µL	Sartorius
PIPETBOY acu 2	INTEGRA Biosciences
QuantStudio 3D Digital PCR Instrument	Thermo Fisher Scientific
QuantStudio 3D Digital PCR 20K Chip	Thermo Fisher Scientific
Quantum-ST4	Vilber Lourmat

Instruments continued	
Description	**Supplier**
Qubit Fluorometric Quantification Device	Thermo Fisher Scientific
Rainin Pipet-Lite™ Multichannel Pipette (200 µL)	Mettler-Toledo Rainin
ShakeTemp SW22 Water bath	Julabo
Thermomixer comfort	Eppendorf
TKA GenePure xCAD Dispenser	Huberlab
Vortex-Genie 2	Scientific Industries
ProFlex 2x Flat PCR System	Thermo Fisher Scientific
4-16 Centrifuge	Sigma Laborzentrifugen
PLJ 3500-2NM	Kern & Sohn

2.1.4 Drugs

Table 4 | Drugs

Description	Supplier
3-Deazaneplanocin A hydrochloride	Sigma-Aldrich
Entinostat	Selleckchem
Hygromycin B	Life Technologies
Puromycin	Sigma-Aldrich
Staurosporine	Roche

2.1.5 Reagents for molecular biology

Table 5 | Reagents for molecular biology

Description	Supplier
100 bp DNA Ladder (500 µg/mL)	New England BioLabs
10x Titanium® Taq PCR Buffer	Takara Bio
150 mM NaCl Solution Sterile Filtered (0.2 µm)	Polyplus-transfection® SA
1kb DNA Ladder (500 µg/mL)	New England BioLabs
7-Aminoactinomycin D for flow cytometry	Sigma-Aldrich
Agarose NEEO Ultra-Quality	Roth
Ambion ddH_2O	Thermo Fisher Scientific
Ampicillin	Sigma-Aldrich
AnnexinV Binding Buffer 5x	Thermo Fisher Scientific
AnnexinV - Phycoerythrin for flow cytometry	Sigma-Aldrich
Blood and Cell Culture DNA Midi Kit	Qiagen
CellTiter blue cell viability assay	Promega
Dimethyl sulfoxide	Sigma-Aldrich
DNAase I (RNase-free)	New England BioLabs
dNTP Mix (2.5 mM)	Invitrogen
Ethidium bromide solution 1 %	Roth
Glutamax	Life Technologies
Lipofectamine transfection reagent	Invitrogen
Lentifuge Viral Concentration Reagent	Cellecta

Reagents for molecular biology continued	
Description	**Supplier**
$MgCl_2$ solution, 5mM, sterile	Sigma-Aldrich
Packaging Plasmid Mix	Cellecta
Penicillin Streptomycin Solution	Life Technologies
Phenol:Chloroform:Isoamyl Alcohol 25:24:1	Sigma-Aldrich
Plasmid Maxi Kit	QIAGEN
PlusReagent	Life Technologies
Polybrene	Sigma-Aldrich
Proteinase K	Sigma-Aldrich
QIAquick gel purification kit	QIAGEN
QIAquick PCR purification kit	QIAGEN
QuantStudio™ 3D Digital PCR Master Mix	Thermo Fisher Scientific
RediPlate™ 96 RiboGreen ® RNA Quantitation Kit	Molecular Probes
RevertAid H Minus First Strand cDNA Synthesis Kit	Thermo Fisher Scientific
RNeasy Mini Kit	QIAGEN
Sodium dodecyl sulfate (SDS) Pellets	Roth
TaqMan Genotyping Assay	Thermo Fisher Scientific
Titanium® Taq DNA Polymerase	Takara Bio
Titanium® Taq PCR Buffer	Takara Bio
Tris-acetate-EDTA (TAE) Buffer (50x)	Thermo Fisher Scientific
Universal Probe Libraries for Real-time RT-qPCR No. 65 (*CHD8*), 20 (*EP300*), 39 (*CCL20*)	Roche

2.1.6 TaqMan genotyping assay

Table 6 | TaqMan genotyping assay

Gene	Assay ID	COSMIC ID	Amino acid change	Nucleotide change	Wild type allele (VIC label)	Mutant allele (FAM label)
KRAS	AH0JEUD	COSM516	p.G12C	c.34G>T	G	T

2.1.7 Primer

Primers for amplification of sgRNA and high throughput sequencing were HPLC purified by the supplier.

Table 7 | Primer for sgRNA amplification and high throughput sequencing

Name	Target sequence	Supplier
NGS_FSeq-gRNA	GGCTTTATATATCTTGTGGAAAGGACGAAACACCG	Sigma-Aldrich
First-Round_FwdU6-1	CAAGGCTGTTAGAGAGATAATTGGAA	Sigma-Aldrich
Sec-Round_P5-NFG16	AGATACGGCGACCACCGAGATCTACACGAGATGGACTATCATATGCTTAC CGTAACTTGAA	Sigma-Aldrich
First-Round_Rev1	CGACAACAACGCAGAAATTTTGAAT	Sigma-Aldrich
Sec-Round_P7-NRG1_6IND-A	CAAGCAGAAGACGGCATACGAGATGCACGACGAGACGCAGACGAATACGA CACAGTGTATGTCTGTTGCTATTATGTCTACT	Sigma-Aldrich
Sec-Round_P7-NRG1_6IND-B	CAAGCAGAAGACGGCATACGAGATGCACGACGAGACGCAGACGAACTGAT GACAGTGTATGTCTGTTGCTATTATGTCTACT	Sigma-Aldrich
Sec-Round_P7-NRG1_6IND-C	CAAGCAGAAGACGGCATACGAGATGCACGACGAGACGCAGACGAAGCATC AACAGTGTATGTCTGTTGCTATTATGTCTACT	Sigma-Aldrich
Sec-Round_P7-NRG1_6IND-D	CAAGCAGAAGACGGCATACGAGATGCACGACGAGACGCAGACGAAAGTCG TACAGTGTATGTCTGTTGCTATTATGTCTACT	Sigma-Aldrich
NGS-Index_P7-RSeqIND	AGATGCACGACGAGACGCAGACGAA	Sigma-Aldrich

Primers for real-time RT-qPCR were desalted by the supplier.

Table 8 | Primer for real-time RT-qPCR

Name	Gene symbol	Target sequence	Corresponding UPL (Roche)	Supplier
*CHD8*_1_NM_001170629.1_fwd	*CHD8*	AGCCTGCAGTTACACTGACCT	65	Sigma-Aldrich
*CHD8*_1_NM_001170629.1_rev	*CHD8*	TTGGGGACCTCCAGACTGT	65	Sigma-Aldrich
*EP300*_fwd	*EP300*	GCAGCCTGCAACTCCACT	20	Sigma-Aldrich
*EP300*_rev	*EP300*	GAGGATTTGATACCTGTCCTTCA	20	Sigma-Aldrich
*CCL20*fwd	*CCL20*	GCTGCTTTGATGTCAGTGCT	39	Sigma-Aldrich
*CCL20*rev	*CCL20*	GCAGTCAAAGTTGCTTGCTG	39	Sigma-Aldrich
GAPDH-fwd	*GAPDH*	AGCCACATCGCTCAGACAC	60	Sigma-Aldrich
GAPDH-rev	*GAPDH*	GCCCAATACGACCAAATCC	60	Sigma-Aldrich

2.1.8 Cell culture

Table 9 | Cell lines

Description	Driver mutation	Supplier
HEK-293T/17 (CRL-11268)	N/A	ATCC
NCI-H1944 (CRL-5907)	KRAS p.G13D	ATCC
NCI-H1975 (CRL-5908)	EGFR p.L858R, p.T790M, TP53 p.R273H	ATCC
NCI-H2030 (CRL-5914)	KRAS p.G12C, TP53 p.G262V	ATCC

Table 10 | Consumables for cell culture

Description	Supplier
Dimethyl sulfoxide analytical standard	Sigma-Aldrich
Dulbecco's Modified Eagle Medium	Thermo Fisher Scientific
Gibco™ Dulbecco´s Phospate-Buffered Saline (no calcium, no magnesium)	Thermo Fisher Scientific
Gibco™ Fetal Bovine Serum, qualified, heat inactivated, E.U.-approved, South America Origin	Thermo Fisher Scientific
Gibco™ RPMI 1640 medium	Thermo Fisher Scientific
0.25 % Trypsin-EDTA Solution for Cell Culture	Thermo Fisher Scientific
HEPES Buffer 1M, aseptically filled	Sigma-Aldrich
Trypan Blue Solution (0.4 %)	Sigma-Aldrich

2.1.9 Gene silencing and knockout

Table 11 | Small interfering RNAs

All siRNAs were purchased in a quantity of 1 nmol.

Description	Target sequence	Supplier
Gene Solution siRNA Hs_*CCL20*_7	CAGACCGTATTCTTCATCCTA	Qiagen
Gene Solution siRNA Hs_*CCL20*_6	AAGGCTGTGACATCAATGCTA	Qiagen
Gene Solution siRNA Hs_*CCL20*_5	AACAGACTTGGGTGAAATATA	Qiagen
Gene Solution siRNA Hs_*CCL20*_1	CTCACTGGACTTGTCCAATTA	Qiagen
Gene Solution siRNA Hs_*CHD8*_9	CCCAAGGATCGTGTCCTGATA	Qiagen
Gene Solution siRNA Hs_*CHD8*_8	CAGAACATACCTCGAGTCTTA	Qiagen
Gene Solution siRNA Hs_*CHD8*_7	CAAGATAGTGTTACAGGGCAA	Qiagen
Gene Solution siRNA Hs_*CHD8*_7	CCGGATCCCTGTCATCAATAA	Qiagen
Gene Solution siRNA Hs_*EP300*_10	GTCCTACGCAGTCAACAGGTA	Qiagen
Gene Solution siRNA Hs_*EP300*_9	GTGCTAGTCCTATGGGAGTAA	Qiagen
Gene Solution siRNA Hs_*EP300*_7	GTGCATCTTCTCGACAAATAA	Qiagen
Gene Solution siRNA Hs_*EP300*_3	GAAACAGTGGCACGAAGATAA	Qiagen
AllStars Negative Control siRNA	Not disclosed by supplier, stated as having no homology to any known mammalian gene	Qiagen

Table 12 | Materials for pooled CRISPR/Cas9 screening

Description	Supplier
CRISPR Human Genome Knockout Library, Module 1 (55K) in pRSG16-U6-sg-HTS6C-UbiC-TagRFP-2A-Puro vector (see supplement materials for vector map). 200 µg Plasmid in 200 µL solution. Lot # 15082402	Cellecta
CRISPRtest™ Functional Cas9 Activity Kit	Cellecta
Packaging Plasmid mix (psPAX2, pMD2.G), 250 µg, 0.5 µg/µL, Lot #: 13040801	Cellecta
pR-CMV-Cas9-2A-Hygro plasmid	Cellecta

2.1.10 Software

Table 13 | Software

Description	Supplier
2100 Expert Software	Agilent Technologies
AxioVision Rel. 4.8	Carl Zeiss
COSMIC Cell Lines Project Database v90	Wellcome Sanger Institute
CRISPRAnalyzer v.1.5 (www.crispranalyzer.com)	Developed by Jan Winter and Marc Schwering, Boutros Lab, DKFZ Heidelberg
DAVID Bioinformatics Resources 6.8	Laboratory of Human Retrovirology and Immunoinformatics (LHRI), Frederick National Laboratory for Cancer Research
FACSDiva Software V. 8.0	BD Biosciences
FlowJo® 10.4	FlowJo
i-control 1.6	Tecan
Image Lab™ 6.0	Bio-Rad Laboratories
LightCycler®480 Software release 1.5.0	Roche
NanoDrop ND-1000 V3.7.1	Peqlab Biotechnologie
Prism 8	GraphPad
QuantStudio Analysis Suite software	Thermo Fisher Scientific
SnapGene	GSL Biotech
Universal Probe Library Assay Design Center	Roche

2.2 Methods

2.2.1 Cell cultivation

All work was performed in Class II Biological Safety Cabinets. Unless stated otherwise, cell culture medium as well as Dulbecco´s Phospate-Buffered Saline (DPBS) and Trypsin-EDTA were heated to 37 °C before they were applied to cells. All cell lines were cultured at 37 °C and 5 % CO_2 in the BBD 6220 CO_2-Incubator.

2.2.1.1 Passaging of cells

NCI-H2030, NCI-H1975, and NCI-H1944 cell lines were purchased from the American Type Culture Collection (ATCC) and cultured in RPMI 1640 medium, supplemented with 10 % fetal bovine serum (FBS). HEK-293T/17 cells were purchased from the ATCC and cultured in Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10 % FBS.

Cells were grown to 90 % confluency in T25, T75, and T175 cell culture flasks. When reaching the desired confluency level, the medium was aspirated and the cells were washed twice with 3 mL (T25 flasks), 7 mL (T75 flasks), or 10 mL (T175 flasks) of DPBS. The cells were dissociated from the culture flask surface by incubation with 1 mL (T25 flasks), 2 mL (T75 flasks), or 4 mL (T175 flasks) of Trypsin-EDTA for 3 - 10 minutes. Incubation was stopped by addition of cell culture medium to inactivate Trypsin-EDTA once all cells were detached. The cell suspension was transferred in a 15 mL Falcon™ Conical Centrifugation Tube and centrifuged at 400 × g for 5 minutes in a Heraeus Megafuge 16 centrifuge. The supernatant was aspirated and discarded, and the cell pellet was resuspended in cell culture medium.

Following this, the cells were diluted in cell culture medium and seeded back into the cell culture flasks at a density that would result in 90 % confluency after 5 - 7 days. Alternatively, the cells were counted and used for subsequent experiments. After a maximum of ten passages, cells were discarded and replaced by an aliquot of freshly thawn cells.

Cells were regularly tested for Mycoplasma contamination by submission to an external facility (Multiplexion GmbH, Heidelberg), following the provided protocol for sample preparation. Cell lines were authenticated before use. For this, DNA was extracted following the manufacturer’s instructions.

2.2.1.2 Cell counting

To establish a defined number of cells in an experiment, cells were washed with DBPS, trypsinized, centrifuged, and taken up in cell culture medium as described in the previous section. Following this, 10 µL of the cell suspension was transferred in a 0.2 mL Safe-Lock Microcentrifuge tube and mixed with 10 µL of 0.4 % Trypan Blue Solution. Ten µL of the mixture was loaded into a Countess Cell counting chamber slide. The fraction of living cells was determined by the Countess II FL Automated Cell Counter.

2.2.1.3 Freezing and thawing of cells

To generate long-term stocks of all cell lines, cells were washed with DBPS, trypsinized, centrifuged and taken up in cell culture medium as described above. Following this, the cells were counted as described in the previous section and diluted in cell culture media to a concentration of 1 x 10^6 cells per mL. The cell suspension was transferred into a 2.0 mL Nunc CryoTube in a quantity of 1.0 mL per tube. Each tube was supplemented with 10 % dimethyl sulfoxide (DMSO), mixed by pipetting up-and-down several times, placed into a Mr. Frosty freezing container and stored at -80 °C. After 24 hours, the CryoTubes were transferred into a liquid nitrogen tank.

Previously stored cells were suspended in 3 - 5 mL of cell culture medium by gradually adding 37 °C pre-warmed media and slowly pipetting up-and-down several times. While gradually adding cell culture media, the already thawn cells were transferred into a 15 mL Falcon Conical Centrifugation Tube, and remaining frozen cells were submitted to additional volumes of pre-warmed media. Once all cells were thawn and transferred into the centrifugation tube, they were centrifuged at 400 × g for 5 minutes in the Heraeus Megafuge 16 centrifuge, followed by the removal of the supernatant. The cell pellet was re-suspended in cell culture medium, transferred into a T25 or T75 cell culture flask and incubated at 37 °C until 90 % confluency. Following this, the cells were passaged as described above.

2.2.2 Generation of Cas9 stable cells

In order to perform a knockdown of genes, the CRISPR technology requires Cas9 nuclease to be present in the target cell. Hence, the cell lines in scope for pooled gene knockout screening were stably transfected to express Cas9.

2.2.2.1 Production of lentiviral Cas9 particles

A total of 12.5 × 10^6 HEK-293T/17 cells was seeded in a total of four T175 flasks with 25 mL DMEM per flask. The cells were incubated for 24 hours at 37 °C. The following day, 240 µL Lentiviral Packaging Plasmid Mix, 24 µL pR-CMV-Cas9-2A-Hygro plasmid (1 µg/µL) and 240 µL PLUS Reagent were mixed with 4.8 mL DMEM without serum or antibiotics in a 15 mL Falcon Conical Centrifugation Tube. The mixture was incubated at room temperature for 15 minutes. In parallel, 360 µL Lipofectamine was diluted in 4.8 mL DMEM in a separate 15 mL Falcon Conical Centrifugation Tube. The diluted Lipofectamine was added to the DNA / PLUS Reagent complex, vortexed thoroughly and incubated at room temperature for 15 minutes. Two and a half mL of the resulting DNA / PLUS Reagent / Lipofectamine complex was added to each T175 flask and cells were incubated for 24 hours at 37 °C. Twenty-four hours post-transfection, the medium was replaced with 25 mL of fresh DMEM supplemented with 10 % FBS, DNase I (1 U/mL), $MgCl_2$ (5 mM), and 20 mM HEPES, pH 7.4, followed by incubation for 24 hours at 37 °C. Forty-eight hours post-transfection, the virus-containing medium was aspirated and filtered into two 50 mL Falcon Conical Centrifugation Tubes through a 0.22 µm Millex-GP syringe filter unit to remove debris and floating cells.

To increase the concentration of lentiviral Cas9 particles, 16 mL of the virus-containing medium was transferred into six Nalgene polycarbonate centrifuge tubes each (capacity: 30 mL). Each tube was supplemented with 16 µL of LentiFuge Viral Concentration Reagent and incubated for 1 hour at 4 °C. Afterwards, the lentiviral supernatant / LentiFuge mixture was centrifuged at 40,000 × g for 8 hours at 4 °C in an Avanti J-30I centrifuge using a JS-24 fixed angle rotor. Following centrifugation, the supernatant was aspirated carefully and discarded. The remaining faint pellet was resuspended in 35 µL of DPBS and transferred into 0.2 mL Safe-Lock microcentrifuge tubes in aliquots of 20 µL.

2.2.2.2 Titer estimation of lentiviral Cas9 particles

To generate cells that stably express Cas9 nuclease, first, the titers of lentiviral Cas9 particles in different cell lines were established. For this, H2030 or H1975 cells were resuspended to a density of 1 × 10^5 cells per mL in RPMI 1640 medium supplemented with 10 % FBS and 5 µg/mL Polybrene. One hundred µL (1 × 10^4 cells) of the cell suspension were aliquoted per well in a 96 well plate. Zero µL, 1 µL, 3.3 µL and 10 µL

of a 1:10 dilution of Cas9 lentiviral stock were added to four different wells each. The suspension was carefully mixed by pipetting up-and-down and incubated at 37 °C. After 24 hours, the medium was replaced with fresh RPMI 1640 medium supplemented with 10 % FBS without Polybrene. The cells were incubated under the same conditions for additional 24 hours. Next, the medium was replaced with fresh RPMI 1640 medium supplemented with 10 % FBS. For each dilution of the Cas9 lentiviral stock solution, two of four wells were treated with 600 µg/mL Hygromycin (for H2030 cells) or 400 µg/mL Hygromycin (for H1975 cells). The cells were incubated at 37 °C. After 72 hours, the cells were harvested, and the cell viability was assessed using the CellTiter blue assay, as described in section 2.2.7.1. For each dilution of the lentiviral Cas9 stock solution, the Hygromycin-treated cells were compared to the wells without Hygromycin. The percent transduction for each dilution was calculated according to

$$Percent\ Transduction = 100 \times \frac{cell\ viability\ in\ sample\ with\ Hygromycin}{cell\ viability\ in\ sample\ without\ Hygromycin}$$

The virus dilution that gave 50 % Hygromycin-resistant cells (HR-50) was selected as a starting point for stable transfection of cell lines with Cas9 nuclease.

2.2.2.3 Stable Cas9-transfection of cell lines

H2030 or H1975 cells were resuspended to a density of 1×10^5 cells per mL in RPMI 1640 medium supplemented with 10 % FBS and 5 µg/mL Polybrene. One mL of the cell suspension was aliquoted per well in a 12 well plate. Subsequently, two wells each were transduced with increasing amounts of virus corresponding to 1 × HR-50, 2 × HR-50, 4 × HR-50, 8 × HR-50, and 16 × HR-50. The suspension was carefully mixed by pipetting up-and-down and incubated at 37 °C. After 24 hours, the medium was replaced with fresh RPMI 1640 medium supplemented with 10 % FBS without Polybrene. Seventy-two hours after transduction, selection with Hygromycin was started using 400 µg/mL Hygromycin for H2030 and H1975 cells. The cells were grown for two weeks. During that period, the cells were passaged into cell culture flasks if needed in order to not discard any cells. Following this, the cells were counted using a Countess device. The sample transduced with the highest amount of virus, which yielded the highest number of living cells was chosen for further selection and split into three samples. For this, the cells were centrifuged at 400 × g for 5 minutes in a Heraeus Megafuge 16 centrifuge, with subsequent removal of the supernatant. The cell pellet was resuspended in 45 mL RPMI 1640 medium supplemented with 10 % FBS, split into

three T75 cell culture flasks and incubated at 37 °C. The three individual cell culture flasks containing were supplemented with 1x, 2x, and 3x the initial concentration of Hygromycin. All cells were incubated at 37 °C for one week, harvested and counted as described above. The sample that survived the highest Hygromycin concentration was expanded and used for the generation of frozen stocks.

2.2.2.4 Cas9 activity validation

The activity of Cas9 nuclease in stably Cas9 transduced cell lines was assessed using a CRISPRtest Functional Cas9 Activity Kit according to the manufacturer's instructions. H2030_Cas9 or H1975_Cas9 cells (1.3×10^6) were suspended in 13 mL of RPMI 1640 medium supplemented with 10 % FBS. For both cell lines, non-transduced parental cells were equally processed. The medium was supplemented with 13 µL CRISPRtest transduction reagent. For each cell line, 1 mL of cell suspension was aliquoted per well in a 12-well plate. Four wells each were transduced with 0 µL, 3 µL, 10 µL and 30 µL of CRISPRtest virus and incubated for 24 hours at 37 °C.

Twenty-four hours post-transduction, the cell growth medium was replaced with fresh RPMI 1640 medium supplemented with 10 % FBS, and the cells were incubated for additional 48 hours at 37 °C and 5 % CO_2. Three days post-transduction, the growth medium was aspirated, and the cells were washed with 0.5 mL PBS. The cells were detached from the surface by incubation with 0.3 mL Trypsin-EDTA for 5 minutes at 37 °C and 5 % CO_2, afterwards suspended in 2 mL RPMI 1640 medium supplemented with 10 % FBS, and transferred into 2 mL Safe-Lock microcentrifuge tubes. One mL of the cell suspension was transferred into a new 12-well plate and incubated at 37 °C and 5 % CO_2. The other half was centrifuged at 400 × g for 5 minutes in a Heraeus Megafuge 16 centrifuge. The supernatant was discarded, the cells were suspended in 2 mL PBS and transferred into a 5 mL Polystyrene Round Bottom FACS Tube. Samples, each containing 1×10^5 cells, were analyzed by flow cytometry using a FACSCanto II with the following settings (Table 14):

Table 14 | Settings for flow cytometric analysis of GFP and RFP

Channel	Excitation wavelength	Emission wavelength
GFP	488 nm	530 nm
RFP	561 nm	600 nm

For each cell line, the virus dilution where the percentage of RFP was closest to 10 % was chosen for further steps. All other samples were discarded. The cells were grown for additional seven days and kept in the log phase by splitting when microscopic appearance indicated 80 % confluency, following the steps as described in section 2.2.1.1. At day 10, the cells were analyzed by flow cytometry using the same settings as on day 3, and the GFP to RFP ratio was calculated. The Cas9 activity in the stable Cas9 transduced cells was calculated based on the following formula:

$$\%\ Knockout = 1 - \frac{(GFP{:}\,RFP)_{Cas9\ transduced\ cells}}{(GFP{:}\,RFP)_{Parental\ cells}}$$

2.2.2.5 Digital PCR

An analysis of the *KRAS* mutation G12C in H2030_Cas9 and H2030_WT cells was performed by digital PCR using the QuantStudio 3D digital PCR System, following the manufacturer's instructions. The *KRAS* G12C mutation was determined with a validated TaqMan Genotyping Assay (Table 6). Cells (4×10^6) were grown in a T175 cell culture flask. The cells were washed with PBS and detached by incubation with 4 mL Trypsin at 37 °C and 5 % CO_2 for 5 minutes. The cells were centrifuged at 1500 × g for 5 minutes in a Heraeus Megafuge 16 centrifuge and resuspended in 4 mL of cold PBS. Genomic DNA was extracted using a Blood and Cell Culture DNA Midi Kit according to the instructions of the manufacturer in a UV-treated clean PCR workstation at room temperature. The DNA was diluted to a concentration of 3 ng/µL in cold PBS. QS3D Master Mix v2 (7.25 µL), 0.71 µL TaqMan assay and 6.54 µL DNA, equivalent to 20 ng, were mixed and loaded onto the QuantStudio 3D Digital PCR 20K Chip. Thermo cycling was performed in a Thermocycler ProFlex 2x Flat PCR System according to the following settings (Table 15):

Table 15 | Thermal cycle settings for digital PCR

Step	Temperature	Duration
1	96 °C	10 min
2	60 °C	2 min
Repeat 44 x		
3	98 °C	0.5 min
4	60 °C	2 min
5	10 °C	hold

Following amplification, fluorescence signals were detected by the QuantStudio 3D Digital PCR Instrument. Absolute numbers of mutant and wildtype *KRAS* allele copies and the resulting Variant Allele Frequency (VAF) were determined using the QuantStudio Analysis Suite software.

2.2.3 Pooled CRISPR/Cas9 gene knockout

2.2.3.1 Production of lentiviral particles

A number of 12.5×10^6 HEK-293T/17 cells was seeded per flask in a total of six T175 flasks with 25 mL DMEM per flask and incubated for 24 hours at 37 °C. The following day, 360 µL Lentiviral Packaging Plasmid Mix, 36 µL Plasmid sgRNA library (1 µg/µL) and 360 µL PLUS Reagent were mixed with 7 mL DMEM without serum or antibiotics in a 50 mL Falcon Conical Centrifugation Tube. The mixture was incubated at room temperature for 15 minutes.

In parallel, 600 µL Lipofectamine was diluted in 7 mL DMEM in a 50 mL Falcon Conical Centrifugation Tube. The diluted Lipofectamine was added to the DNA / PLUS Reagent complex, vortexed thoroughly and incubated at room temperature for 15 minutes. The resulting DNA / PLUS Reagent / Lipofectamine complex was added to each T175 flask (2 mL / flask), and cells were incubated for 24 hours at 37 °C and 5 % CO_2.

Twenty-four hours post-transfection, the medium was replaced with 25 mL of fresh DMEM supplemented with 10 % FBS, DNase I (1 U/mL), $MgCl_2$ (5 mM), and 20mM HEPES, pH 7.4, followed by incubation for 24 hours at 37 °C and 5 % CO_2. Forty-eight hours post-transfection, the virus-containing medium was aspirated and filtered into 50 mL Falcon Conical Centrifugation Tubes through a 0.22 µm Millex-GP syringe filter unit to remove debris and floating cells. The lentivirus-containing supernatant was transferred into 5 mL aliquots in 15 mL Falcon Conical Centrifugation Tubes and stored at -80 °C.

2.2.3.2 Titration of lentiviral particles

To assess the viral transduction efficacy, 2×10^6 cells were seeded in a T175 flask. After 24 hours, the cells were transduced with 40 µL, 80 µL, 160 µL, 320 µL, 640 µL, 1280 µL, and 2560 µL of pooled lentiviral particles. In transduced cells, RFP is

expressed from the same promoter as the sgRNA, allowing for quantification of the fraction of RFP(+) cells by flow cytometry.

2.2.3.3 Pooled gene knockdown

H2030_Cas9 or H1975_Cas9 cells (45×10^6 each) were trypsinized and resuspended to a density of 80×10^5 cells per mL in RPMI 1640 medium supplemented with 10 % FBS and 5 µg/mL Polybrene. The cell suspension was aliquoted in 23 (H2030_Cas9) or 32 (H1975_Cas9) T175 flasks and transduced with 80 µL (H2030_Cas9) or 340 µL (H1975_Cas9) of lentiviral CRISPR library particles (section 2.2.3.1) per flask, respectively. After 72 hours, the cells of one representative flask were harvested and the fraction of transduced / RFP(+) cells was assessed by flow cytometry. Afterwards, the cells of all flasks were treated with Puromycin to deplete non-transduced cells according to the previously established IC-90 for each cell line. The selected cells were expanded for further 72 hours and split into a control / no-drug arm and a drug treatment arm, each consisting of 55×10^6 cells in 32 T175 flasks according to the protocol in section 2.2.1.1. The cells were treated with the drug for 14 days in the treatment arm and the equivalent of DMSO in the control arm. Upon splitting of the cells, the number of cells in each arm was kept not lower than 55×10^6 cells. The cell culture medium was replaced with fresh RPMI 1640 medium supplemented with 10 % FBS and the appropriate concentration of the drug.

2.2.4 Extraction of genomic DNA and amplification of sgRNAs

2.2.4.1 Isolation of genomic DNA

After 14 days of incubation the cells were harvested. All steps were executed in batches of eight cell-culture flaks. First, the medium of each cell culture flasks was discarded, and the cells were washed with 10 mL of DPBS. The cells were dissociated from the culture flask surface by incubation with 4 mL of Trypsin-EDTA for 5 minutes at 37 °C and 5 % CO_2 in a BBD 6220 CO_2-Incubator. Incubation was stopped once all cells were detached by addition of 4 mL cell culture medium supplemented with FBS. The detached cells from each screening arm were pooled, transferred into three 50 mL Falcon Conical Centrifugation Tubes and centrifuged at 400 × g for 5 minutes in a Heraeus Megafuge 16 centrifuge. Following centrifugation, the supernatant was aspirated and discarded, and the cells were resuspended in 4 mL of DBPS per tube. The

cell suspensions were pooled in a 15 mL Falcon Conical Centrifugation Tube and centrifuged at 400 × g for 5 minutes. Afterwards, the supernatant was aspirated and discarded. The cells were lysed using 0.025 % Sodium dodecyl sulfate (SDS) in 5 mL of Buffer P1 with RNAse A per tube and incubated at RT for 5 minutes.

Using a FS-300N Ultrasonic Homogenizer Sonicator, the DNA was sheared for 3 minutes with settings of 60 % power and four pulses of 10 seconds until the suspension became less viscous. Afterwards, the proteins were digested for 15 minutes at RT using 10 µL of Proteinase K. Next, 5 mL of Phenol:Chloroform:Isoamyl (25:24:1) Alcohol solution was added to each tube, and the suspension was mixed by inverting the tube 10 times. The solution was centrifuged at 8,000 × g for 1 hour at 20 °C in an Avanti J-30I centrifuge. Following centrifugation, 4 mL of the upper phase was transferred into a new 15 mL Falcon Conical Centrifugation Tube. The DNA was precipitated by adding 0.5 mL of 3M sodium acetate and 4 mL of isopropanol. After mixing the solution by inverting the tube 10 times, the solution was incubated overnight at RT. Afterwards, the precipitated DNA was centrifuged at 8,000 × g for 30 minutes at 20 °C in an Avanti J-30I centrifuge. Following centrifugation, the supernatant was discarded and 10 mL 70 % ethanol was added to the tube. The solution was centrifuged at 8,000 × g for 5 minutes at 20 °C. The latter step was repeated once. The supernatant was discarded, and the pellet was air-dried until no remaining ethanol was visible. Finally, the pellet was dissolved in 500 µL pre-heated (80 °C) Ambion ddH_2O and transferred to a 1.5 mL Safe-Lock microcentrifuge tube. After incubation at 80 °C for 30 minutes, the DNA concentration was determined by a Nanodrop UV-spectrophotometer.

2.2.4.2 Amplification of sgRNAs

sgRNAs were amplified from genomic DNA with a 2-step nested PCR. All reactions were performed in a GeneTouch Thermal Cycler. The components of the first round PCR were mixed according to the following protocol (Table 16) and split into eight 0.2 mL PCR microtubes with 50 µl reaction volume each.

Table 16 | First-round PCR master mix for sgRNA amplification

Volume	Component
X µL	200 µg of genomic DNA
12 µL	First-Round_FwdU6-1 primer(10 µM)
12 µL	First-Round_Rev1primer (10 µM)
8 µL	50x dNTP Mix (10 mM each)
40 µL	10x Titanium Taq Buffer
8 µL	50x Titanium Taq Polymerase
320 – X	Deionized water

The first-round PCR was performed with the following settings (Table 17):

Table 17 | Thermal cycle settings for first-round amplification of sgRNAs

Step	Temperature	Duration
1	94 °C	3 min
2	94 °C	30 sec
3	65 °C	10 sec
4	72 °C	20 sec
Go to 2, repeat 15x		
5	68 °C	2 min
6	4°C	hold

For the second-round PCR, the 8 × 50 µL first-round PCR reactions were combined, and 200 µL of the first-round PCR product was used for the second round PCR. The remaining first-round PCR product was stored at -80 °C as a backup. The components were mixed according to the following protocol (Table 18) and split into six 0.2 mL PCR microtubes with 50 µl reaction volume each.

Table 18 | Second-round PCR master mix for sgRNA amplification

Volume	Component
200 µL	First-round PCR product
10 µL	Sec-Round_P5-NFG16 primer (10 µM)
10 µL	Sec-Round_P7-NRG1_6IND-A – D primer (10 µM)
4 µL	50x dNTP Mix (10 mM each)
20 µL	10x Titanium Taq Buffer
4 µL	50x Titanium Taq Polymerase
52 µL	Deionized water

The second-round PCR was performed with the following settings (Table 19):

Table 19 | Thermal cycle settings for second-round amplification of sgRNAs

Step	Temperature	Duration
1	94 °C	3 min
2	94 °C	30 sec
3	65 °C	10 sec
4	72 °C	20 sec
Go to 2, repeat 17x		
5	68 °C	2 min
6	4°C	hold

The samples were combined and purified with the QIAquick PCR purification kit according to the manufacturer's protocol, followed by a preparative agarose gel electrophoresis (section 2.2.6.1) and subsequent extraction using a QIAquick gel purification kit according to the manufacturer's protocol. The DNA concentration was measured using a Qubit Fluorometric Quantification Device.

2.2.4.3 High-throughput deep sequencing of sgRNAs

Sequencing of sgRNAs was performed on a HiSeq 2000 instrument by the High Throughput Sequencing Unit in the Genomics and Proteomics Core Facility of the German Cancer Research Center (DKFZ).

Pooled sgRNAs (amplified and purified as described in section 2.2.4.2) were adjusted to a concentration of 10 nM in Buffer EB in a quantity of 40 µL. The mass concentration was transferred into the molar concentration using an excel form that was kindly provided by the core facility. All samples were stored on ice until the submission to the Core Facility.

2.2.4.4 Assessment of sgRNA readcounts

Sequencing results were provided in the *.fastq file format. To access the sgRNA readcounts, the files were opened in an editor and divided into ten smaller files. Each resulting *.fastq file was individually uploaded to the *CRISPRAnalyzeR* platform developed by Jan Winter at the Division of Signaling and Functional Genomics at the DKFZ (http://crispr-analyzer.dkfz.de). The reads were mapped to corresponding genes using the Bowtie2 tool, resulting in ten readcount files per sample. The readcount files corresponding to the same sample were merged, and the reads were normalized to a total readcount of 100 million reads per sample. The merged readcount files were re-uploaded to CRISPRAnalyzeR, and statistically significant hits were identified according to the platform's instructions.

2.2.5 siRNA-mediated transient gene silencing

The siRNA transfection was performed under the Maxisafe 2020 Class II Biological Safety Cabinets. The lyophilized siRNAs were re-suspended in 100 µL of nuclease-free water to obtain 10 µM stock solutions. Target sequences of the siRNAs used are listed in Table 11.

Depending on the cell line and the desired readout different cell counts (5×10^5, 10×10^5 or 20×10^5) were plated per well in 1 mL RPMI 1640 medium supplemented with 10 % FBS, in Falcon Polystyrene 12-well plates. After 24 hours, the culture medium was aspirated and replaced with 950 µL fresh RPMI 1640 Medium supplemented with 10 % FBS per well.

Lipofectamine 2000 Transfection reagent was diluted 1:12 in Gibco Opti-MEM I Reduced Serum Medium. In parallel, the siRNA was diluted in Gibco Opti-MEM I Reduced Serum Medium to the desired concentration and mixed in a ratio of 1:1 with the diluted Lipofectamine Transfection reagent. After 10 minutes of incubation at room temperature, 50 µL of the mixture were added per well. The cells were incubated at 37 °C (5 % CO2) in a BBD 6220 CO2-Incubator until further processing, depending on the desired read-out.

2.2.6 Nucleic acid analysis

2.2.6.1 DNA agarose gel electrophoresis

Agarose NEEO Ultra-Quality (3.5 g) was dissolved in 100 mL 1x Tris-acetate-EDTA (TAE) buffer by boiling in a microwave equivalent to a 3.5 % gel. The solution was cooled down at room temperature followed by supplementation with one drop of 1 % ethidium bromide solution, and poured into a gel electrophoresis tray with a well comb. After solidification, the gel tray was submerged in an Electrophoresis chamber with 1x TAE buffer. Prior to sample loading, 1 µL of 6x Gel Loading Dye was supplemented to 5 µL of sample. The samples were then loaded in volumes of 5 µL per well. Gel electrophoresis was carried out at 200 V for approximately 45 minutes. Fragments were visualized by UV-light using the Quantum-ST4 gel documentation instrument. For preparative gel extraction of DNA, the desired fragment bands were cut from the gel with a scalpel.

2.2.6.2 RNA extraction

One to 3×10^6 cells were washed in DPBS and trypsinized as described in section 2.2.1.1. Total RNA isolation was performed using the RNeasy Mini Kit according to the manufacturer's instructions. All centrifugation steps were performed in a Heraeus Fresco Microcentrifuge. The total RNA was eluted in 30 µL of nuclease-free water.

2.2.6.3 RNA and DNA quantification

Nucleic acid concentrations were determined spectrophotometrically using a NanoDrop 1000 Spectrophotometer. After a blank reading with the diluent, the absorption at 260 nm was measured (1 × OD260 = 40 µg RNA or 50 µg DNA). The purity of nucleic acids was assessed by calculating the 260 nm /280 nm absorbance ratio. Ratios of 1.8-2.0 indicated the absence of protein contaminations.

2.2.6.4 cDNA synthesis

cDNA was synthesized from 100 ng total RNA diluted in 11 µL nuclease-free water using the RevertAid H Minus First Strand cDNA Synthesis Kit. Random hexamer primers were used to prime the cDNA synthesis. The synthesis was completed

following the manufacturer's First Strand cDNA Synthesis protocol. The cDNA was diluted to a concentration of 1 ng/µL.

2.2.6.5 Real-time RT-qPCR

Primer pairs for all real-time RT-qPCR assays were designed using the Universal ProbeLibrary Assay Design Center. A list of primer sequences is provided in Table 8. Real-time RT-qPCR amplification was performed in the LightCycler 480 II Real Time PCR System using LightCycler 480 384 Multiwell Plates. For each reaction, 5 µL of the previously diluted cDNA (corresponding to 10 ng) were mixed with 5.5 µL enzyme master mix, 0.11 µL For/Rev primer, 0.11 µL Universal Probe Library, and 0.28 µL nuclease-free water. Each real-time RT-qPCR run contained a negative control (nuclease-free water) and a positive control (Universal Human Reference RNA from Agilent Technologies). All reactions were set up in triplicates. Following sample loading, the LightCycler 480 384 Multiwell Plate was sealed using the LightCycler 480 Sealing Foil, followed by centrifugation at 1,000 × g for 15 seconds. Afterwards, the plate was placed into the LightCycler 480 II real-time PCR System and the run was started according to the following protocol (Table 20):

Table 20 | Thermal cycle settings for real-time RT-qPCR

Step	Temperature	Duration
1	95 °C	2 min
2	95 °C	10 sec
3	60 °C	25 sec
Go to 2, repeat 39x		
4	40 °C	1 min
5	4°C	hold

Data analysis was done with the LightCycler®480 Software 1.5.0 using the second derivative method for calculation of the sample Ct-values. The expression level of the gene of interest was determined relatively to the expression level of the housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*). First, ΔCt-values were calculated for each gene of interest according to the following formula:

$$\Delta Ct = Ct_{Gene\ of\ Interest} - Ct_{GAPDH}$$

Afterwards, the relative expression ratio (R) of the gene of interest was calculated with the following equation:

$$R = 2^{-(\Delta Ct_{Gene\ of\ Interest} - \Delta Ct_{Neg.Control\ Gene}} = 2^{-\Delta\Delta Ct}$$

2.2.6.6 Microarray analysis

Cells were harvested and the mRNA was isolated using the RNeasy Mini Kit (Qiagen, USA). The RNA concentration was determined with a Nanodrop spectrophotometer (Peqlab, Inc.) by measuring absorption at 260 nm. RNA samples were diluted to a concentration of 10 ng/µL to 50 ng/µL and submitted to the Microarray Unit of the Genomics and Proteomics Core Facility of the German Cancer Research Center in a quantity of > 150 ng in a volume of > 10 µL. The integrity of the samples was confirmed prior to submission using an RNA 6000 Nano Chip following the manufacturer's protocol. The samples were analyzed using a HumanHT-12 v4 Expression BeadChip Kit. The mean probe values were normalized to the non-target control sample and deregulated genes were called with fold changes of > 1.5 for upregulated and < 0.66 for downregulated genes. Shortlisted genes were submitted to the *Database for Annotation, Visualization and Integrated Discovery* (DAVID), and annotation cluster settings were adjusted to include *Biocarta pathways*, *Kyoto Encyclopedia of Genes and Genomes pathways* (KEGG) and *Gene Ontology annotations* (GO) with an adjusted p-value < 0.05.

2.2.7 Biological assays

All assays were performed at least in two biological replicates unless otherwise stated.

2.2.7.1 Cell viability assay

The CellTiter-Blue Cell Viability Assay measures the metabolic activity of cells, which is a proportional indicator for cell viability, by the conversion of resazurin to resorufin. The latter is highly fluorescent (Excitation maximum: 579 nM / Emission maximum: 584 nM) and can be used for the quantification of cell viability. Non-viable cells rapidly lose metabolic capacity, do not reduce the indicator dye, and do not generate a fluorescent signal. Thus, the CTB signal is directly proportional to metabolic activity and can be utilized as a marker for cell viability. For the assay, cells were plated at an appropriate density (depending on the cell type) in a 12-well plate with 1 mL RPMI 1640 medium supplemented with 10 % FBS. The cells were cultured at 37 °C and 5 % CO_2 in the BBD 6220 CO_2-Incubator. After 24 hours, the cell culture medium was replaced and supplemented with the desired drug concentration. A list of drugs used in this study is provided in Table 4. The cells were incubated for 72 hours until

performing the cell viability measurement. For this, the cell culture medium was discarded, and the cells were washed with 500 µL DPBS. Each well was supplemented with 1 mL of fresh RPMI cell culture medium and 200 µL of CellTiter-Blue Cell Viability Assay reagent, following incubation at 37 °C and 5 % CO_2 for 60 minutes. Afterwards, the fluorescence signal was measured using the Infinite M200 multiplate reader at a wavelength of 590 nm and a gain of 70. The background signal (medium plus assay reagent without cells) was deducted from the readout values.

2.2.7.2 Combined transient gene silencing and drug treatment assay

To study the effects of combined drug treatment and gene knockdown, the target gene expression was transiently silenced using siRNA, as described in section 2.2.5. Afterwards, the cells were treated with the drug of interest, as described in section 2.2.7.1. Subsequent steps depended on the readout (RT-qPCR: 2.2.6.2, Microarray: 2.2.6.6, Cell viability assay: 2.2.7.1, Analysis of apoptosis by flow cytometry: 2.2.7.3).

2.2.7.3 Analysis of apoptosis by flow cytometry

After treatment, the cells were washed twice with cold DPBS and resuspended in 1 × Binding Buffer at a concentration of 1×10^6 cells / mL. One hundred µL of the solution were transferred into a 5mL Polystyrene Round Bottom FACS Tube. Five µL of Annexin V –Phycoerythrin (PE) and 5 µL of 7-Aminoactinomycin D (7-AAD) dye were added. In normal viable cells, PS is located on the inner surface of the cell membrane. However, in apoptotic cells, PS is flipped from the inner to the outer leaflet of the plasma membrane. Annexin V, labelled with PE, can identify apoptotic cells by binding to PS exposed on the outer leaflet. 7-AAD cannot enter an intact or early apoptotic cell, as it is unable to penetrate the cell membrane. However, when a cell's membrane is damaged due to cell death, the dye can enter the cell and stain its DNA, serving as a marker for cell death. After staining with both Annexin V - PE and 7-AAD, viable cells will be non-fluorescent, cells actively undergoing apoptosis will stain positive for Annexin V - PE, late apoptotic cells will stain positive for both PE and 7-AAD and dead cells will stain positive for solely 7-AAD. The tube was gently vortexed, and the cells were incubated on ice for 15 minutes in the dark. Four hundred µL of 1 × Binding Buffer were added to each tube, and the samples were analyzed using a Canto II instrument with the following settings (Table 21):

Table 21 | Settings for flow cytometry analysis

Dye	Excitation/Emission	Excitation laser	Fluorescence channel
Annexin V –Phycoerythrin (PE)	496 / 576 nm	488 nm	Yellow
Aminoactinomycin D (7-AAD)	550 / 650 nm	488 nm	Red

3 Results

3.1 Establishment of a pooled lentiviral CRISPR/Cas9 screen

This project aimed to identify novel essential viability genes and synthetically lethal genes with epigenetic drugs by pooled lentiviral CRISPR/Cas9 screenings. The first milestone of this work was to establish a platform for pooled screenings. This process involved the generation of stable Cas9 nuclease-expressing cells with subsequent testing of the nuclease activity, and further the validation of an unchanged *KRAS* mutation status after the integration of Cas9. This part of the work was concluded by the determination of appropriate drug concentrations in NSCLC cell lines, which would be used for the screenings that were conducted in the next phase of the study.

3.1.1 Generation of stable Cas9 nuclease expressing cells

For CRISPR screens, cells need to express Cas9 nuclease in conjunction with sgRNAs. The CRISPR sgRNA library used in this study (provided by Cellecta, Inc.) expresses solely sgRNAs, so it was necessary to engineer target cells that express the appropriate Cas9 variant for the screen. First, the sensitivity of the cells to Hygromycin was established, because Hygromycin would be used as a selection drug for Cas9-transduced cells. Next, lentiviral Cas9 particles were produced, and cells were transduced with the particles. Finally, the transduced cells were selected using the previously established concentration of Hygromycin.

3.1.1.1 Titration of Hygromycin for stable Cas9 transduction of cell lines

To utilize Hygromycin as a selection marker for the generation of stably transfected Cas9 expressing cells, the concentration of Hygromycin, which would eliminate > 90 % of H2030 cells (IC-90) was established. For this, 10,000 cells/well were seeded in a 12-well plate and treated with increasing concentrations of Hygromycin. After 72 hours, cell viability was assessed with CellTiter Blue (Promega) staining (Figure 6).

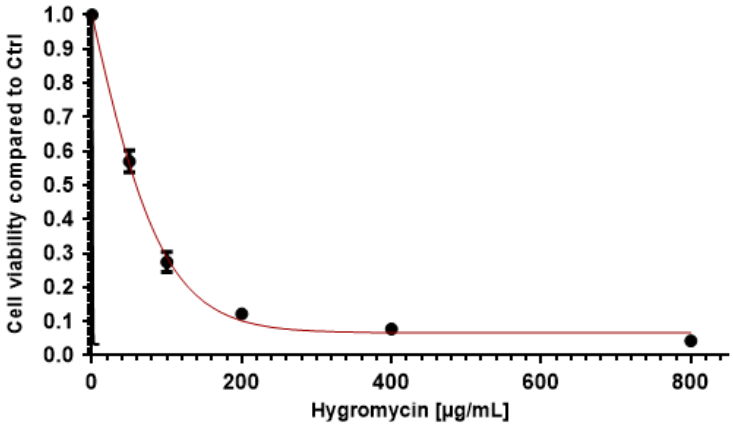


Figure 6 | Titration of Hygromycin in H2030 cells

To determine the concentration of Hygromycin that would result in 90 % or more non-viable H2030 cells, the cells were treated with increasing concentrations of Hygromycin. After 72 hours, cell viability was assessed with CellTiter Blue staining. 400 µg/mL Hygromycin resulted in 85 % of dead cells compared to the untreated control, and 800 µg/mL resulted in 88 % of dead cells.

Using 400 µg/mL or 800 µg/mL Hygromycin, 85 % or 88 % of the treated cells were dead compared to the untreated control, respectively. Since doubling the Hygromycin concentration from 400 to 800 µg/mL added only a minor effect to cell-death, 400 µg/mL was chosen for subsequent selection steps within the process of generating stable Cas9 expressing H2030 cells.

In the same fashion as for H2030 cells, the IC-90 of Hygromycin was determined for H1975 cells (Figure 7).

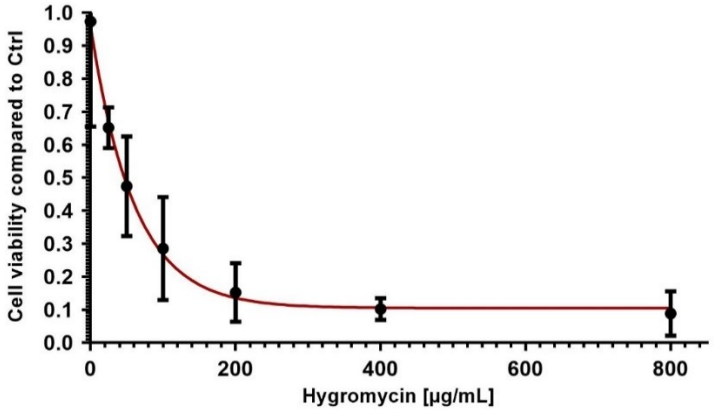


Figure 7 | Titration of Hygromycin in H1975 cells

To determine the concentration of Hygromycin that would result in 90 % or more non-viable H1975 cells, the cells were treated with increasing concentrations of Hygromycin. After 72 hours, cell viability was assessed with CellTiter Blue staining. 400 µg/mL Hygromycin resulted in 90 % of dead cells compared to the untreated control.

Here, by using 400 µg/mL 90 % of the treated cells were dead compared to the untreated control. Consequently, also here 400 µg/mL was chosen to be used in subsequent steps within the process of generating stable Cas9 expressing H1975 cells.

3.1.1.2 Stable transduction of cell lines with Cas9 nuclease

To establish Cas9 expressing cell lines, 1.35 mL of lentiviral Cas9 particles capable of transducing the Cas9 gene into target cells were produced. After having produced lentiviral Cas9 particles, the viral titer was established in different cell lines. In H2030 cells, 3.3 µL of a 1:10 dilution of Cas9 lentiviral stock resulted in 50 % Hygromycin-resistant cells (HR-50), in H1975 cells, HR-50 was observed with 10 µL of the diluted virus stock (data not shown). In the next step, the cells were stably transduced with Cas9 nuclease. The cells were grown for two weeks and counted afterwards (Table 22).

Table 22 | Number of cells after two weeks of Hygromycin selection

	Number of living cells	
Amount of Cas9 Virus	H2030_Cas9	H1975_Cas9
1 x HR-50	3×10^6	4×10^6
2 x HR-50	6×10^6	8×10^6
4 x HR-50	1×10^7	3×10^7
8 x HR-50	1×10^7	2×10^7
16 x HR-50	8×10^6	9×10^6

As depicted in Table 22, in H2030_Cas9 cells, both 4 x HR-50 and 8 x HR-50 yielded the highest number of living cells. Since 8 x HR-50 exhibited a higher number of Cas9 integrations than that of 4 x HR-50, the 8 x HR-50 sample was chosen for further selection steps. In H1975_Cas9 cells, transduction with 4 x HR-50 yielded the highest number of living cells and was chosen for further selection procedures.

The selected sample of H2030_Cas9 cells was split into three samples and supplemented with 600 µg/mL, 1200 µg/mL, or 1800 µg/mL Hygromycin. The selected sample of H1975_Cas9 cells was split into three samples and supplemented with 400 µg/mL, 800 µg/mL, or 1600 µg/mL Hygromycin. After one week, the H2030_Cas9 and H1975_Cas9 cells that were treated with 600 µg/mL and 400 µg/mL, respectively, were chosen for CRISPR/Cas9 screening experiments and expanded. H2030_Cas9 and H1975_Cas9 cells that received higher concentrations of Hygromycin did not survive.

3.1.2 Cas9 activity and cell line integrity after integration of Cas9

The nuclease activity of Cas9 is proportional to the number of *Cas9* integrations into the genome of a cell, with more integrations leading to a higher nuclease activity. A high number of integrations and thus a high Cas9 activity is favorable for efficient CRISPR–based knockdown experiments; consequently, the nuclease activity needed to be confirmed before starting knockdown experiments, ensuring an efficient gene knockout.

After having stably transduced H2030 and H1975 cells with Cas9 nuclease, the Cas9 activity in these cell lines was assessed using two pre-mixed lentiviral-packaged vectors, the first containing a *green fluorescent protein* (GFP) marker and sgRNA sequence targeting an essential gene (human *PCNA*) and the second containing a *red fluorescent protein* (RFP) marker and a non-targeting sgRNA (Figure 8).

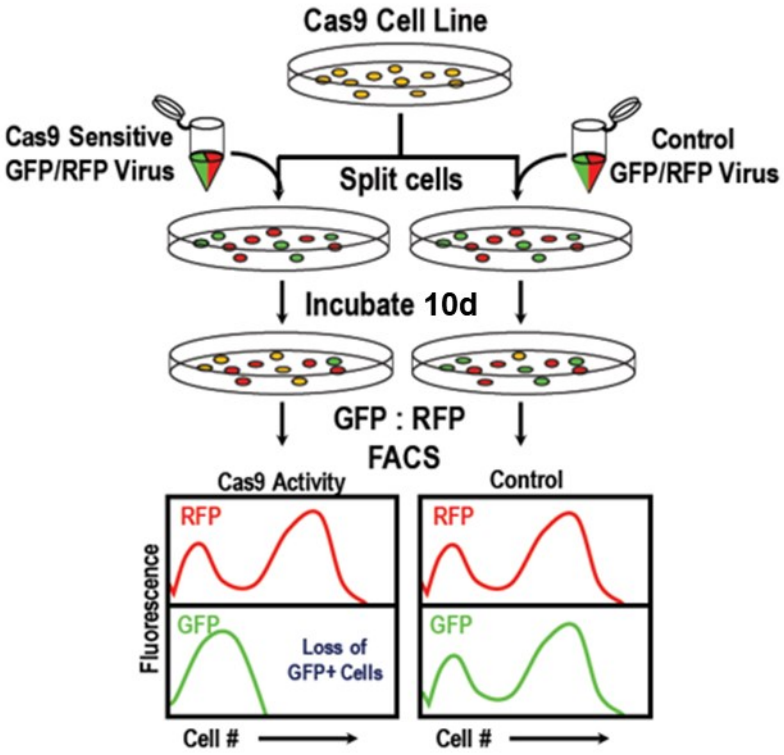


Figure 8 | Overview of the CRISPRtest workflow

The Cas9 cell line and parental cells were transduced with the CRISPRtest virus. 10 days after transduction, a portion of the cells was analyzed by flow cytometry to determine GFP:RFP ratio. These ratios were then used to calculate the knockout percentage, which is an indicator of Cas9 activity. (Illustration is courtesy of and has been adapted from Cellecta, Inc.)

The transduction of CRISPRtest virus into cells would result in an initial ratio of GFP positive cells to RFP positive cells, which is analyzed by flow cytometry 3 days after transduction. Upon cell growth in the presence of active Cas9, the ratio of GFP positive

cells to RFP positive cells will decrease proportionally to Cas9 activity. However, in cells that do not express Cas9, the ratio would not be altered.

3.1.2.1 Cas9 activity in H2030_Cas9 cells

Three days post-transduction, an aliquot of the cells from each well was analyzed by flow cytometry, and the percentage of RFP positive but GFP negative cells was assessed. Ideally, 10 % of the cells would fulfil this requirement. The cells that were transduced with 3 µL of the virus showed 9.58 % (H2030_Cas9) and 12.0 % (H2030 WT) of RFP(+) / GFP(-) cells and were selected for continuation (Figure 9, top row, Q1). All other dilutions were discarded.

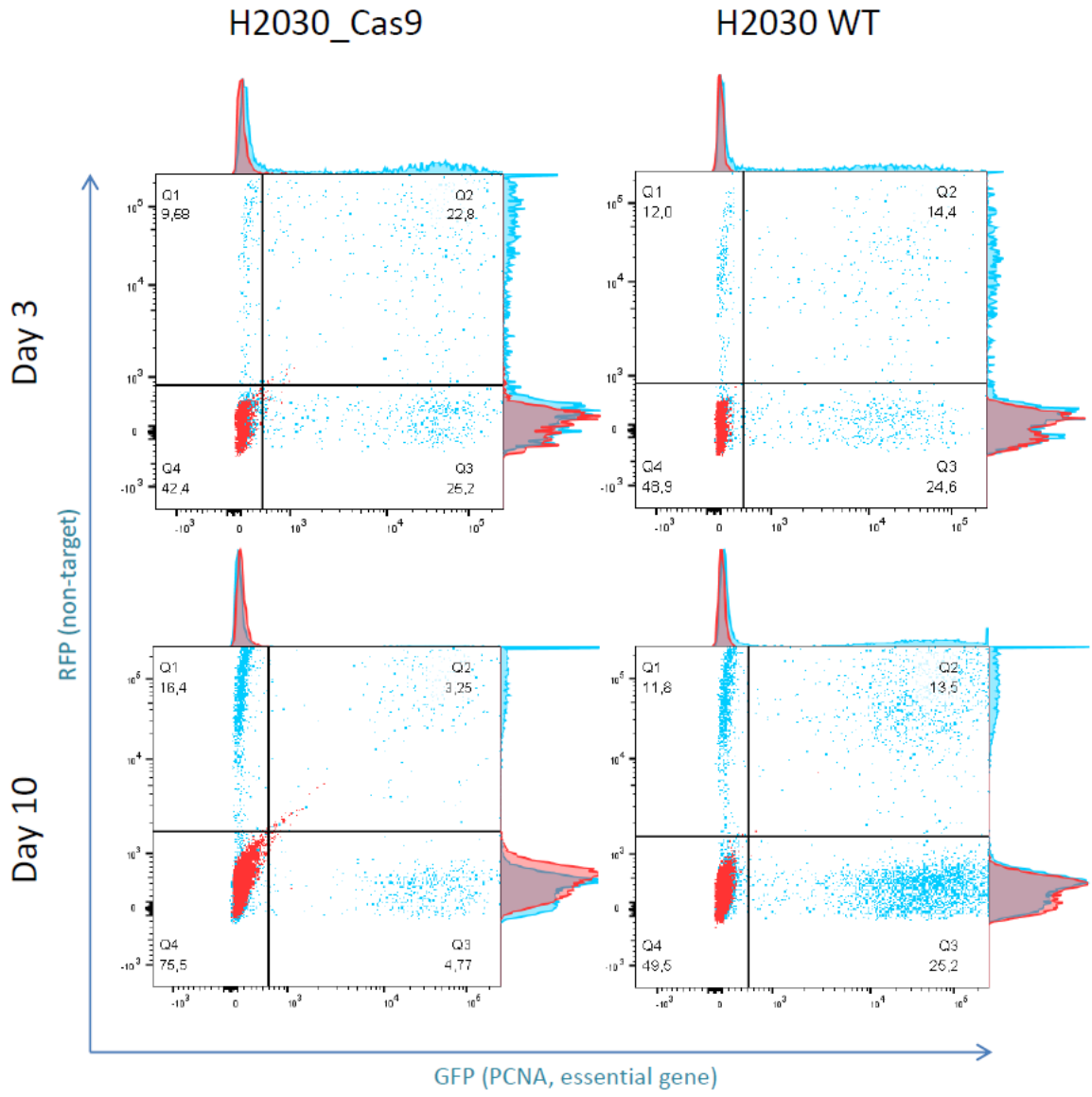


Figure 9 | Flow cytometry analysis of Cas9 nuclease activity in H2030_Cas9 and H2030 WT cells

The Cas9 activity in H2030 cells was assessed using two pre-mixed lentiviral-packaged vectors (Cellecta), the first containing a green fluorescent protein (GFP) marker and sgRNA sequence targeting an essential gene (human *PCNA*) and the second containing a red fluorescent protein (RFP) marker and a non-targeting sgRNA. RFP and GFP signal intensity of H2030_Cas9 and parental H2030 WT cells were analyzed at day 3 and day 10 post-transduction. The ratio of GFP to RFP signal intensity at day 10 was used to calculate the Cas9 activity in H2030_Cas9 cells, which was determined to be 86 %.

At day 10, the cells were analyzed by flow cytometry using the same settings as on day 3 (Figure 9, lower panel). The GFP to RFP ratio was calculated in both cell lines, and the Cas9 activity was assessed, the results are shown in Table 23.

Table 23 | Results of the flow cytometry analysis of Cas9 nuclease activity in H2030_Cas9 cells

Calculations	H2030 WT	H2030_Cas9
RFP positive cells (%)	11.8	16.4
GFP positive cells (%)	25.2	4.77
Ratio GFP to RFP	2.14	0.29
Knockout (%)	1-(0.29/2.14) = 0.86 = 86 %	

The calculated knockout value is the percentage of cells in the H2030_Cas9 cell line where the *PCNA* gene was knocked out (i.e., where CRISPR knockout was effective for both alleles). In H2030_Cas9 cells, the *PCNA* gene was knocked out in both alleles in 86 % of the cells.

3.1.2.2 Cas9 activity in H1975_Cas9 cells

H1975_Cas9 cells that were transduced with 30 µL of the virus presented 2.91 % (H1975 WT) or 3.05 % of RFP(+) / GFP(-) cells after three days (Figure 10, top row, Q1).

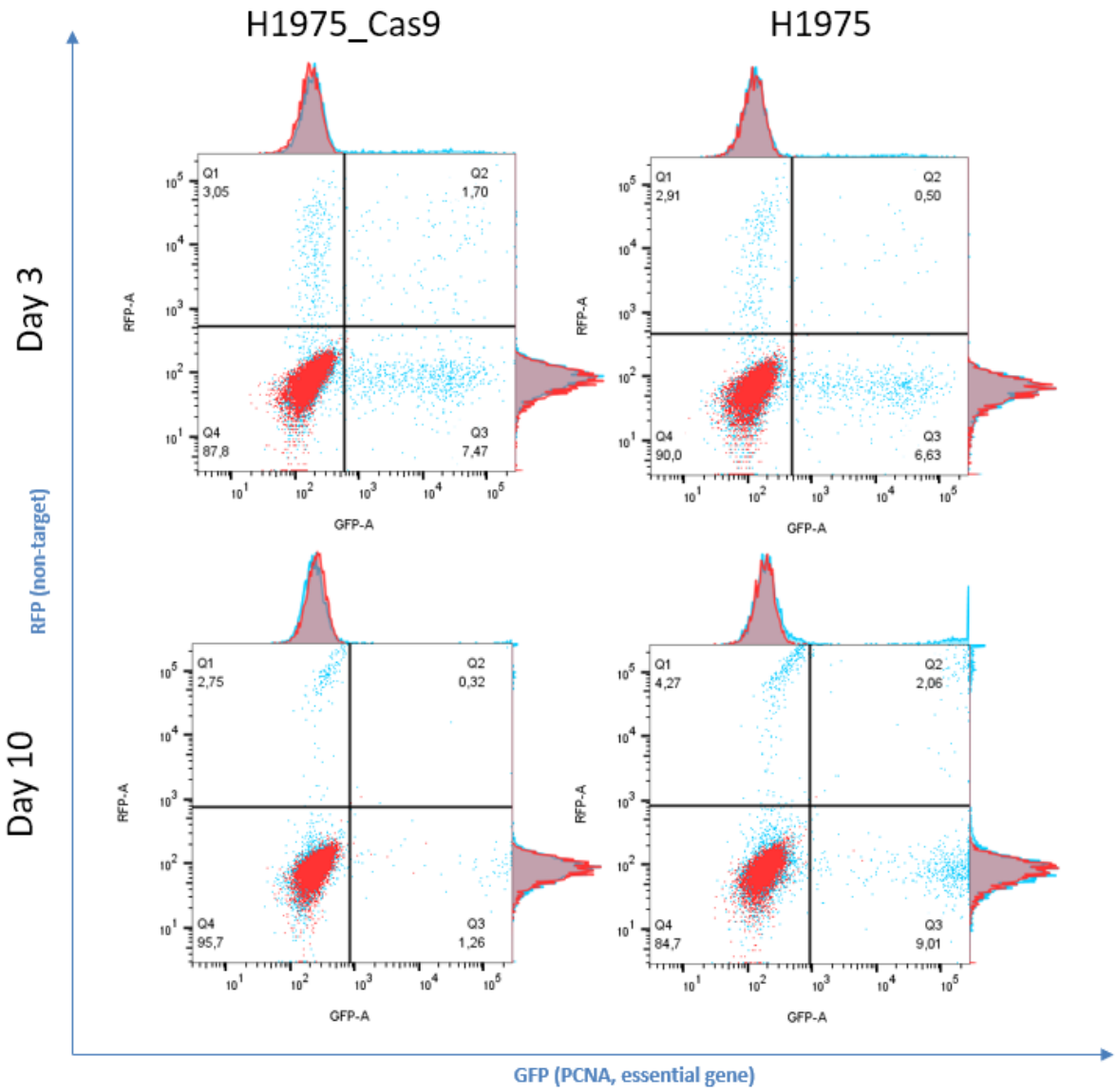


Figure 10 | Flow cytometry analysis of Cas9 nuclease activity in H1975_Cas9 and H1975 WT cells

The Cas9 activity in H1975 cells was assessed using two pre-mixed lentiviral-packaged vectors (Cellecta), the first containing a green fluorescent protein (GFP) marker and sgRNA sequence targeting an essential gene (human *PCNA*) and the second containing a red fluorescent protein (RFP) marker and a non-targeting sgRNA. RFP and GFP signal intensity of H2030_Cas9 and parental H2030 WT cells were analyzed at day 3 and day 10 post-transduction. The ratio of GFP to RFP signal intensity at day 10 was used to calculate the Cas9 activity in H2030_Cas9 cells, which was determined to be 78 %.

Although these numbers were not in the, according to the supplier, ideal range of around 10 %, they were comparable between the WT and the Cas9 variant of H1975 cells and therefore selected for further incubation. At day 10, the cells were analyzed by flow cytometry using the same settings as on day three (Figure 10, lower panel). The GFP to RFP ratio was calculated in both cell lines, and the Cas9 activity was assessed (Table 24).

Table 24 | Results of the flow cytometry analysis of Cas9 nuclease activity in H1975_Cas9 cells

Calculations	H1975 WT	H1975_Cas9
RFP positive cells (%)	4.27	2.75
GFP positive cells (%)	9.01	1.26
Ratio GFP to RFP	2.11	0.46
Knockout (%)	1-(0.46/2.11) = 0.78 = 78 %	

The *PCNA* gene was knocked out in both alleles in 78 % of the cells, being 8 % lower compared to H2030_Cas9 cells. This finding indicates fewer integrations of the Cas9 gene into this cell line compared to H2030_Cas9; however, a 78 % knockout rate would be considered as efficient editing and suitable for knockout experiments[142]. Taken together, the knockout efficiency in both cell lines proved functional Cas9 gene editing and suitability for pooled CRISPR/Cas9 knockout screening experiments.

3.1.2.3 *KRAS* G12C mutation status in Cas9 transduced H2030 cells

Since Cas9 randomly integrates into the genome of a cell, viral DNA coding for Cas9 may have integrated at the genomic locus of *KRAS* in H2030 cells, hereby compromising the mutation. To ensure that the integration of Cas9 did not alter the allele frequency of the *KRAS* mutation p.G12C, this frequency was determined by a digital PCR in H2030_Cas9 cells and compared to parental H2030 cells. Absolute numbers of mutant and wildtype *KRAS* allele copies and the resulting Variant Allele Frequency (VAF) were determined (Figure 11).

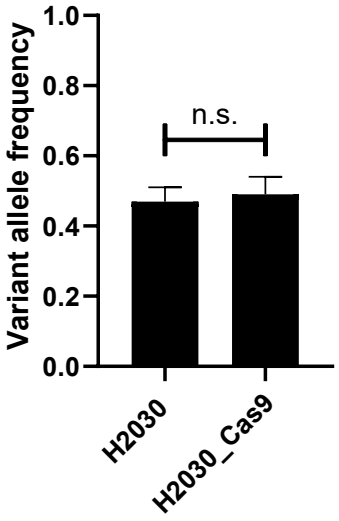


Figure 11 | Digital PCR of *KRAS* G12C mutation status in H2030 and H2030_Cas9 cells.

Genomic DNA was amplified with different dyes for the *KRAS* wildtype or mutant allele. The Variant allele frequency (VAF) was calculated, with 1.0 being equivalent to a homozygous- and 0.5 to a heterozygous allele distribution. The analysis of H2030 and H2030_Cas9 cells resulted in VAF values around 0.5, indicating no changes of the *KRAS* G12C mutation status after the stable integration of Cas9 into the cells and a heterozygous genotype in both samples.

Stable integration of Cas9 into H2030 cells did not change the VAF of *KRAS* G12C, indicating an unaltered mutation status after Cas9 integration. Both parental H2030 and H2030_Cas9 cells appeared to be heterozygous for the mutation, as indicated by a VAF value of 0.47 and 0.49, respectively. This finding is not in accordance with results published at cancer.sanger.ac.uk, where H2030 cells are reported to be homozygous for the *KRAS* G12C mutation[143]. If the mutation was homozygous in the samples, VAF values would have been around 1.0. However, since the results showed an unchanged status in H2030_Cas9 compared to parental H2030 cells, it was shown that the integration of Cas9 did not affect the G12C mutation. Consequently, the H2030_Cas9 cells were used for subsequent pooled CRISPR/Cas9 screening.

3.1.3 Determination of drug concentrations for pooled CRISPR/Cas9 screens

In order to see differential effects of synthetic lethality of genes with epigenetic drug treatment, a prerequisite is that the concentration of the epigenetic drug used for screening should only affect a minor fraction of the cells. Hence, the drug sensitivity of H2030_Cas9 as well as H1975_Cas9 cell lines was assessed to establish the highest possible dose that would not affect the cells.

3.1.3.1 DZNep

To determine the optimal concentration for pooled CRISPR/Cas9 screens, H2030 and H2030_Cas9 cells were treated with increasing concentrations of DZNep. After 72 hours, cells were stained with resazurin. analyzed using excitation at 570 nm and normalized to untreated cells (Figure 12).

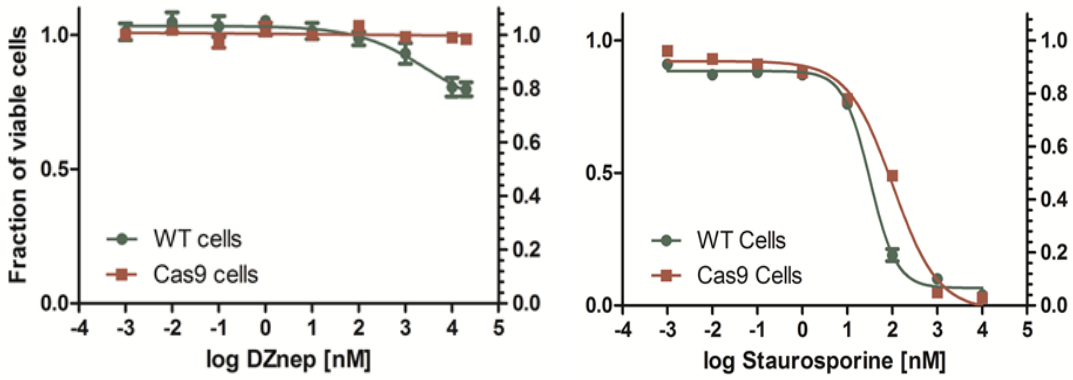


Figure 12 | DZNep titration in H2030 and H2030_Cas9 cells

To determine the optimal concentration for pooled CRISPR/Cas9 screens, H2030 and H2030_Cas9 cells were treated with increasing concentrations of DZNep or Staurosporine as a positive control. After 72 hours, the cell viability was measured by CellTiter Blue and normalized to untreated cells.

As shown in Figure 12 (left), the fraction of viable H2030_Cas9 cells remains unchanged even when treated with 10 µM DZNep. The viability of WT H2030 cells decreased to 80% when treated with 10 µM of the drug. At a drug concentration of 1 µM, H2030_WT cells showed only a minimal response to DZNep, suggesting the most suitable concentration for screening synthetically lethal genes. In contrast, using 10 µM DZNep would cause increased cell death due to the sole usage of the drug rather than due to synthetic lethality. Of note, a concentration of 1 µM DZNep has also been used in H2030 cells in a study in which synthetic lethality of DZNep and Topoisomerase-Inhibitors was assessed[124].

To confirm that the integration of Cas9 did not perturb the ability to induce apoptosis after drug treatment, the cells were treated with the non-selective protein kinase inhibitor Staurosporine, which is widely used as a general method to induce apoptosis[144] (Figure 12, right). With 50 % of viable cells compared to 20 % in parental H2030 cells, H2030_Cas9 cells showed decreased sensitivity to Staurosporine at 100 nM. However, at 1 µM, the fraction of viable cells in both cell lines decreased to less than 10 %.

3.1.3.2 Entinostat

Similar to the titration of DZNep, H2030_Cas9 and H1975_Cas9 cells, along with their parental lines, were treated with increasing concentrations of Entinostat to determine the appropriate concentration for pooled screens (Figure 13).

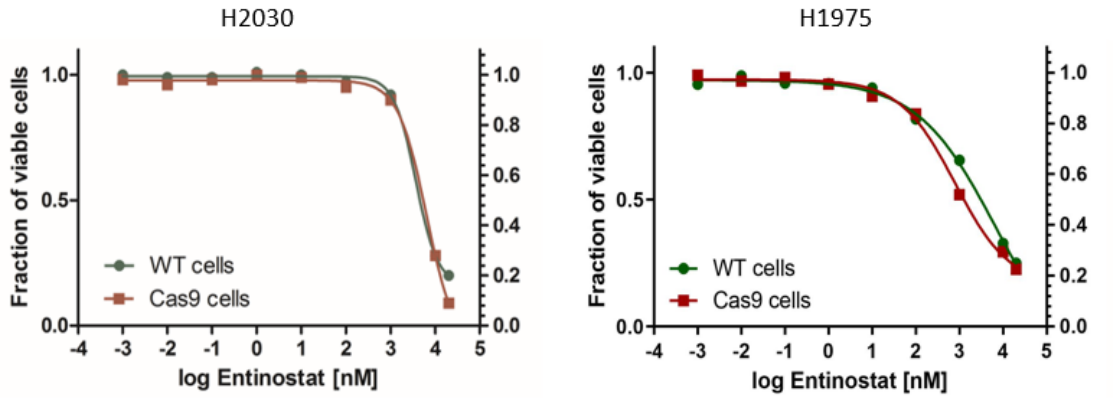


Figure 13 | Titration of Entinostat in H2030, H2030_Cas9, H1975, and H1975_Cas9 cells

To determine the optimal concentration for pooled CRISPR/Cas9 screens, H2030, H2030_Cas9, H1975, and H1975_Cas9 cells were treated with increasing concentrations of Entinostat. After 72 hours, the cell viability was measured by CellTiter Blue and normalized to untreated cells. In H2030_Cas9 cells, cell viability decreased above concentrations of 1 µM, and the viability of H1975 cells decreased above concentrations of 100 nM Entinostat.

In H2030_Cas9 cells, cell viability was decreased by concentrations above 1 µM Entinostat. However, as the cells appeared to be highly sensitive to the drug above this concentration, 300 nM was chosen for synthetic lethality screenings to not risk increased cell death by the sole use of the compound. H1975 cells showed to be more sensitive to Entinostat treatment. A concentration of 100 nM was chosen for a CRISPR/Cas9 screen for synthetically lethal genes in this cell line, since at this concentration, both H1975 and H1975_Cas9 showed no major response to the drug and the slope of the curve did not indicate a highly dynamic response above this concentration as observed in H2030 cells.

3.1.3.3 Puromycin

The lentiviral sgRNA library that was used in this study featured a Puromycin resistance marker; hence, Puromycin could be utilized as a selection drug for successfully transduced cells in pooled CRISPR/Cas9 screens. To evaluate the optimal concentration of Puromycin that would kill over 90 % of non-transduced H2030_Cas9 cells (IC-90),

the cells were treated with increasing concentrations of Puromycin and the cell viability was assessed (Figure 14).

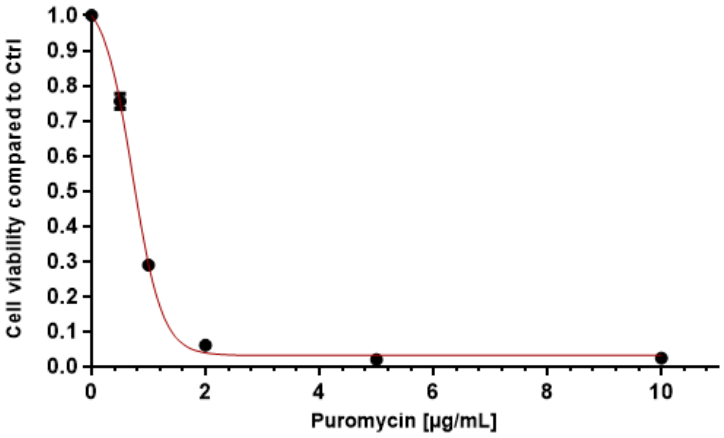


Figure 14 | Titration of Puromycin in H2030 cells

To determine the concentration to be used for selection after transduction with pooled lentiviral particles, H2030 cells were treated with increasing concentrations of Puromycin for 72 hours. Cell viability was afterwards assessed using a Cell Titer Blue assay. 2 µg/mL of Puromycin resulted in 6 % of remaining viable cells.

Using 2 µg/mL of Puromycin resulted in 6 % of remaining viable cells after 72 hours of treatment and was chosen to be used during selection steps in pooled CRISPR/Cas9 screens with H2030 cells. In the same fashion as for H2030 cells, the IC-90 of Puromycin was determined for H1975 cells (Figure 15).

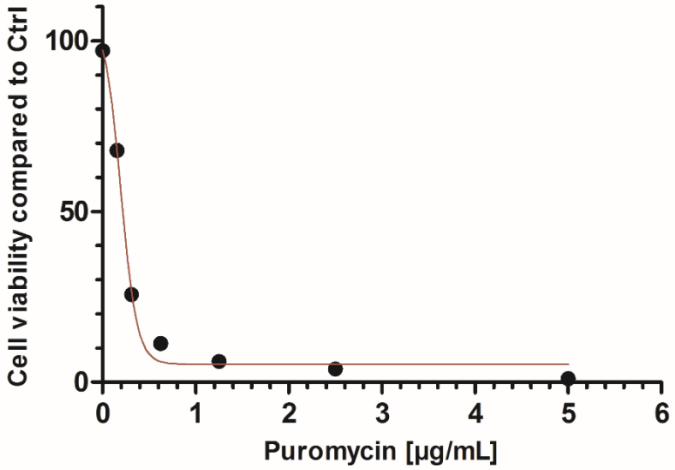


Figure 15 | Titration of Puromycin in H1975 cells

To determine the concentration to be used for selection after transduction with pooled lentiviral particles, H1975 cells were treated with increasing concentrations of Puromycin for 72 hours. Cell viability was afterwards assessed using a Cell Titer Blue assay. 1.25 µg/mL Puromycin resulted in 6 % of remaining viable cells.

Treatment of H1975 cells with 1.25 µg/mL Puromycin resulted in 6 % of remaining viable cells after 72 hours of treatment and was chosen to be used during selection steps in pooled CRISPR/Cas9 screens with H1975 cells.

In summary, the establishment of a pooled lentiviral CRISPR/Cas9 screen was successfully accomplished. Both cell lines that were transduced for stable expression of Cas9 showed high nuclease activity levels, making them suitable for pooled CRISPR screens. Further, the unchanged *KRAS* mutation status of H2030_Cas9 cells was validated, and drug concentrations for pooled screens in both cell lines were established. In the next step, the generated stably Cas9 expressing cell lines H2030_Cas9 and H1975_Cas9 were used for the identification of essential viability genes and synthetically lethal genes in pooled CRISPR/Cas9 screens.

3.2 Identification of essential viability genes and candidates for synthetic lethality with epigenetic drugs by pooled CRISPR/Cas9 screening

3.2.1 Pooled lentiviral gene knockout

To identify genes that are synthetically lethal with epigenetic drug treatment, the previously generated H2030_Cas9 and H9175_Cas9 cells were transduced with the *CRISPR Human Genome Knockout Library Module 1*, a pooled lentiviral library targeting 6,500 genes that have been annotated to be involved in various diseases, such as neurodevelopment, cardiovascular diseases or cancer (library design by Cellecta Inc, USA). Each gene was targeted by 6 - 8 different sgRNAs, resulting in a total of 52,549 sgRNAs. After the depletion of non-transduced cells by Puromycin, the cells were expanded and split into a no-drug arm and a treatment arm, where cells were treated with either DZNep or Entinostat for 14 days. Afterwards, the genomic DNA was isolated from the cells and the abundance of sgRNAs in each arm was determined by next-generation sequencing, allowing the identification of genes whose depletion impacted a cell's ability to confer resistance to the drug. An overview of the screening workflow is shown in Figure 16.

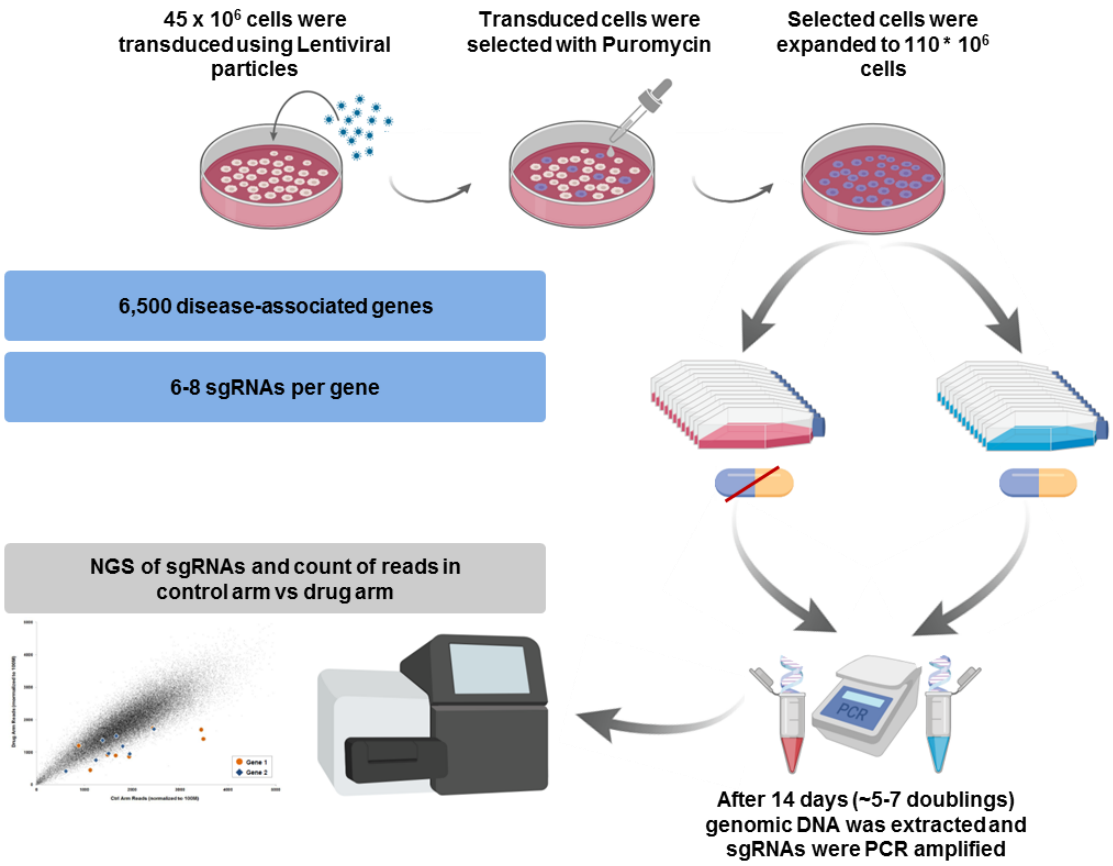


Figure 16 | Workflow of a pooled lentiviral gene knockout screen to identify genes acting synthetically lethal with epigenetic drug treatment

To start the screening process, 45 × 10^6 cells were transduced with the lentiviral particles and, after depletion of non-transduced cells by Puromycin, the cells were expanded to 110 × 10^6 cells and split into a no-drug arm and a drug treatment arm, where cells were treated with either DZNep or Entinostat for 14 days. Afterwards, the genomic DNA was isolated from the cells and the abundance of sgRNAs in each arm was determined by next-generation sequencing to compare the readcounts in the two arms.

3.2.1.1 Production and titration of lentiviral sgRNA particles

Before conducting the screen, the lentiviral particles were produced in HEK-293T/17 cells. For this study, 140 mL of lentiviral supernatant containing a lentiviral sgRNA library targeting 6,500 genes simultaneously was produced. This volume would be sufficient to run approximately 100 screens. Hereafter, the volume of pooled lentiviral sgRNA particles that would result in approximately 30 % of transduced cells was established. At this concentration, most cells would not receive any lentiviral particle, but cells that are transduced would predominantly receive only one viral particle and have one particular gene knocked out. For pooled sgRNA library screens, it is important to ensure that only a minority of cells receives two or more viral particles, as this may bias the results by having multiple knockouts within the same cell. Following a

transduction of H2030_Cas9 cells with increasing amounts of lentiviral particles, the fraction of RFP-positive cells was assessed by flow cytometry (Figure 17).

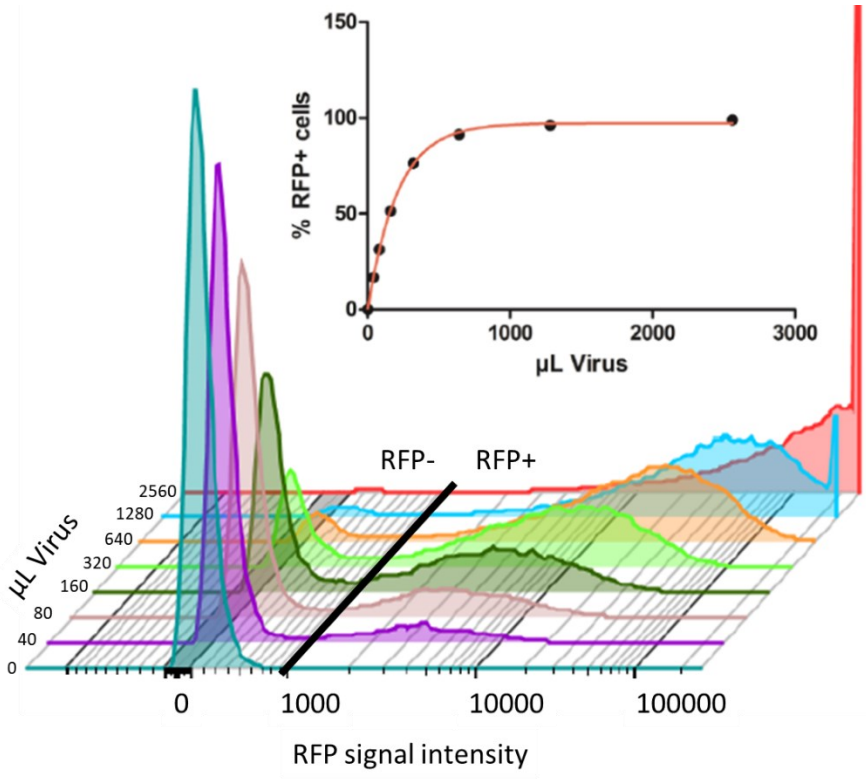


Figure 17 | Determination of the optimal virus volume in H2030_Cas9 cells by flow cytometry

To establish the volume of pooled lentiviral sgRNA particles that would result in approximately 30 % of transduced cells, H2030_Cas9 cells were transduced with increasing volumes of pooled lentiviral particles. The fraction of RFP(+) cells was assessed by flow cytometry 72 hours post-transduction.

As indicated in Figure 17, using 80 µL of viral particles per 2×10^6 H2030_Cas9 cells resulted in 31.3 % of cells stained positive for RFP, and this volume was chosen for viral transduction in this cell line.

Similarly, the fraction of RFP(+) H1975_Cas9 cells was determined by flow cytometry 72 hours post transduction (Figure 18).

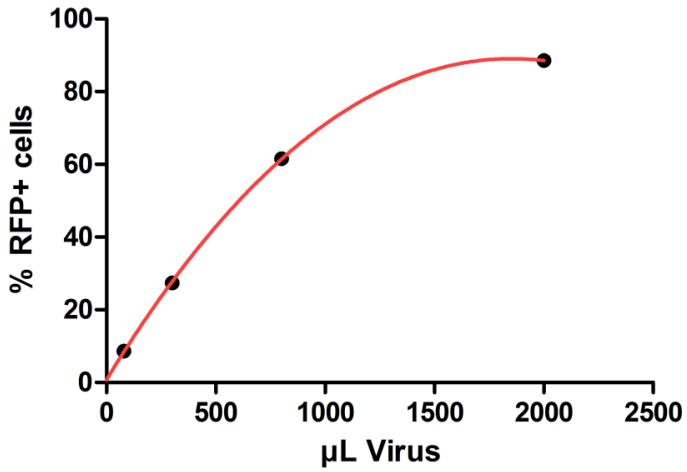


Figure 18 | Determination of the optimal virus volume in H1975_Cas9 cells

To establish the volume of pooled lentiviral sgRNA particles that would result in approximately 30 % of transduced cells, H1975_Cas9 cells were transduced with increasing volumes of pooled lentiviral particles. The fraction of RFP+ cells was assessed by flow cytometry 72 hours post-transduction.

Transduction of H1975_Cas9 cells with 300 µL of pooled viral particles resulted in 27 % of RFP^+ cells. Linear regression resulted in a volume of 340 µL of pooled lentiviral particles per 1.4×10^6 H1975_Cas9 cells to achieve the desired 30 % of RFP^+ cells.

3.2.1.2 Pooled gene knockdown

Using the established volumes of viral particles, a pooled sgRNA library screen was set up with H2030_Cas9 cells and the fraction of RFP(+) cells was assessed by flow cytometry 72 hours post transduction (Figure 19).

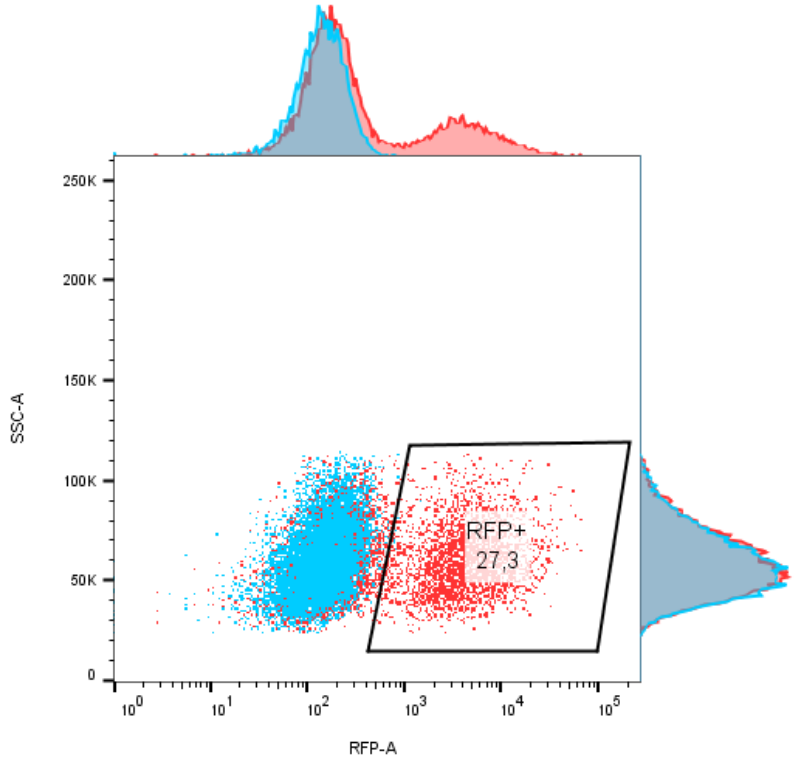


Figure 19 | Determination of the fraction of transduced cells by flow cytometry

The fraction of RFP(+) H2030_Cas9 cells was assessed 72 hours post-transduction. 27.3 % of cells stained positive for RFP.

As expected from the previously performed assessment of lentiviral transduction efficacy, around 30 % of cells stained positive for RFP (27.3 %) in H2030_Cas9 cells. The cells were treated with 2 µg/mL Puromycin to deplete non-transduced cells. The selected cells were expanded for 72 hours and split into a control / no-drug arm and a DZNep treatment arm, each consisting of 55×10^6 cells in 32 T175 flasks. The cells were treated with 1 µM DZNep or 300 nM Entinostat for 14 days in the treatment arm and the equivalent of DMSO in the no-drug arm.

Similar to H2030_Cas9 cells, H1975_Cas9 cells were seeded and transduced with lentiviral CRISPR library particles with the previously established volume. After 72 hours, the fraction of RFP(+) cells assessed by flow cytometry was 19.1 % (Supplementary figure 1).

Following this, the H1975_Cas9 cells were treated with 100 nM Entinostat and cultured using the same protocol as for H2030_Cas9 cells.

3.2.1.3 Amplification of sgRNAs

Upon transduction, sgRNAs integrate into the genomic DNA of a cell. To analyze the abundance of sgRNAs in the cell population after the phenotypic screen is complete, the genomic DNA was extracted, and the sgRNAs were amplified for subsequent high-throughput deep sequencing. A representative agarose gel of the PCR products of the no-drug arm of the DZNep screen in H2030_Cas9 cells is shown in Figure 20.

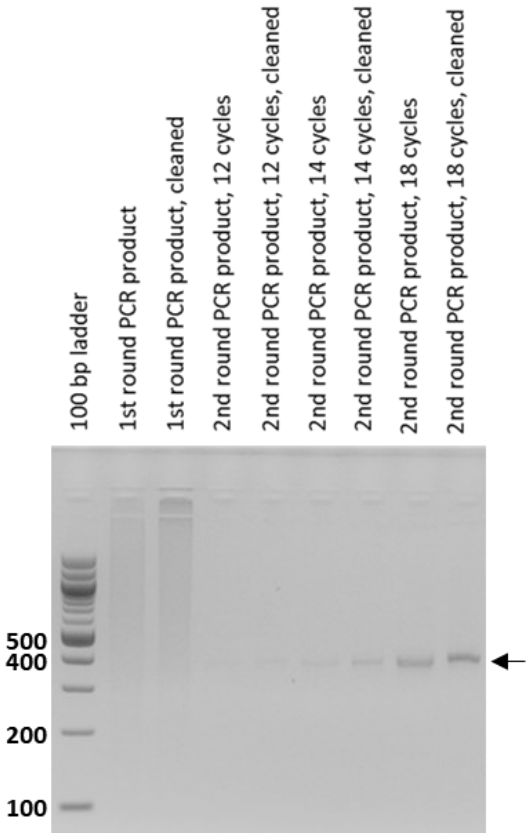


Figure 20 | 3.5 % Agarose/TAE gel of 1st and 2nd round PCR products in H2030_Cas9 cells

The sgRNAs were amplified using a two-step nested PCR to decrease genomic carryover, and 5 µg of DNA was loaded per lane. The arrow indicates the amplified product with a size of 383 bp, corresponding to the expected fragment size.

The agarose gel showed a band of approx. 400 bp that corresponded to the estimated PCR product of 383 bp (Supplementary figure 2). As expected, the intensity of the band increased with the number of cycles of the 2nd PCR round. Cleaning of the PCR product did not reduce the band intensity. As it gave the strongest signal, the PCR product generated with 18 cycles was cut out, the DNA was extracted from the gel and the DNA concentration was quantified (Table 25) and adjusted to 10 nM.

Table 25 | DNA concentration of pooled CRISPR/Cas9 screening samples after agarose gel extraction

Cell line	Sample	Mass concentration (ng/µL)	Molar concentration (nM)
H2030_Cas9	No-drug arm	4.84	24.82
H2030_Cas9	DZNep arm	4.34	22.26
H2030_Cas9	Entinostat arm	3.40	17.44
H1975_Cas9	No-drug arm	3.53	18.10
H1975_Cas9	Entinostat arm	3.38	17.33

The sgRNAs were sequenced on a HiSeq2000 instrument (Illumina) at the Genomics and Proteomics Core Facility at the German Cancer Research Center (DKFZ). The number of resulting reads is shown in Table 26.

Table 26 | Number of reads resulting from deep sequencing of sgRNAs

Cell line	Sample	Reads	Average reads per sgRNA
H2030_Cas9	No-drug arm	134454823	2558
H2030_Cas9	DZNep arm	133182842	2534
H2030_Cas9	Entinostat arm	101908491	1939
H1975_Cas9	No-drug arm	129378625	2462
H1975_Cas9	Entinostat arm	128952501	2453

In conclusion, the sgRNAs were successfully amplified from genomic DNA, and a high number of reads per sgRNA indicated a robust screen. Next, the screen robustness was validated by the analysis of internal control sgRNAs.

3.2.1.4 Evaluation of screen robustness

In order to assess the reliability of the results, the abundance of 88 positive (cell death) control sgRNAs and 172 negative (non-target) control sgRNAs was compared to sequencing results of the lentiviral vector library (provided by Cellecta). For this, the sgRNA reads were mapped to the corresponding genes using the *CRISPRAnalyzeR* platform[145] The read counts of the no-drug screen arm and the vector library were normalized to 100 million reads, and the log2 foldchange of the no-drug arm to the vector library was calculated for all sgRNAs. The sgRNAs were ranked from negative log2 foldchange (indicating depleted sgRNAs in the no-drug screen arm vs. the vector library) to positive log2 foldchange (indicating enriched sgRNAs in the screen arm vs. the vector library). The results for H2030_Cas9 and H1975_Cas9 cells are shown in Figure 21.

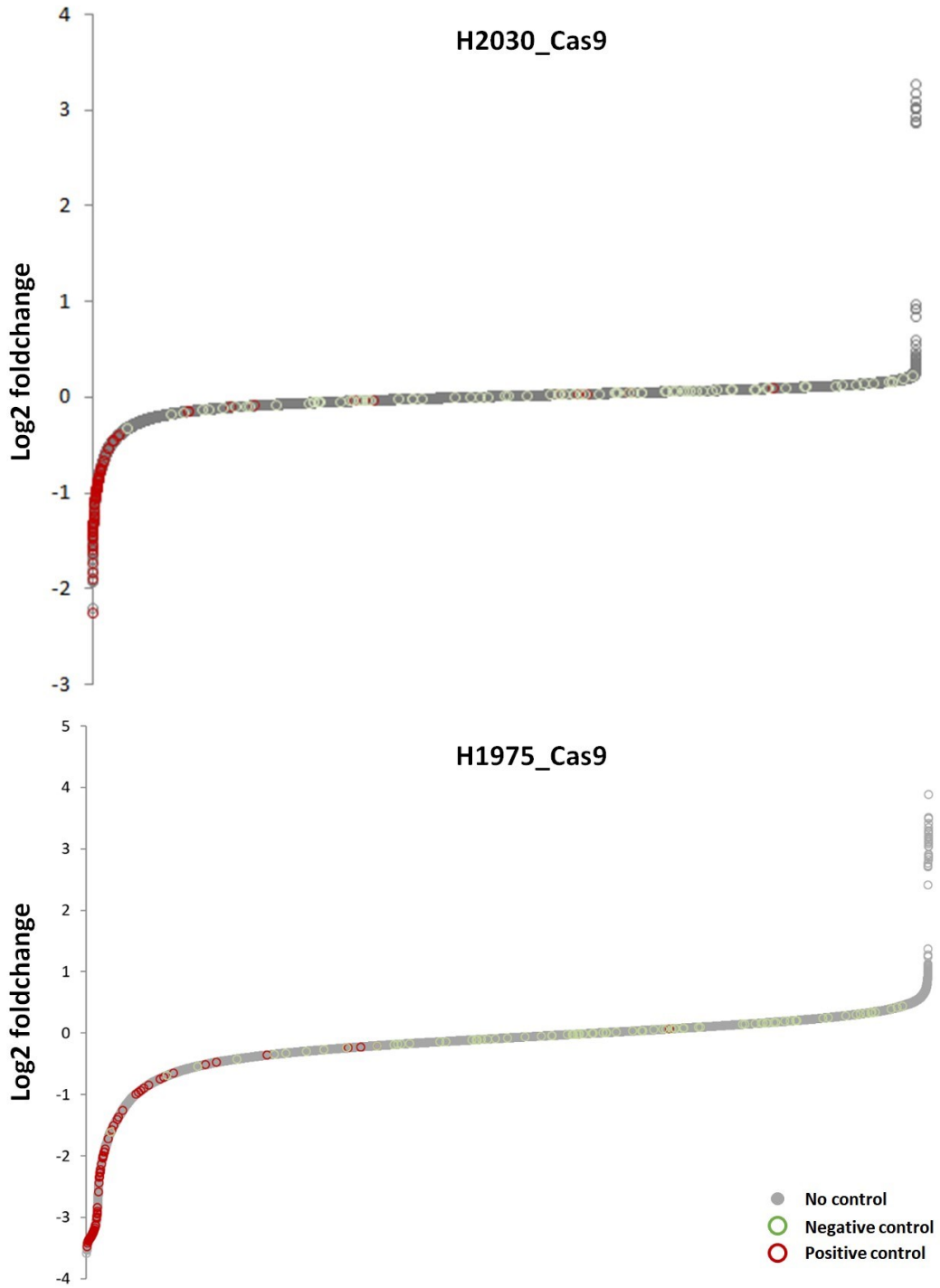


Figure 21 | Ranked representation of log2 sgRNA foldchanges of the no-drug screen arm vs. normalized vector library readcounts in H2030_Cas9 and H1975_Cas9 cells

The readcounts of all samples were normalized to 100 million, and the log2 foldchange was calculated. sgRNAs were ranked from negative log2 foldchange (indicating depleted sgRNAs in the no-drug arm vs. the vector library) to positive log2 foldchange (indicating enriched sgRNAs in the no-drug screen arm vs. the vector library). Red circles represent positive (cell death) controls, and green circles represent negative (non-target) controls. The vector library readcounts were kindly provided by Cellecta, Inc. For both cell lines, positive controls are accumulated on the left, proving successful gene knockout. Negative sgRNA controls were randomly distributed in both cell lines, indicating the absence of off-target effects.

As shown in Figure 21, in both cell lines, most sgRNA positive controls (red) were enriched at the left side, indicating that a knockout of eleven positive controls[A] led to cell death and, as a result, to the depletion of the sgRNAs from the cell pool. Because the positive controls are essential for cell viability of eukaryotic cells, the accumulation on the left of the ranked representation indicated that Cas9 had successfully facilitated gene knockout. On the other hand, it was expected that the negative (non-target) control sgRNAs would not be depleted or accumulated. Accordingly, as seen in Figure 21, they were distributed across the entire range, indicating that non-target sgRNAs did not impair cell viability and ruling out off-target effects. Taken together, these results showed that the read counts obtained by next-generation sequencing after pooled CRISPR/Cas9 screening in H2030 and H1975 cells were suitable for further analysis.

3.2.2 Identification of essential viability genes

After having ensured screening data quality, essential viability genes of the cell lines were identified by calling genes that were depleted in the no-drug screen arm compared to the vector library. To ensure the comparability of both samples, they were normalized to a total read count of 100 million reads per sample. Identification of essential viability genes served two purposes: First, it would be a good indicator for screening robustness if published viability genes for these cell lines were reproduced. Second, the identification of new relevant genes may be a starting point for further studies, as a loss of such genes results in impaired cell viability, making them ideal targets to improve the understanding of tumor cell biology in *EGFR* and *KRAS*-mutant NSCLC cells.

3.2.2.1 Essential viability genes of H2030_Cas9 cells

The read counts of the no-drug screen arm vs. the lentiviral library were analyzed to identify essential viability genes in H2030_Cas9 cells. Genes that were significantly depleted in the no-drug screen arm vs. the lentiviral library were identified by five independent algorithms (Wilcoxon signed-rank test, EdgeR, sgRSEA, MAGeCK, and

[A] *Cell Division Cycle 16 (CDC16); General Transcription Factor IIB (TF2B); Heat Shock Protein Family A (Hsp70) Member 5 (HSPA5); Heat Shock Protein Family A (Hsp70) Member 9 (HSPA9); Platelet Activating Factor Acetylhydrolase 1b Regulatory Subunit 1 (PAFAH1B1); Proliferating Cell Nuclear Antigen (PCNA); RNA Polymerase II Subunit L (POLR2L); Ribosomal Protein L9 (RPL9); Splicing Factor 3a Subunit 3 (SF3A3); Splicing Factor 3b Subunit 3 (SF3B3); Tubulin Beta Class I (TUBB)*

DESeq) using the *CRISPRAnalyzeR* platform[145] (Figure 22, see Supplementary figure 3 for a scatterplot representation of the read counts).

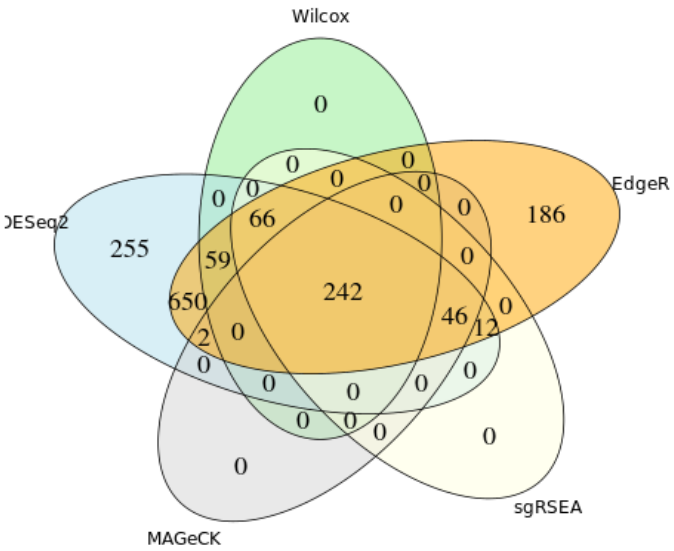


Figure 22 | Venn diagram of viability genes in H2030_Cas9 cells

The read counts of the no-drug screen arm vs. the lentiviral library were analyzed using the CRISPRAnalyzeR platform, using five algorithms and resulted in 242 significantly depleted genes in the no-drug screen arm.

As depicted in Figure 22, 242 genes were identified by all five algorithms. A list of all hits is enclosed in Supplementary Table 1. All positive controls (n = 11) of the library appeared among the identified hits and were removed, resulting in 231 shortlisted viability genes. Notably, previously reported genes that are essential for the viability of *KRAS* mutant NSCLC appeared among the shortlisted genes. For example, the relative median readcount of 8 sgRNAs targeting *Polo-like kinase 1* (*PLK1*) was reduced by 88 % compared to the lentiviral library, indicating a strong dependence of the H2030_Cas9 cells on this gene. This finding is in agreement with published results[72]. Likewise, the relative median read count of sgRNAs for *Aurora kinase A*, a previously identified target in *KRAS* mutant NSCLC[146], was reduced by 79 %. It has been reported that *KRAS* mutant / *LKB1* deficient H2030 NSCLC cells exhibit increased sensitivity to *MAPK* and *mechanistic Target of Rapamycin* (*mTOR*) signaling inhibition[147]. In agreement with this report, the relative median readcount of 8 sgRNAs targeting mTOR was reduced by 59 % compared to the plasmid library. The relative median read count of sgRNAs targeting the *KRAS* gene was reduced by 52 %. A representation of foldchanges of *KRAS* signaling-related genes can be found in the supplementary materials of this work (Supplementary figure 4).

To assess the functional implications of the 231 shortlisted essential viability genes, KEGG and Biocarta pathways were annotated using DAVID[148] v6.8 (Figure 23).

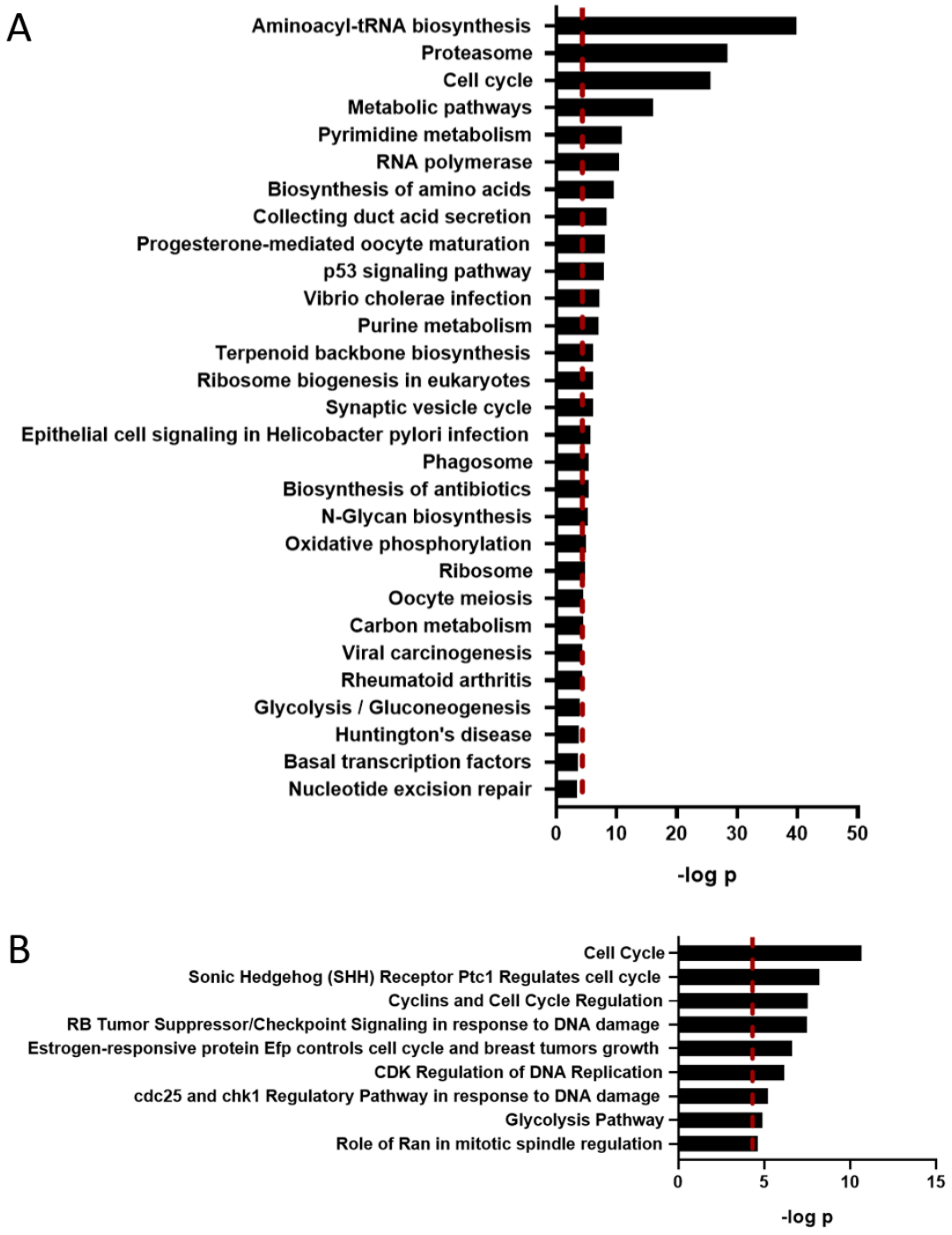


Figure 23 | Functional annotation of essential viability genes of H2030_Cas9 cells

Identified essential viability genes in H2030_Cas9 cells were functionally annotated using (A) KEGG and (B) Biocarta pathways. -log p-values larger than 4.32 (red line, equivalent to p = 0.05) indicate a significant gene enrichment.

According to the KEGG analysis, genes involved in *Aminoacyl-tRNA biosynthesis* were most enriched among the identified viability genes (represented by the highest -log p-value in Figure 23 A), followed by *proteasome-*, *cell cycle-, metabolic pathways- and pyrimidine metabolism*-related genes. Noteworthy, *p53 signaling* was also significantly enriched within the set of identified viability genes. In the Biocarta pathway analysis (Figure 23 B), an enrichment of H2030_Cas9 viability genes involved in the *cell cycle* was found to be most significant. Also, *Cyclins and cell cycle regulation*, as well as *CDK regulation and DNA replication* were among the enriched Biocarta pathways, thereby supporting the role of the identified genes as key viability genes for H2030_Cas9 cells.

3.2.2.2 Essential viability genes of H1975_Cas9 cells

Essential viability genes of the H1975_Cas9 cell line were identified by calling genes that were depleted in the normalized no-drug screen arm compared to the normalized vector library (see Supplementary figure 5 for a scatterplot representation).

Similar to the identification of essential viability genes of H2030_Cas9 cells, the normalized read counts of the no-drug arm of transduced H1975_Cas9 cells were compared to the normalized read counts of the lentiviral library using the *CRISPRAnalyzeR* platform. Genes that were significantly depleted in the no-drug screen arm vs. the lentiviral library were identified by the five independent algorithms Wilcoxon signed-rank test, EdgeR, sgRSEA, MAGeCK, and DESeq[145] (Figure 24).

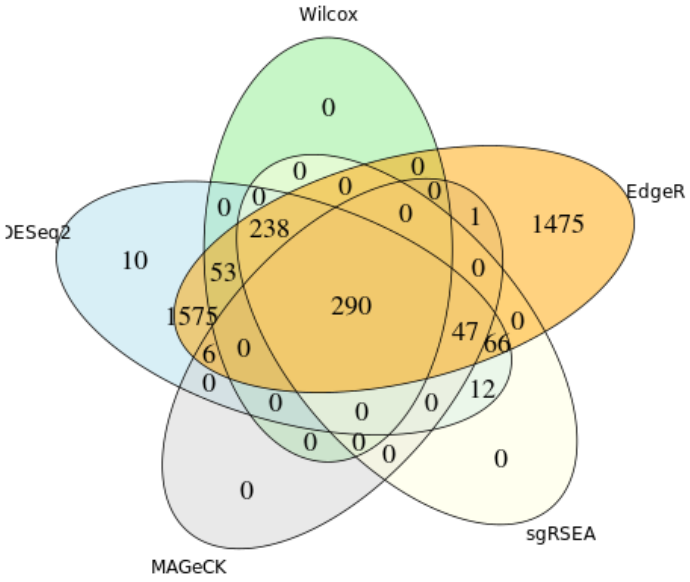


Figure 24 | Venn diagram of viability genes in H1975_Cas9 cells

Read counts of the no-drug screen arm of H1975_Cas9 cells vs. the lentiviral library were analyzed using the CRISPRAnalyzeR platform, using five algorithms and resulted in 290 significantly depleted genes in the no-drug screen arm.

290 genes were identified as critical for the viability of H1975_Cas9 cells. A list of all hits is enclosed in Supplementary table 2. As for H2030_Cas9 cells, all positive controls (n = 11) of the library appeared among the identified hits and were removed. Importantly, the relative median read count of the sgRNAs targeting the *EGFR* gene was reduced by 85 % in the transduced cells when compared to the plasmid library, indicative of a strong dependence of H1975_Cas9 cells on this driver gene[149].

To elucidate the functional implications of the shortlisted viability genes, KEGG and Biocarta pathways were annotated using DAVID (Figure 25).

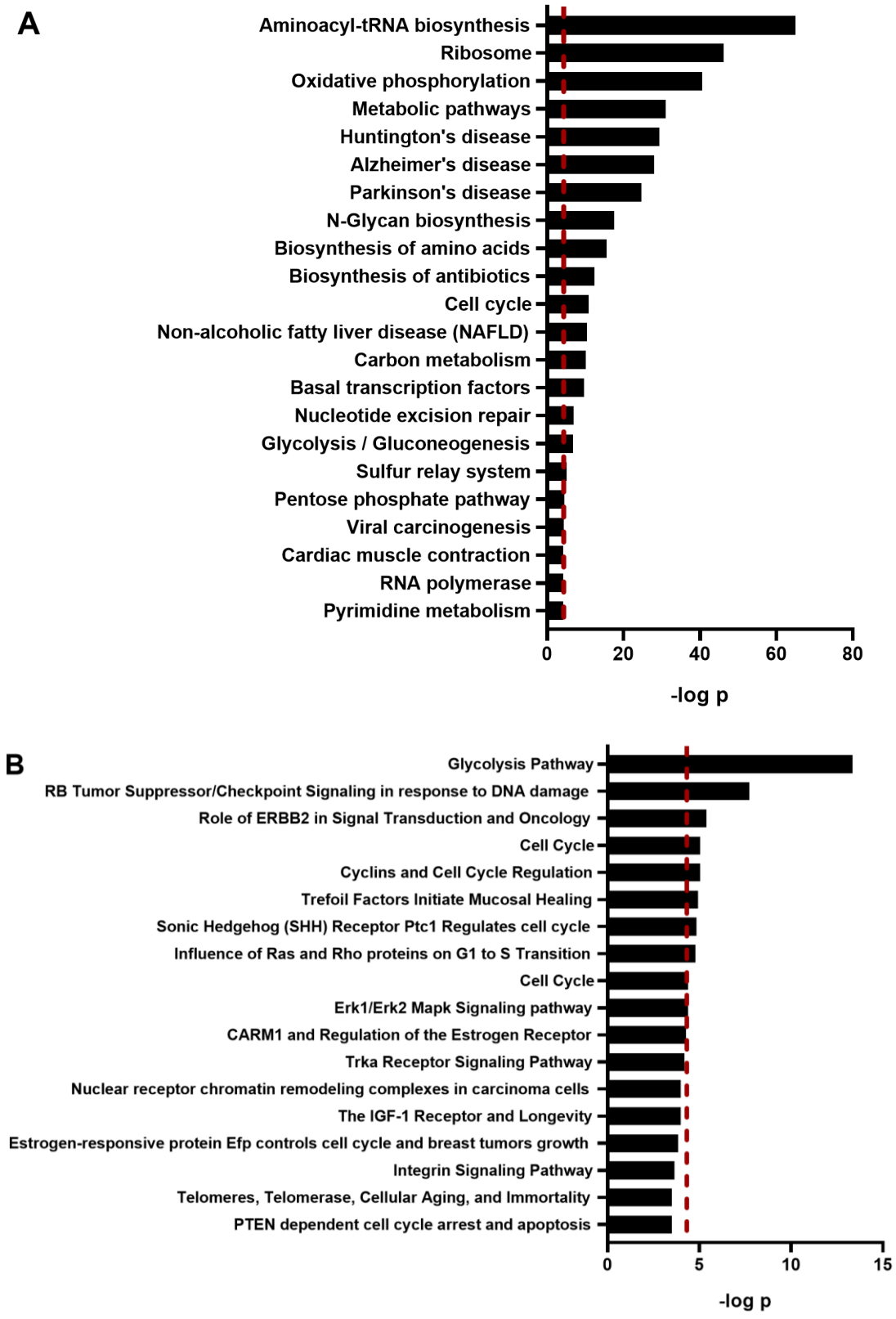


Figure 25 | Functional annotation of essential viability genes of H1975_Cas9 cells

Identified essential viability genes of H1975_Cas9 cells were functionally annotated using (A) KEGG and (B) Biocarta pathways. -log p-values larger than 4.32 (red line, equivalent to p = 0.05) indicate a significant gene enrichment.

The KEGG pathways analysis (Figure 25 A) revealed significant enrichment of genes associated to *Aminoacyl-tRNA biosynthesis*, followed by genes related to functions of the *Ribosome*, *oxidative phosphorylation* and *metabolic pathways*. The analysis also revealed significantly enhanced enrichment of H1975's essential viability genes in *Alzheimer's-*, *Huntington's-* and *Parkinson's disease*.

In the Biocarta analysis (Figure 25 B), genes related to the *glycolysis pathway* were most enriched among the identified viability genes. Other significantly enriched pathways indicative for the role of the identified genes as viability genes included *RB Tumor Suppressor/Checkpoint Signaling in response to DNA damage*, *Role of ERBB2 in Signal Transduction and Oncology*, *Cell Cycle*, and *Cyclins and Cell Cycle Regulation*. Genes involved in *PTEN dependent cell cycle arrest and apoptosis* showed a trend towards enrichment (-log p = 3.52), in accordance with findings that a deletion of *PTEN* in *EGFR*-mutant NSCLC correlates with decreased progression-free survival[150].

3.2.2.3 Comparison of H2030_Cas9 and H1975_Cas9 viability genes

To understand the distribution of identified viability genes between the two cell lines, the overlap was calculated as shown in Figure 26.

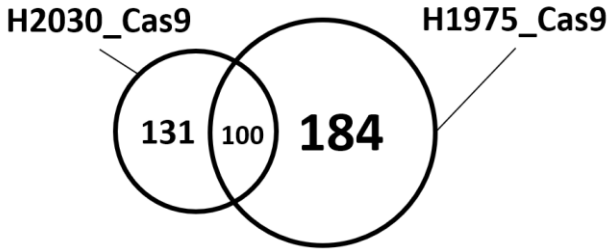


Figure 26 | Venn diagram of shortlisted essential viability genes

The Venn diagram represents the overlap of all identified viability genes in H2030_Cas9 and H1975_Cas9 cells. 75 % of all genes are non-overlapping, and 25 % of genes were identified as viability genes in both cell lines.

The majority of shortlisted essential viability genes of *KRAS* mutant H2030_Cas9- and *EGFR* mutant H1975_Cas9 NSCLC cells appeared to be cell line-specific, with 131 genes found exclusively in H2030_Cas9- and 184 genes in H1975_Cas9 cells. In H2030 cells, due to the lower number of total identified viability genes (n = 231), 43 % of the genes (n = 100) were also found to be viability genes in H1975 cells, whereas this was only the case in 35 % of the genes that were identified in H1975_Cas9 cells (n = 284).

However, considering the total number of identified genes, 25 % of all identified genes appeared in both cell lines, and 75 % appeared in one cell line exclusively, suggesting a specific function for the majority of those genes and tailored to the genetic background of the cells, rather than being genes which generally affect cell viability.

Taken together, both *KRAS* and *EGFR* could be confirmed as essential viability genes in the respective cell lines bearing activating mutations in the *KRAS*- (H2030_Cas9) or *EGFR* (H1975_Cas9) gene. Additionally, various published targets for *KRAS* mutant NSCLC cells were reproduced, while at the same time a set of novel potential target genes for both *KRAS* and *EGFR* mutant NSCLC cells was identified. However, these targets would need to be investigated in subsequent research projects, as the main focus of this study was to identify synthetically lethal genes with epigenetic drugs.

3.2.3 Synthetically lethal genes with epigenetic drugs

The identification of essential viability genes confirmed previously published targets and pathways and may be used to shortlist novel targets that are essential in *KRAS* and *EGFR* mutated NSCLC. However, the main objective of this work was to identify gene targets that do not impair cell viability when being depleted but act synthetically lethal when depletion is combined with epigenetic drug treatment.

3.2.3.1 Synthetically lethal genes with DZNep treatment in H2030_Cas9 cells

In order to identify synthetically lethal genes with 1 µM DZNep, the read counts of the no-drug arm and the DZNep treatment arm of the screen were normalized to 100 million reads. The read counts were submitted to *CRISPRAnalyzeR,* and genes that were significantly depleted in the DZNep arm vs. the no-drug arm were called by five different algorithms: Wilcoxon signed-rank test, EdgeR, sgRSEA, MAGeCK, and DESeq. *Protein arginine N-methyltransferase 1* (*PRMT1*), *Chromodomain-helicase-DNA-binding protein 8* (*CHD8), Lysine-specific demethylase 5C* (*KDM5C)* and *Histone acetyltransferase p300* (*EP300)* were identified by all algorithms and shortlisted for further experiments (Figure 27 and Table 27).

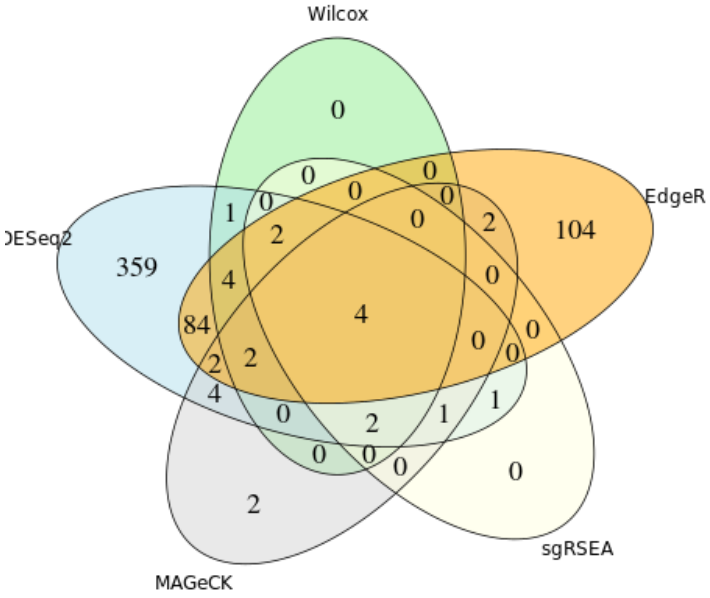


Figure 27 | Venn diagram of algorithmic identification of synthetically lethal genes with 1 µM DZNep in H2030_Cas9 cells

The readcounts of the no-drug screen arm vs. the DZNep treatment art were analyzed using the CRISPRAnalyzeR platform, using five algorithms, and resulted in four significantly depleted genes in the DZNep treatment arm.

Table 27 | Foldchanges between the 1 µM DZNep arm and the no-drug arm in H2030_Cas9 cells

	Fold Change	P-values				
		Wilcox	DESeq2	MAGeCK	sgRSEA	EdgeR
PRMT1	0.35	0.008	5.54E-243	0.00099	0.003	0.024
CHD8	0.56	0.037	1.05E-131	0.00099	0.003	0.037
KDM5C	0.66	0.027	9.435E-79	0.00248	0.003	0.024
EP300	0.70	0.022	1.828E-51	0.01845	0.043	0.024

PRMT1 exhibited the lowest median fold change (0.35, indicating 35 % abundance of sgRNAs for *PRMT1* in the drug arm compared to the no-drug arm), followed by *CHD8 (*0.56), *KDM5C* (0.66) and *EP300* (0.70). However, analysis of the no-drug screen arm revealed that knockout of *PRMT1* resulted in significantly depleted sgRNA read counts even without DZNep treatment (data not shown). Although DZNep treatment further decreased sgRNA read counts significantly, the gene was excluded from further analysis due to being an essential cell viability gene. Because *KDM5C* was also identified as synthetically lethal with the HDAC inhibitor Entinostat in a parallel pooled CRISPR/Cas9 screen with H2030_Cas9 cells, thereby indicating an unspecific mechanism rather than a direct link to DZNep treatment, the gene was not chosen for further analysis but still may be an interesting candidate for future studies. *CHD8* and

EP300 were selected for further analysis. A scatterplot and a ranked representation of sgRNAs for *CHD8* and *EP300* are shown in Figure 28.

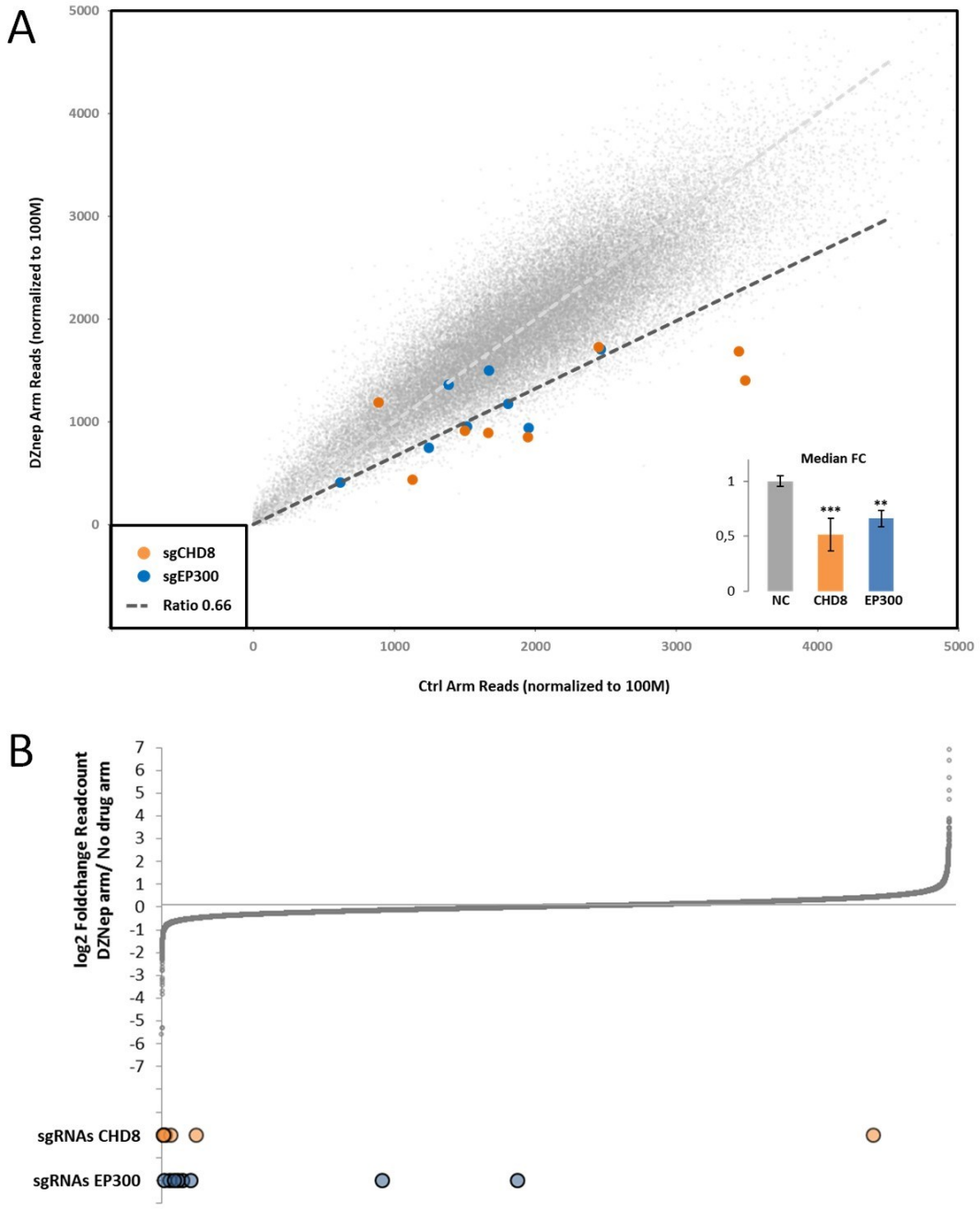


Figure 28 | (A) sgRNA scatterplot in H2030_Cas9 cells treated with DZNep vs. the no-drug arm (B) Ranked representation of *CHD8* and *EP300* sgRNAs

(A) The readcounts of the DZNep- and the no-drug screen arm were normalized to 100 million reads. The abundance of all individual sgRNAs in both study arms was compared, with each dot representing one sgRNA. sgRNAs targeting *CHD8* and *EP300* were highlighted in orange (*CHD8*) and blue (*EP300*). Statistical significance of median foldchange of sgRNAs was calculated using a student`s t-test. **(B)** The logarithmic foldchange of sgRNA abundance in the DZNep arm vs. the no-drug arm was ranked from most negative value (left, indicating the most depleted sgRNAs in the DZNep arm) to the most positive value (right, indicating enriched sgRNAs in the DZNep arm).

As depicted in the scatterplot of Figure 28 (A), 6 out of 8 sgRNAs for *CHD8* and 5 out of 8 sgRNAs for *EP300* exhibited fold changes below 0.66, indicating high synthetic lethality. The abundance of non-target sgRNAs (NC) was not altered between the no-drug and the DZNep arms (barplot, grey bar), whereas sgRNAs targeting *CHD8* or *EP300* were significantly depleted upon treatment of the cells with 1 µM DZNep (barplot, orange and blue bars). In part B of the figure, sgRNAs for *CHD8* and *EP300* are enriched at the left side of the ranked representation, again illustrating diminished read counts in the DZNep arm compared to the no-drug arm of the screen.

CHD8 was initially identified as a negative regulator of the Wnt/b-catenin signaling pathway[151] and is a member of the chromodomain helicase (CHD) family that belongs to the SNF2 superfamily. It contains two chromodomains: a helicase/ATPase domain and a DNA-binding domain[152]. CHD proteins have important roles during development, neurogenesis and are reported to be of importance for a variety of diseases[153]. Notably, *CHD8* is frequently mutated, deleted, or silenced in various cancer entities[154–156], making *CHD8* an interesting target for further elucidation of synthetic lethality in lung cancer cells.

EP300, also known as the *E1A-associated protein p300*, functions as a histone acetyltransferase[157] and regulates transcription via chromatin remodeling[158]. It mediates acetylation of histone H3 at Lysine residue 122 (H3K122ac), thereby enhancing gene transcription. Additionally, *EP300* mediates acetylation of histone H3 at lysine residue 27 (H3K27ac), which is also described as an active transcriptional enhancer mark[159]. *EP300* has been linked to lung adenocarcinoma, colon adenocarcinoma, bladder urothelial carcinoma, breast invasive ductal carcinoma, and endometrial endometrioid adenocarcinoma[160] and is involved in the regulation of fundamental biological processes such as proliferation, cell cycle, cell differentiation, and the DNA damage response[161], making it an attractive target for synthetic lethality.

3.2.3.2 Synthetically lethal genes with Entinostat treatment in H2030_Cas9 cells

In order to identify synthetically lethal genes with 300 nM Entinostat in H2030_Cas9 cells (as determined in 3.1.3.2), the read counts of the no-drug arm, as well as the Entinostat treatment arm of the screen were normalized to 100 million reads. The read counts were submitted to *CRISPRAnalyzeR,* and genes that were significantly depleted in the DZNep arm vs. the no-drug arm were called by five different algorithms.

However, no overlapping hits were identified by all five algorithms, calling for an alternative strategy for the identification of potential synthetically lethal genes with Entinostat in this cell line. By applying a selection criterion of a minimum of three sgRNAs with foldchanges of < 0.6, *Glutathione peroxidase 4* (*GPX4*) and *Lysine-specific demethylase 5C* (*KDM5C*) were shortlisted (Figure 29).

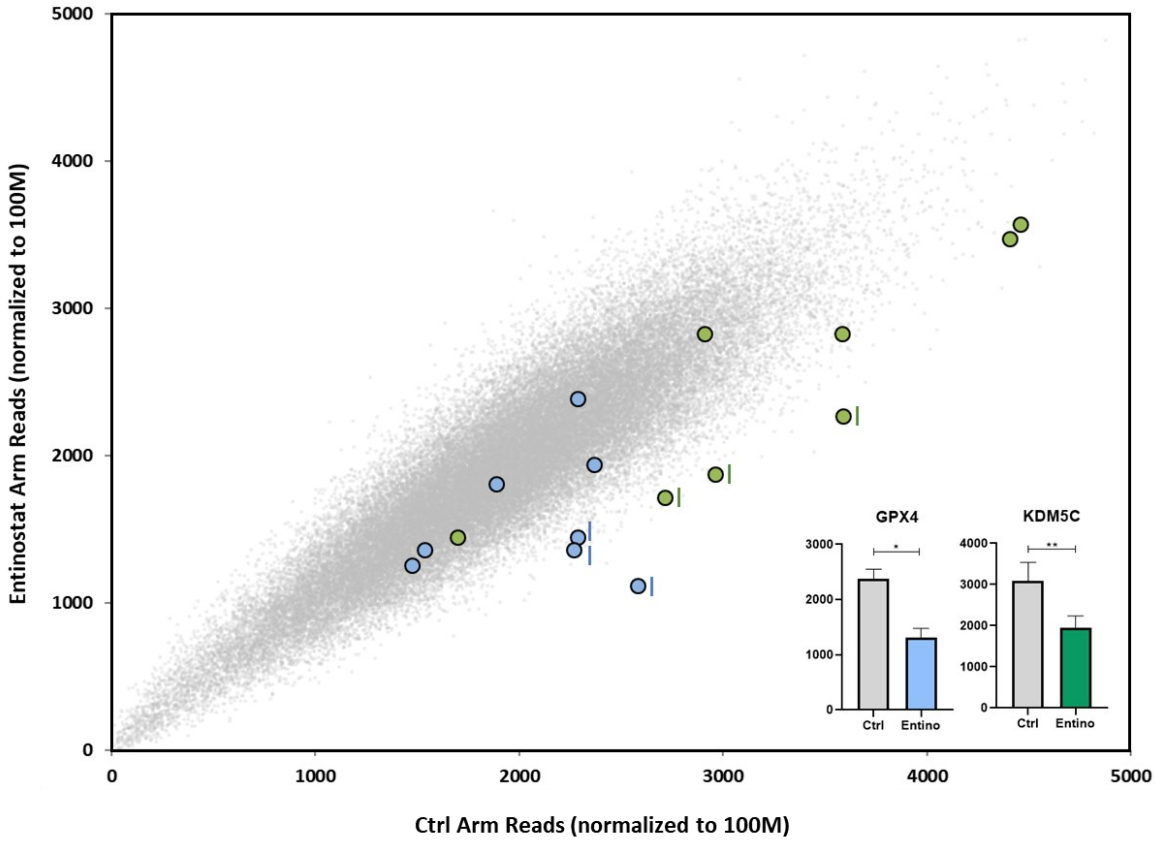


Figure 29 | sgRNA scatterplot in H2030_Cas9 cells treated with Entinostat vs. the no-drug arm

The readcounts of both study arms were normalized to 100 million reads. The abundance of all individual sgRNAs in both study arms was compared, with each dot representing one sgRNA. sgRNAs that were considered for the analysis are marked by a bar next to the dot. sgRNAs targeting *GPX4* and *KDM5C* were highlighted in blue (*GPX4*) and green (*KDM5C*).

GPX4, a phospholipid hydroperoxidase, protects cells against membrane lipid peroxidation causing oxidative stress. Loss of *GPX4* leads to an accumulation of lipid peroxides, resulting in cell death[162]. Consequently, considering the implications of a depletion of *GPX4* on healthy cells, the gene does not appear to be a promising candidate for further studies of synthetic lethality in lung cancer cells.

KDM5C, a member of the *KDM5* family, is an oxygenase acting as histone 3 lysine 4 trimethyl (H3K4me3) demethylase. It is involved in regulating proliferation, stem cell self-renewal, and differentiation. *KDM5C* plays a role in neuronal development and

shows highest expression in neuronal tissues. It has been shown that *KDM5C* activity is linked to myeloma[163].

3.2.3.3 Synthetically lethal genes with Entinostat treatment in H1975_Cas9 cells

In order to identify synthetically lethal genes with 100 nM Entinostat in *EGFR* mutant H1975_Cas9 cells (as determined in 3.1.3.2), the readcounts of the Entinostat treatment arm of the screen, as well as the control arm, were normalized to 100 million reads (Figure 30).

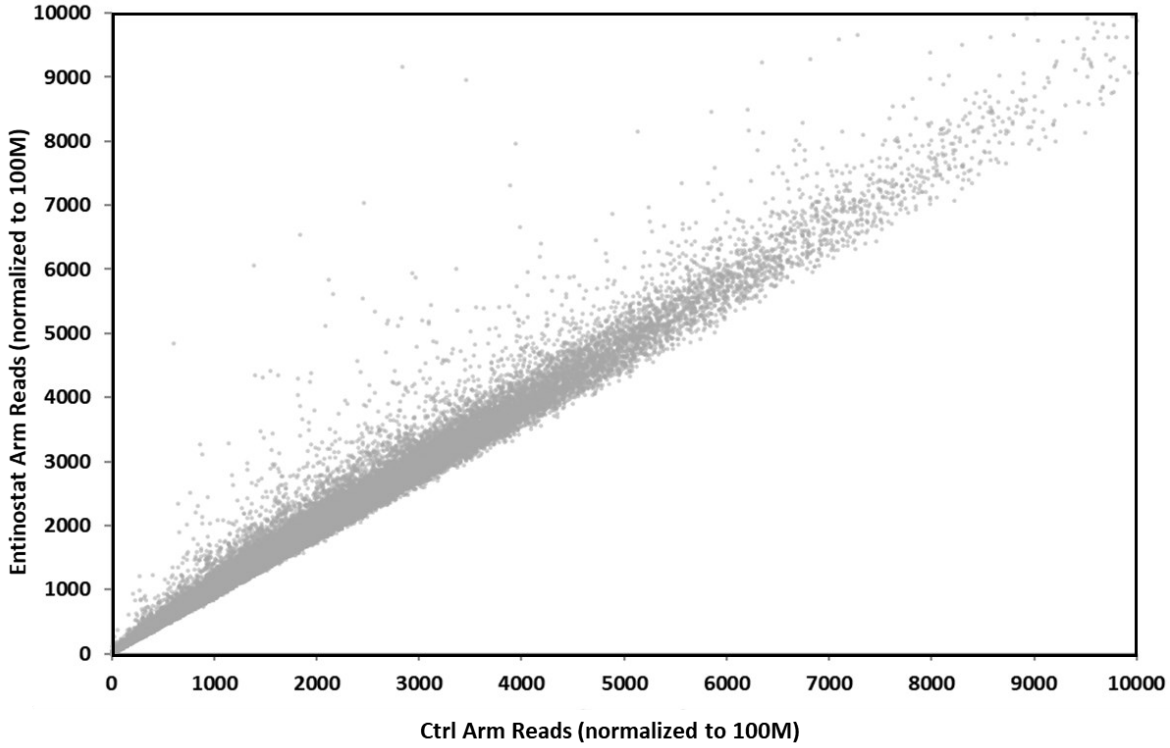


Figure 30 | Scatterplot in H1975_Cas9 cells treated with Entinostat vs. the no-drug arm

The readcounts of both study arms were normalized to 100 million reads. The abundance of all individual sgRNAs in both study arms was compared. Each dot represents one sgRNA. No sgRNAs were identified as possible synthetically lethal target genes with Entinostat.

The read counts were submitted to *CRISPRAnalyzeR,* and genes that were significantly depleted in the DZNep arm vs. the no-drug arm were called by five different algorithms. However, no overlapping hits were identified by all five algorithms. Moreover, by applying a selection criterion for potential synthetically lethal genes (a minimum of 3 sgRNAs with foldchanges of < 0.6), no genes were identified as potential synthetically lethal. Given that the internal positive and negative controls of the library indicated a robust screen, a replicate of the screen would be needed to verify the result. However, at

this time of the study, the emphasis was put on synthetically lethal genes with DZNep in *KRAS* mutant H2030_Cas9 cells.

3.3 Analysis of *CHD8* and *EP300* in the Cancer Genome Atlas dataset

CHD8 and *EP300* were identified as synthetically lethal with DZNep treatment in *KRAS* mutant H2030_Cas9 cells by a pooled CRISPR/Cas9 screen. Among all identified hits for synthetic lethality with DZNep or Entinostat in H2030 cells, *CHD8* and *EP300* showed high confidence levels, as they were detected by five independent algorithms and showed significant p-values. Furthermore, both genes did not show evidence for being generally essential to cell viability. Hence, the study focused on *CHD8* and *EP300* to elucidate their synthetically lethal effects with DZNep.

In a first approach, the publicly available TCGA adenocarcinoma dataset was used to analyze the relation of *CHD8* and *EP300* as well as their involvement in cellular processes in NSCLC cells. For this, the mRNA levels of *CHD8* and *EP300* in 542 patients were correlated (Figure 31).

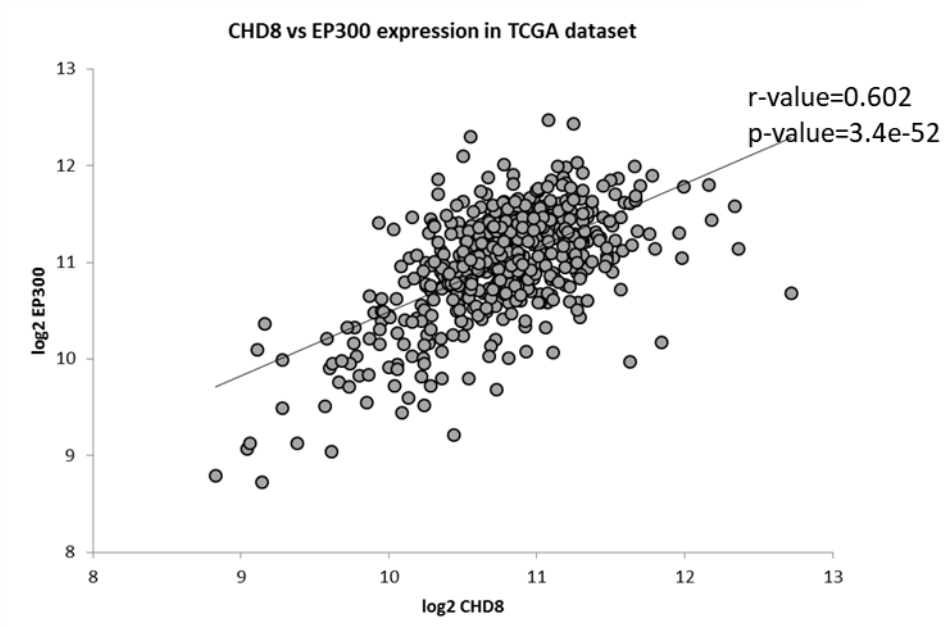


Figure 31 | Scatterplot of *EP300* and *CHD8* expression in the TCGA dataset

The expression of *CHD8* and *EP300* correlated significantly in patients with NSCLC with a r-value of 0.602 and a p-value of 3.4e-52.

The expression of *CHD8* and *EP300* significantly correlated in NSCLC patients with a r-value of 0.602 and a p-value of 3.4×10^{-52}. However, this correlation was not unique

to *KRAS* mutant tumors and neither *CHD8* nor *EP300* expression correlated with the *KRAS* mutation status or *KRAS* expression levels (data not shown). Next, the genes that significantly positively correlated ($p < 0.01$) with either *CHD8* or *EP300* were called and analyzed for overlaps (Figure 32).

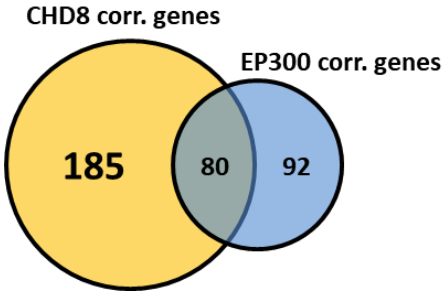


Figure 32 | Venn diagram of overlapping genes whose expression correlated with CHD8 or EP300 in the TCGA dataset

Genes whose expression positively correlated with *CHD8* or *EP300* in the TCGA dataset were called and analyzed for overlaps. 80 genes whose expression correlated with *CHD8* and *EP300* expression levels were identified.

Eighty genes whose expression correlated with *CHD8* and *EP300* expression levels were identified. Noteworthy, these encompassed almost half of all genes correlating with *EP300*. As these results suggested that *CHD8* and *EP300* might share biological functions in NSCLC cells, the *CHD8*-correlating and *EP300*-correlating gene sets were functionally annotated using DAVID v6.8 (Figure 33).

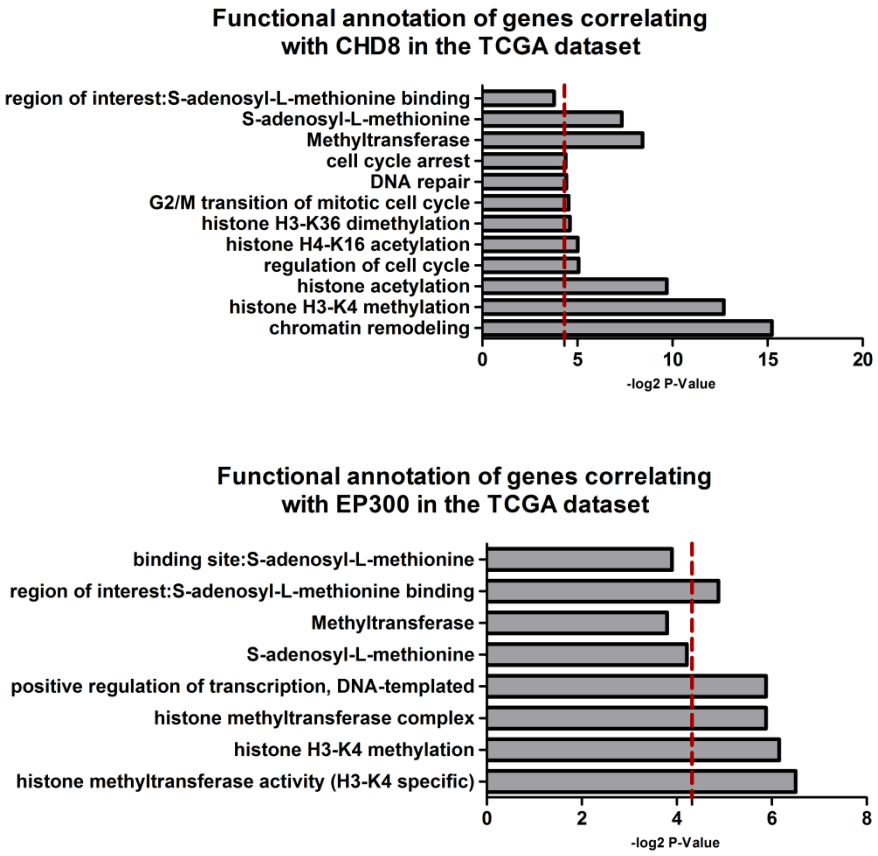


Figure 33 | Functional annotation of overlapping genes whose expression correlated with CHD8 and EP300 in the TCGA dataset

The biological implications of genes whose expression correlated with *CHD8* or *EP300* in the TCGA dataset were analyzed using DAVID v.6.8. Values above -log2 = 4.32 (red line, equivalent to p = 0.05) indicate a significant enrichment of genes. A functional correlation to histone methylation, S-adenosyl-methionine (SAM) and methyltransferase activity was found in both gene sets.

Both genesets show a functional correlation to histone methylation, as *Histone H3-K4 methylation* was found among the highest -log2 p values for both genesets. Further, based on this analysis, S-adenosyl-methionine (SAM) and methyltransferase activity is related to both genesets. Interestingly, DZNep acts as an inhibitor of *S-adenosyl-L-Homocysteine Hydrolase* (*AHCY*), an enzyme that catalyzes the hydrolysis of S-adenosyl-homocysteine (SAH) to adenosine and L-homocysteine. DZNep treatment causes an accumulation of SAH in the cell, and SAH is a strong competitive inhibitor of methyltransferases, which require SAM as a substrate[164,165].

3.4 Analysis of synthetic lethality of DZNep with knockdown of *CHD8* or *EP300* in *KRAS*- and *EGFR*-mutant cell lines

3.4.1 Confirmation of CRISPR/Cas9 screening results with RNA interference-mediated gene knockdown

To validate the results of the CRISPR/Cas9 screen and to confirm that the target genes *CHD8* and *EP300* are synthetically lethal with DZNep independently of the method used for gene depletion, both genes were individually knocked down using 50 nM small interfering RNA (siRNA) and treated with 1 µM DZNep. Cell viability was assessed by staining with CellTiter Blue and normalized to untreated cells (Figure 34).

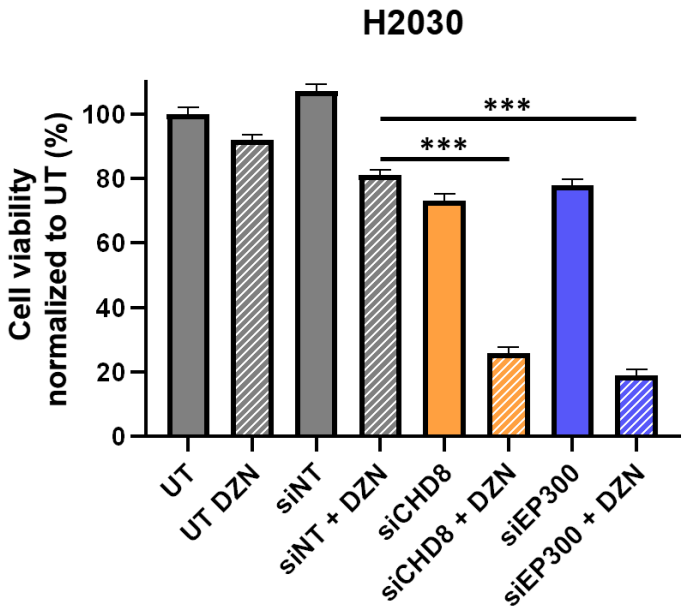


Figure 34 | Measurement of cell viability with CellTiter Blue in H2030 cells after DZNep treatment and knockdown of *CHD8* or *EP300*

H2030 cells were incubated with siRNA for 48 hours and split into a DZNep arm and no-DZNep arm at a ratio of 1:5. After 72 hours the cell viability was measured by CTB and the data was normalized to the untreated (UT) sample. Cells that were treated with DZNep in conjunction with either siCHD8 or siEP300 showed decreased viability compared to cells that were treated with DZNep in conjunction with a non-target control siRNA. Statistical significance of the cell viability changes was evaluated by an unpaired t-test. UT = untreated; DZN = DZNep; siNT = non-target siRNA

As shown in Figure 34, treatment with 1µM DZNep had only a minor effect on cell viability, even in the presence of 50 nM non-target siRNA (siNT). Knockdown of *CHD8* and *EP300* both decreased the fraction of viable cells to 73 % for *CHD8*, and 78 % for *EP300*. Additional treatment with 1µM DZNep decreased the fraction of

viable cells to 26 % for *CHD8* and 19 % for *EP300*. These results were in accordance with the results from the CRISPR/Cas9 screen. Notably, siRNA mediated target gene knockdown and DZNep treatment led to a stronger decrease of cell viability than sgRNA mediated knockout. This could be caused by the duration of the CRISPR/Cas9 screen, which could have allowed some cells to cope with the DZNep treatment. In contrast to Cas9-mediated permanent gene knockout, long incubation times following siRNA-mediated knockdown come with the risk of restored gene expression. As a result, the CTB levels were measured after a considerably shorter incubation time of 72 hours in this siRNA-knockdown experiment.

3.4.2 Cell viability analysis

Following successful siRNA mediated reproduction of synthetic lethality in H2030 cells, other cell lines were tested for synthetic lethality by knockdown of *CHD8* or *EP300* in conjunction with DZNep treatment. The results are shown in Figure 35.

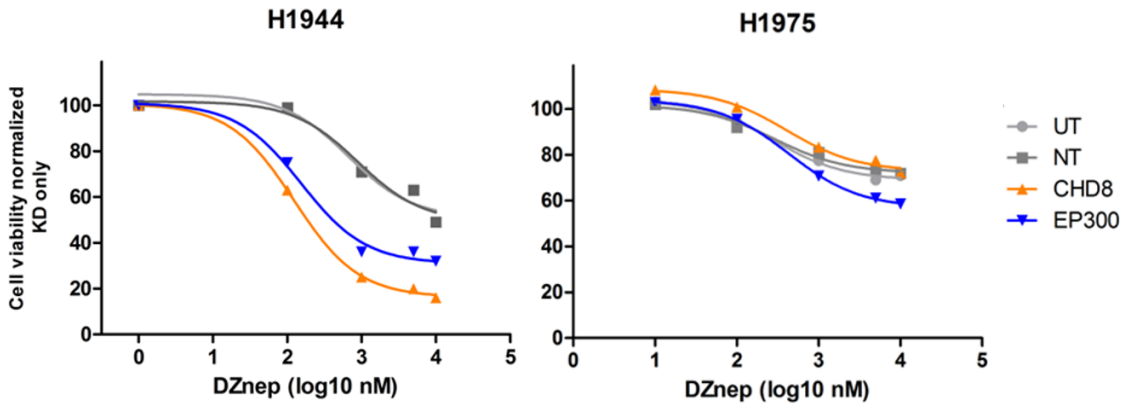


Figure 35 | Measurement of cell viability with CellTiter Blue in H1944 and H1975 cells after treatment with increasing DZNep concentrations and knockdown of *CHD8* or *EP300*

H1944 and H1975 cells were incubated with siRNA for 24 hours, split at a ratio of 1:20 and treated with increasing concentrations of DZNep. After 72 hours the cell viability was measured by CTB and the data were normalized to the untreated cells. In H1944 cells, but not in H1975 cells, the addition of DZNep to si*CHD8* pre-treated cells enhanced cell death compared to DZNep treatment alone. For si*EP300*, the synthetic effect was reduced compared to si*CHD8* in H1944 cells, yet still detectable. In H1975 cells, knockdown of *EP300* led to a slightly stronger response to DZNep. UT = untreated cells; NT = cells treated with non-target siRNA

In H1944 cells, the addition of DZNep to si*CHD8* or si*EP300* pre-treated cells enhanced cell death compared to DZNep treatment alone, and a minor synthetic effect could be observed in *EGFR* mutant H1975 cells for si*EP300*. Of note, H1944 cells harbor a different *KRAS* mutation (p.G13D) than H2030 cells (p.G12C), suggesting that

the observed synthetic lethality does not depend on the *KRAS* mutation subtype for those two cases.

In order to compare the response of H1944 and H1975 cells to H2030 cells at a DZNep concentration of 1 µM (which was used in the DZNep screen with H2030 cells), the cell viability after gene knockdown in conjunction with DZNep treatment was normalized to the corresponding sample that solely received the gene knockdown, thereby illustrating the synthetically lethal effect of the gene knockdown with DZNep treatment (Figure 36).

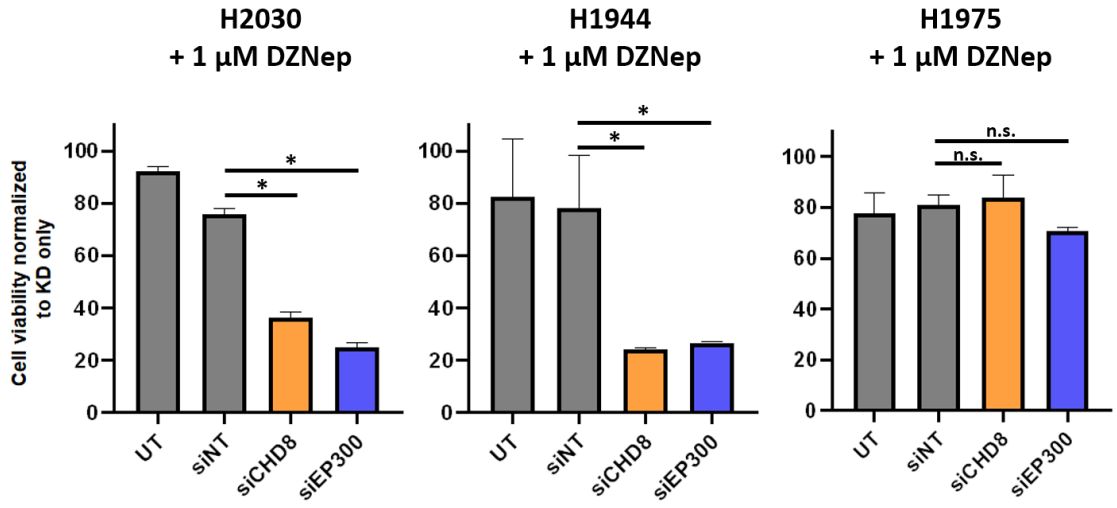


Figure 36 | Comparison of cell viability upon 1 µM DZNep treatment in conjunction with gene knockdown to gene knockdown without additional DZNep treatment

Cells were incubated with siRNA for 48 hours and split into a DZNep arm and no-DZNep arm at a ratio of 1:5. After 72 hours the cell viability was measured by CTB. Bars indicate the percentage of viable cells upon gene knockdown in conjunction with 1 µM DZNep treatment normalized to the respective knockdown without additional DZNep. A knockdown of *CHD8* or *EP300* resulted in decreased cell viability in H2030 and H1944 cells, but not in H1975 cells. Statistical significance of the cell viability changes was evaluated by an unpaired t-test. UT = Cells not treated with any siRNA; siNT = cells treated with non-target siRNA

As demonstrated in Figure 36, 1 µM DZNep reduced the cell viability to 80 % in H1944 cells, and an additional knockdown of *CHD8* further decreased the fraction of viable cells to 24 %. For *EP300*, similar results were seen, with a reduction of viable cells to 27 %. Likewise, an additional knockdown of *CHD8* or *EP300* resulted in decreased viability of H2030 cells, with 36 % (si*CHD8* + DZNep) and 25 % (si*EP300* + DZNep) of viable cells compared to 92 % with DZNep treatment alone. In *EGFR*-mutant H1975 cells, an additional knockdown of neither *CHD8* nor *EP300* resulted in significantly decreased cell viability. Taken together, the results show that synthetic lethality of

DZNep in conjunction with the depletion of *CHD8* or *EP300* could be confirmed in another *KRAS*-mutant, but not in an *EGFR*-mutant cell line.

3.4.3 Analysis of apoptosis by flow cytometry

As seen in the previous section, the Cell-Titer Blue assay provided a useful approach for monitoring the cell viability after gene knockdown in conjunction with DZNep treatment. However, this assay only measures the metabolic capacity of the cells and does not provide information on whether a decreased metabolic capacity is due to elevated apoptosis. Consequently, to determine the fraction of cells undergoing apoptosis, H2030 cells were treated with either DZNep, knockdown of *CHD8*, likewise *EP300*, or both combined, and apoptosis was analyzed by flow cytometry. Instead of using various concentrations of DZNep, 1 µM was used as this concentration has been applied during the screening. A flow cytometry-based analysis of apoptosis due to gene knockdown in conjunction with DZNep treatment in H2030 cells is shown in Figure 37.

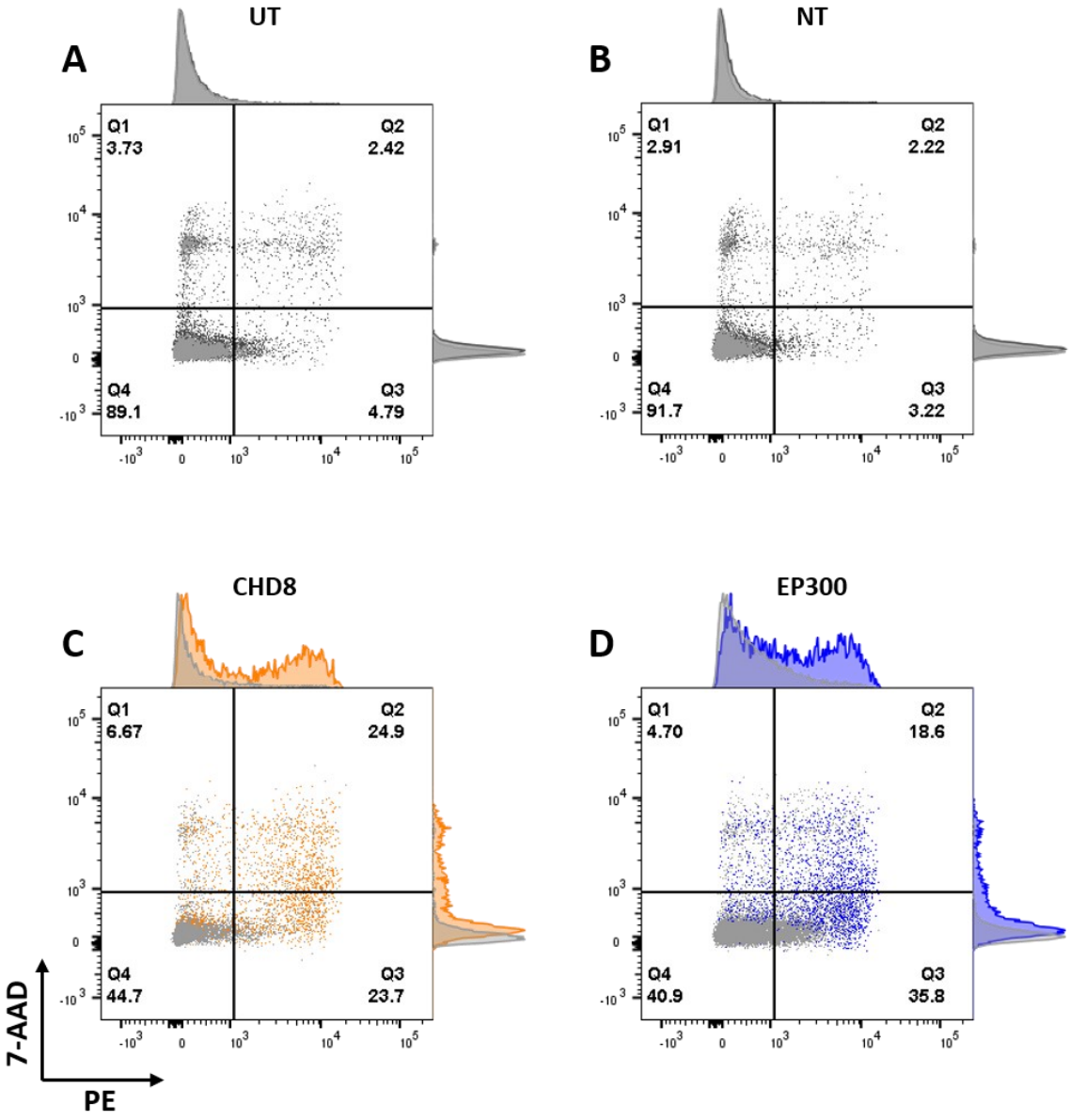


Figure 37 | Flow cytometry-based analysis of apoptosis in H2030 cells

H2030 cells were subjected to either DZNep, knockdown of *CHD8*, *EP300*, or combined treatment. Apoptotic cells were identified through staining with AnnexinV coupled to Phycoerythrin dye (PE) and 7-Aminoactinomycin D (7-AAD). Viable cells appear in Q4, early apoptotic cells in Q3, late apoptotic cells in Q2, and dead cells in Q1. (A) Cells that were not pre-treated with siRNA (UT) showed no difference between and DZNep treated cells (black) and non-DZNep treated cells (light grey). (B) Cells that were pre-treated with a non-target siRNA (NT) showed no difference between and DZNep treated cells (black) and non-DZNep treated cells (light grey). (C) Cells that were pre-treated with siRNA targeting *CHD8* showed elevated levels of early and late-stage apoptosis (Q2, Q3) in DZNep-treated (orange) vs. non-DZNep treated cells (light grey). (D) Cells that were pre-treated with siRNA targeting *EP300* showed elevated levels of apoptosis (Q2, Q3) in DZNep-treated (blue) vs. non-DZNep treated cells (light grey), despite slightly increased apoptosis levels (Q3) also in non-drug treated cells (light grey).

Treatment with 1µM DZNep did not increase apoptosis or cell death (Figure 37A). Similarly, treatment with non-target siRNA did not increase the fraction of apoptotic

cells (Figure 37B). H2030 cells that were treated with si*CHD8* (Figure 37C) or si*EP300* (Figure 37D) showed slightly increased signals of PE or 7-AAD, indicating a slightly increased rate of apoptotic cells. However, when 1 µM DZNep treatment was combined with knockdown of *CHD8* (Figure 37C, orange) or *EP300* (Figure 37D, blue), the frequency of viable cells (Q4) decreased from 91.7 % (siNT plus DZNep, Figure 37B) to 44.7 % (si*CHD8* + DZNep, Figure 37C) or 40.9 % (*EP300* + DZNep, Figure 37D). Concomitantly, the fraction of apoptotic, late apoptotic or dead cells increased from 8.3 % (siNT plus DZNep, as indicated by the sum of Q2, Q3 and Q4 of Figure 37B) to 55.3 % (si*CHD8* + DZNep, as indicated by the sum of Q2, Q3 and Q4 of Figure 37C) or 59.1 % (*EP300* + DZNep, as indicated by the sum of Q2, Q3 and Q4 of Figure 37D).

Next, *KRAS*-mutant H1944 cells were treated with 1 µM DZNep in conjunction with a knockdown of *CHD8* or *EP300* to determine if apoptosis was also elevated similarly to H2030 cells (Figure 38).

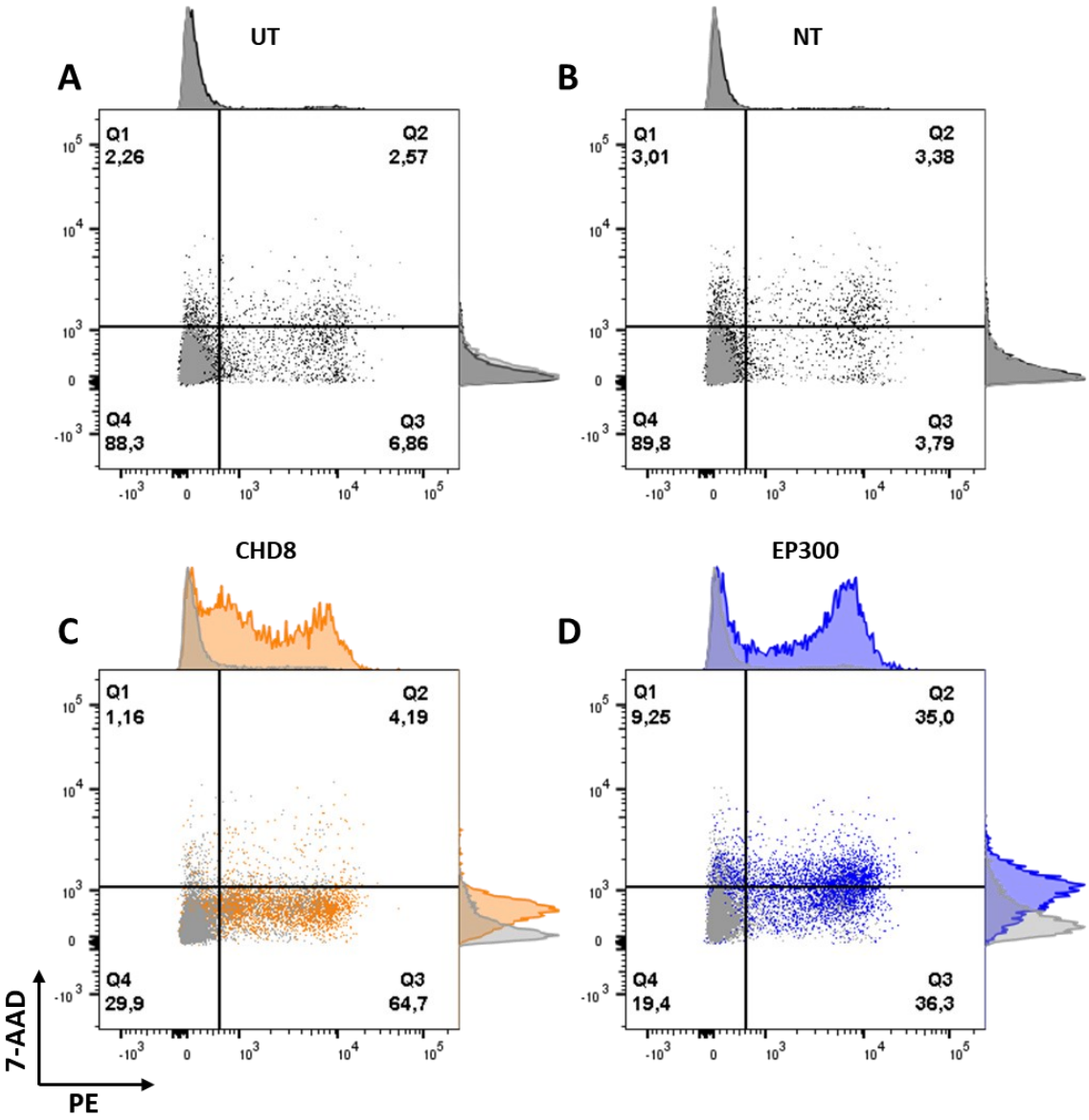


Figure 38 | Flow cytometry-based analysis of apoptosis in H1944 cells

H1944 cells were subjected to either DZNep, knockdown of *CHD8*, *EP300*, or combined treatment. Apoptotic cells were identified through staining with AnnexinV coupled to Phycoerythrin dye (PE) and 7-Aminoactinomycin D (7-AAD). Viable cells appear in Q4, early apoptotic cells in Q3, late apoptotic cells in Q2, and dead cells in Q1. (A) Cells that were not pre-treated with siRNA (UT) showed no difference between and DZNep treated cells (black) and non-DZNep treated cells (light grey). (B) Cells that were pre-treated with a non-target siRNA (NT) showed no difference between and DZNep treated cells (black) and non-DZNep treated cells (light grey). (C) Cells that were pre-treated with siRNA targeting *CHD8* showed elevated levels of early apoptosis (Q3) in DZNep-treated (orange) vs. non-DZNep treated cells (light grey). (D) Cells that were pre-treated with siRNA targeting *EP300* showed elevated levels of apoptosis (Q2, Q3) in DZNep-treated (blue) vs. non-DZNep treated cells (light grey).

As expected from the decreased CTB signal in H1944 cells upon combined treatment with *CHD8* or *EP300* knockdown with DZNep (Figure 35, page 84), the frequency of viable cells decreased from 89.8 % (siNT plus DZNep, as shown in Q1 of Figure 38B) to 29.9 % (si*CHD8* + DZNep, as shown in Q1 of Figure 38C) or 19.4 % (*EP300* +

DZNep, as shown in Q1 of Figure 38D), respectively. The frequency of apoptotic, late apoptotic or dead cells increased from 10.1 % (siNT plus DZNep, as indicated by the sum of Q2, Q3 and Q4 of Figure 38B) to 70.1 % (si*CHD8* + DZNep, as indicated by the sum of Q2, Q3 and Q4 of Figure 38C) or 80.6 % (*EP300* + DZNep, as indicated by the sum of Q2, Q3 and Q4 of Figure 38D).

Finally, *EGFR*-mutant H1975 cells were treated with 1 µM DZNep in conjunction with a knockdown of *CHD8* or *EP300* to determine if apoptosis would also be elevated in this cell line (Figure 39).

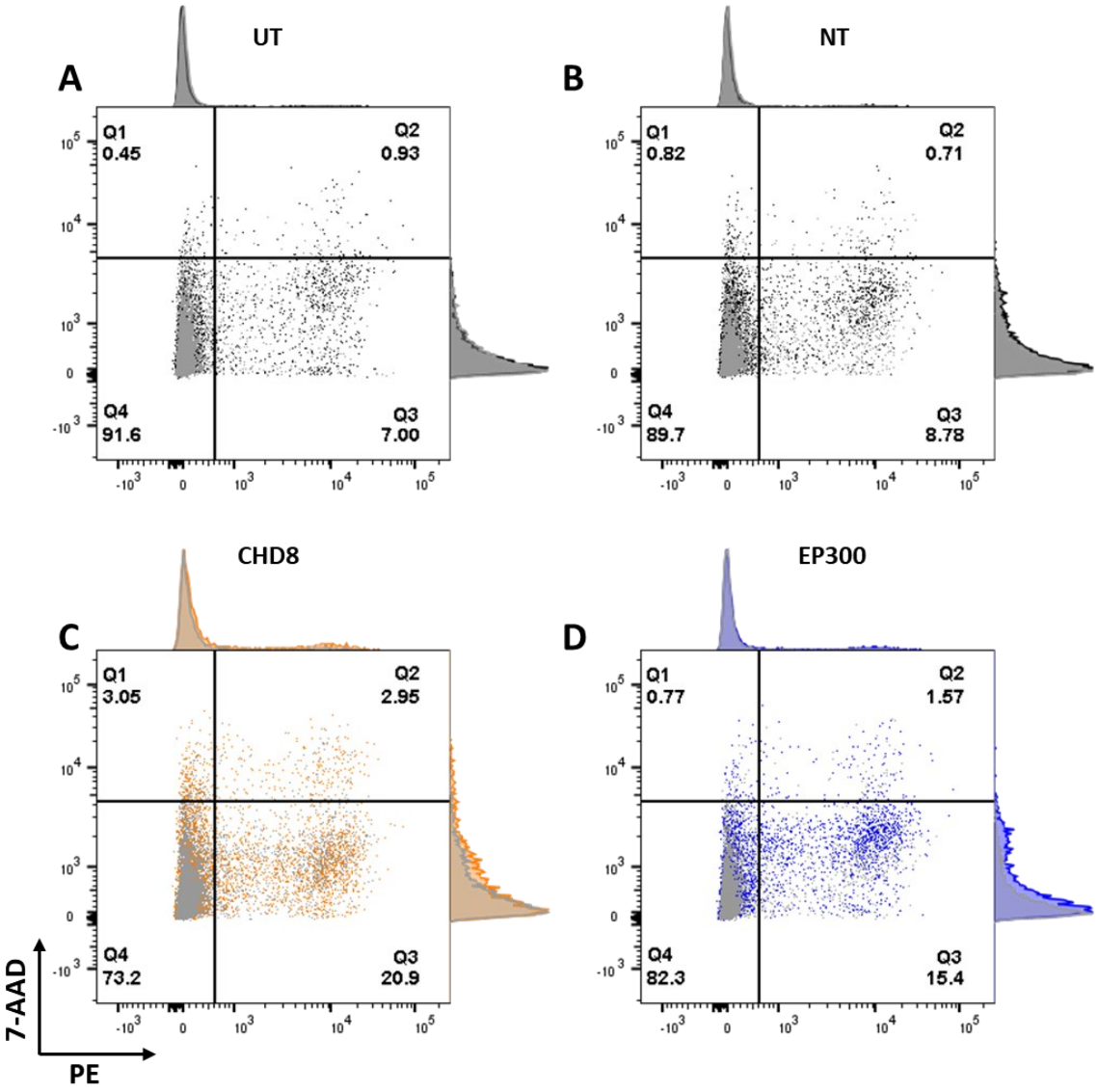


Figure 39 | Flow cytometry-based analysis of apoptosis in H1975 cells

H1975 cells were subjected to either DZNep, knockdown of *CHD8*, *EP300*, or combined treatment. Apoptotic cells were identified through staining with AnnexinV coupled to Phycoerythrin dye (PE) and 7-Aminoactinomycin D (7-AAD). Viable cells appear in Q4, early apoptotic cells in Q3, late apoptotic cells in Q2, and dead cells in Q1. (A) Cells that were not pre-treated with siRNA (UT) showed no difference between and DZNep treated cells (black) and non-DZNep treated cells (light grey). (B) Cells that were pre-treated with a non-target siRNA (NT) showed no difference between and DZNep treated cells (black) and non-DZNep treated cells (light grey). (C) Cells that were pre-treated with siRNA targeting *CHD8* showed slightly elevated levels of apoptosis (Q2, Q3) in DZNep-treated (orange) vs. non-DZNep treated cells (light grey). (D) Cells that were pre-treated with siRNA targeting *EP300* showed slightly elevated levels of apoptosis (Q2, Q3) in DZNep-treated (blue) vs. non-DZNep treated cells (light grey).

H1975 cells showed a slightly increased frequency of apoptotic cells after treatment with 1µM DZNep and parallel knockdown of *CHD8* (Figure 39C) or *EP300* (Figure 39D); however, when compared to H2030 and H1944 cells, this increase was remarkably lower. The frequency of viable cells decreased from 89.7 % (siNT plus

DZNep, as shown in Q1 of Figure 39B) to 73.2 % (si*CHD8* + DZNep, as shown in Q1 of Figure 39C) or 82.3 % (*EP300* + DZNep, as shown in Q1 of Figure 39D), respectively.

Taken together, cell viability was decreased by combinational treatment of *KRAS* mutant H2030 and H1944 cells, but not *EGFR* mutant H1975 cells with 1 µM DZNep and parallel knockdown of *CHD8* or *EP300*. In the same manner, flow cytometry confirmed increased apoptosis in *KRAS* mutant H2030 and H1944 cells upon double treatment with 1µM DZNep and parallel knockdown of *CHD8* or *EP300*. The experiments also confirmed that synthetic lethality of DZNep and loss of *CHD8* or *EP300* does not occur in *EGFR* mutant H1975 cells.

3.5 Gene expression profiling of H2030 cells after target gene knockdown, DZNep treatment, and combined treatment

In order to assess the functional implications of either DZNep treatment, knockdown of *CHD8* or combined treatment in H2030 cells, global changes in gene expression on the mRNA level were measured using a HumanHT-12 Expression BeadChip Kit (Illumina, Inc.) microarray. The mean probe values of si*CHD8*, si*EP300*, si*CHD8* + DZNep and si*EP300* + DZNep samples were normalized to the non-target control siRNA sample and deregulated genes were called with foldchanges of > 1.5 for upregulated and < 0.66 for downregulated genes.

In cells treated with DZNep only, 636 genes were upregulated, and 650 genes were downregulated. Following knockout of *CHD8* or *EP300*, 386 or 603 genes were up- and 433 or 649 genes were downregulated, respectively. Combined treatment of DZNep with knockdown of yielded 1052 (*CHD8*) or 2190 (*EP300*) upregulated and 728 (*CHD8*), or 2175 (*EP300*) downregulated genes.

First, the overlap of deregulated genes upon knockdown of *CHD8* or *EP300* was determined (Figure 40).

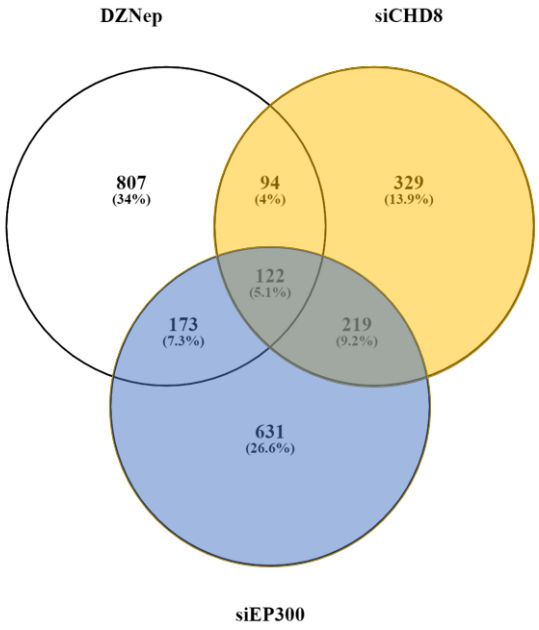


Figure 40 | Venn diagram illustrating the overlap of deregulated genes upon DZNep treatment, knockdown of *CHD8*, and knockdown of *EP300*

Deregulated genes following DZNep treatment, knockdown of *CHD8* (orange), and knockdown of *EP300* (blue) were identified by a microarray. Percentages indicate the respective shares of all deregulated genes in all samples. 44.63 % of genes that were deregulated upon *CHD8* knockdown were also deregulated upon *EP300* knockdown, and 29.78 % of genes were deregulated vice versa.

Because *CHD8* and *EP300* expression correlated in the TCGA dataset (Figure 31) and a large overlap of *CHD8* and *EP300* related gene sets was observed (Figure 32), a large overlap of deregulated genes was expected. Indeed, 44.63 % of all genes that were deregulated upon *CHD8* knockdown were also deregulated upon *EP300* knockdown. Accordingly, 29.78 % of genes that were deregulated upon *EP300* knockdown were also deregulated after knockdown of *CHD8*. This result is in good agreement with the previously shown TCGA dataset findings, corroborating a potential overlap of biological functions or pathways among *CHD8* and *EP300* in lung cancer cells.

Next, genes that were uniquely deregulated after the knockdown of *CHD8* or *EP300* in conjunction with DZNep treatment were quantified. For this purpose, the overlap of deregulated genes after DZNep treatment, si*CHD8* / si*EP300*, or combined treatment was calculated (Figure 41).

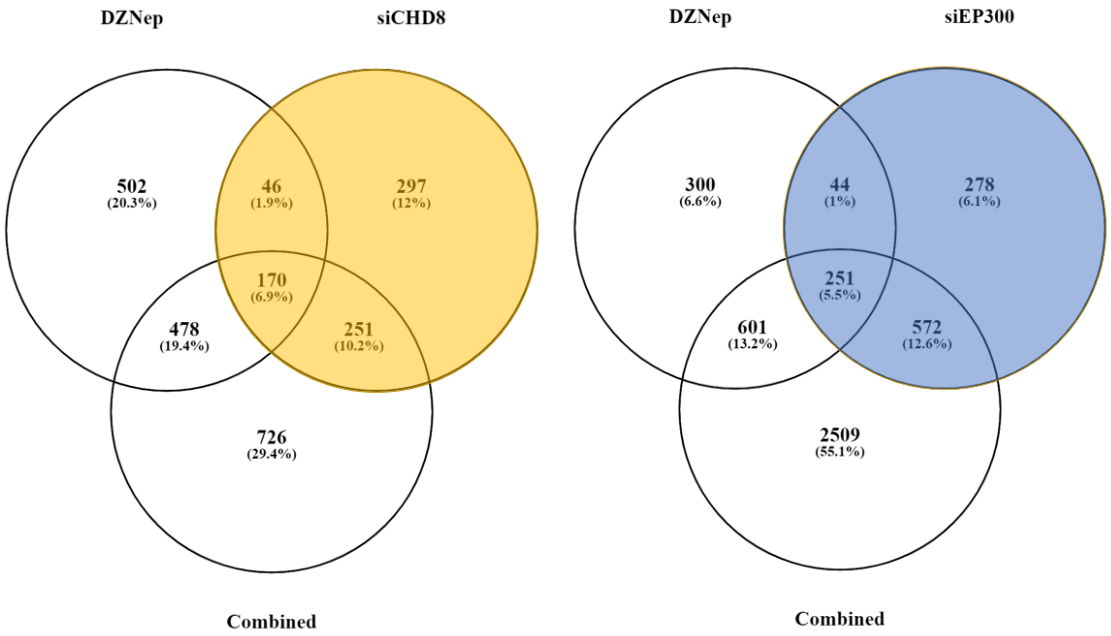


Figure 41 | Venn diagram illustrating the overlap of deregulated genes upon DZNep treatment, knockdown of *CHD8* (orange, left) or *EP300* (blue, right), and combined treatment

Deregulated genes following DZNep treatment, knockdown of *CHD8* (orange), knockdown of *EP300* (blue), and combined treatment were identified by a Microarray. Percentages indicate the respective share of all deregulated genes in all samples. 726 genes were uniquely deregulated by combined DZNep treatment and *CHD8* knockdown, and 2509 genes were uniquely deregulated with combined knockdown of *EP300* and DZNep treatment.

As seen in Figure 41, 726 genes were uniquely deregulated by combined DZNep treatment and *CHD8* knockdown (orange, left), encompassing 44 % of all genes that were deregulated by the combined treatment. Similarly, 2509 genes were uniquely deregulated with combined knockdown of *EP300* and DZNep treatment (blue, right), accounting for 63.79 % of genes that were deregulated upon double treatment and 55.10 % of all genes either deregulated by DZNep or knockdown of *EP300* or combined treatment.

These high percentages suggested that, upon double treatment, gene regulation was affected in a way that led to up- and downregulation of genes that would not be affected by either treatment alone, supporting evidence for synthetic lethality.

Finally, the overlap of genes that were exclusively deregulated upon combined DZNep treatment with either knockdown of *CHD8* or *EP300* was calculated (Figure 42).

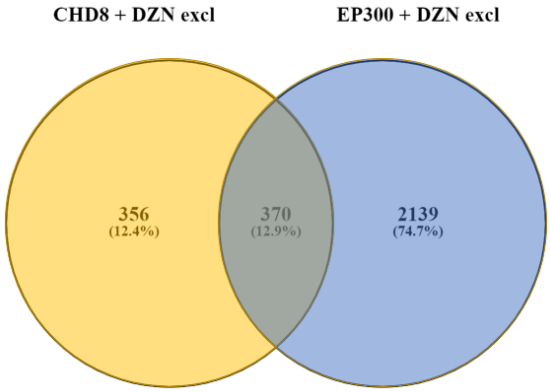


Figure 42 | Venn diagram illustrating the overlap of exclusively deregulated genes upon combined DZNep treatment and knockdown of either *CHD8* (orange, left) or *EP300* (blue, right)

Genes that are only deregulated upon DZNep treatment in conjunction with a knockdown of *CHD8* (orange), or a knockdown of *EP300* (blue) were identified by a microarray. Percentages indicate the respective share of all deregulated genes in all samples. 50.9 % of genes that were only deregulated upon DZNep treatment in conjunction with a knockdown of *CHD8* were also deregulated by a knockdown of *EP300* in conjunction with DZNep treatment, but not by DZNep treatment or gene knockdown alone.

More than half of all genes (50.9 %) that were deregulated upon combined DZNep treatment and knockdown of *CHD8*, but not upon DZNep or knockdown of *CHD8* alone, were also deregulated by combined DZNep treatment and knockdown of *EP300*, but not *EP300* alone.

Taken together, genes that were deregulated by either *CHD8* or *EP300* exhibited a large overlap. Furthermore, a large proportion of deregulated genes was exclusively deregulated by combined but not single treatment with either DZNep or knockdown. Also, genes that were exclusively deregulated by combined DZNep treatment with *CHD8* or *EP300* knockdown exhibited a large overlap, further corroborating the hypothesis that *CHD8* and *EP300* may have common biological functions in NSCLC cells.

3.5.1 Functional annotation of deregulated genes

3.5.1.1 Analysis of Biocarta, KEGG, and Gene Ontology annotations

In order to understand the biological implications of deregulated genes, DAVID was used for functional annotation. Genes that fulfilled the foldchange criteria listed above were submitted to DAVID, and annotation clusters were set to include Biocarta pathways, KEGG pathways and Gene Ontology (GO) annotations with a p-value < 0.05. In order to compare the results among samples that were either DZNep treated, received

a knockdown of *CHD8* or *EP300* or received a combined DZNep and gene knockout, the top cluster identified by DAVID and terms related to cell division and cell cycle were considered. According to these criteria, the functional annotation of deregulated genes in H2030 cells treated with 1 µM DZNep is shown below (Table 28).

Table 28 | Functional annotation of gene expression profiling following 1 µM DZNep treatment of H2030 cells. Deregulated genes were selected using cutoff values of FC > 1.5 for upregulation and FC < 0.66 for downregulation. The DAVID top clusters and cell division or cell cycle-related annotations were selected.

Annotation	Gene count	P-Value
Top cluster (Enrichment Score: 5.68)		
Ribosomal protein	22	2.20E-07
Translation	27	5.00E-07
SRP-dependent co-translational protein targeting to membrane	16	5.70E-07
Viral transcription	17	1.10E-06
Nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	16	1.20E-05
Translational initiation	16	6.20E-05
RRNA processing	20	1.30E-04
Cell division and cell cycle-related annotations		
Cell division	34	2.10E-04
Cell cycle	49	2.60E-04
Mitosis	25	5.10E-04

As seen in Table 28, altered genes were mainly annotated for ribosomal functions and translation, presenting an enrichment score of 5.68. The group enrichment scores intend to put the relative importance of the gene groups in order. A higher score indicates a higher confidence level that group members are involved in important biological functions in the group. Group scores are compared to each other rather than put on an absolute scale[166]. Interestingly, 34 genes involved in cell division, 49 genes related to cell cycle, and 25 genes related to mitosis were identified among the deregulated genes, although H2030 cells treated with DZNep showed almost no reduction in cell viability.

The functional annotation of deregulated genes in H2030 cells that received a knockdown of *CHD8* is shown in Table 29.

Table 29 | Functional annotation of gene expression profiling following knockdown of *CHD8* in H2030 cells. Deregulated genes were selected using cutoff values of FC > 1.5 for upregulation and FC < 0.66 for downregulation. The DAVID top cluster and cell division or cell cycle-related annotations were selected.

Annotation	Gene count	P-Value
Top cluster (Enrichment Score: 4.63)		
Ribosomal protein	22	2.20E-07
Translation	27	5.00E-07
SRP-dependent co-translational protein targeting to membrane	16	5.70E-07
Viral transcription	17	1.10E-06
Nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	16	1.20E-05
Ribonucleoprotein	25	1.30E-05
Translational initiation	16	6.20E-05
rRNA processing	20	1.30E-04
Cell division and cell cycle-related annotations		
Cell cycle	36	8.00E-04
Mitosis	19	1.10E-03
Cell division	23	3.90E-03

Genes that were altered in H2030 cells that received a knockdown of *CHD8* were mainly related to ribosomal functions, highly similar to DZNep treated cells. Also, similarly to DZNep treated cells, genes related to cell cycle (36 genes), mitosis (19 genes), and cell division (23 genes) were identified among the genes that were deregulated upon *CHD8* knockdown while showing 20 % decreased cell viability upon *CHD8* knockdown (Figure 34).

The functional annotation of deregulated genes in H2030 cells that received a knockdown of *EP300* is shown in Table 30.

Table 30 | Functional annotation of gene expression profiling following knockdown of *EP300* in H2030 cells. Deregulated genes were selected using cutoff values of FC > 1.5 for upregulation and FC < 0.66 for downregulation. The DAVID top cluster and cell division or cell cycle-related annotations were selected.

Annotation	Gene count	P-Value
Top cluster (Enrichment Score: 2.69)		
Nucleotide phosphate-binding region: NADP	12	6.10E-04
Oxidoreductase activity	23	6.40E-04
NADP	19	2.20E-03
Glucose dehydrogenase	9	6.00E-03
Short-chain dehydrogenase/reductase, conserved site	6	3.20E-02
Cell division and cell cycle-related annotations		
(None)	N/A	N/A

In this setting, altered genes were mainly related to *Nicotinamide adenine dinucleotide phosphate* (*NADP*), but not to ribosomal processes. Noteworthy, no annotations for cell cycle, mitosis, or cell division were found for genes that were deregulated upon knockdown of *EP300*, and the enrichment score for the top cluster was rather low compared to the other top cluster scores in this experiment.

The functional annotation of deregulated genes that were treated with 1 µM DZNep in conjunction with a knockdown of *CHD8* is shown in Table 31.

Table 31 | Functional annotation of gene expression profiling following 1 µM DZNep treatment in conjunction with a knockdown of *CHD8* in H2030 cells. Deregulated genes were selected using cutoff values of FC > 1.5 for upregulation and FC < 0.66 for downregulation. The DAVID top cluster and cell division or cell cycle-related annotations were selected.

Annotation	Gene count	P-Value
Top cluster (Enrichment Score: 11.34)		
Cell cycle	83	1.4E-11
Mitosis	42	4.7E-9
Cell division	51	7.4E-8
Mitotic nuclear division	35	1.8E-5
Cell division and cell cycle-related annotations		
Sister chromatid cohesion	29	8.40E-06
Centromere	31	5.30E-05
Kinetochore	21	2.50E-03
Microtubule	49	3.00E-04
DNA replication	22	5.00E-04
DNA strand elongation involved in DNA replication	7	4.40E-03
Telomere maintenance via recombination	10	7.90E-03
Nucleotide-excision repair, DNA gap filling	8	1.50E-02
Cyclin	10	1.90E-02

As seen in Table 31, remarkably, the top cluster consisted of annotations related to cell cycle (83 genes), mitosis (42 genes), cell division (51 genes), and mitotic nuclear division (35 genes). Furthermore, the enrichment score for the top cluster was the highest in the cells that were treated with DZNep in conjunction with a knockdown of *CHD8*. Further annotations that were linked to cell division were significantly enriched in the group of deregulated genes. The increased number of genes, as well as the lower p-values compared to cells that were treated with either DZNep or received a knockdown of *CHD8* were in accordance with the previously shown effects of decreased cell viability and enhanced apoptosis upon a combined treatment, thereby further corroborating the previous findings.

The functional annotation of deregulated genes that were treated with 1 µM DZNep in conjunction with a knockdown of *EP300* is shown in Table 32.

Table 32 | Functional annotation of gene expression profiling following 1 µM DZNep treatment in conjunction with a knockdown of *EP300* in H2030 cells. Deregulated genes were selected using cutoff values of FC > 1.5 for upregulation and FC < 0.66 for downregulation. The DAVID top cluster and cell division or cell cycle-related annotations were selected.

Annotation	Gene count	P-Value
Top cluster (Enrichment Score: 6.11)		
Mitochondrion	107	3.30E-07
Transit peptide	59	3.50E-06
Mitochondrial matrix	43	6.30E-06
Cell division and cell cycle-related annotations		
Cell cycle	63	8.00E-05
Cell division	41	3.00E-04
Mitosis	28	2.70E-03
DNA replication	11	2.50E-04
Telomere maintenance via recombination	9	8.10E-04
Nucleotide-excision repair, DNA gap filling	7	3.60E-03
DNA damage response, detection of DNA damage	8	7.60E-03

As shown in Table 32, combined DZNep treatment and knockdown of *EP300* resulted in a top cluster of mitochondria-related genes. Since a knockdown of *EP300* without additional DZNep treatment did not result in any cell-cycle related terms, it is noteworthy that with additional DZNep treatment, 63 cell-cycle related genes and cell division related genes (41 genes) were significantly deregulated. This result is in line with the previous results of decreased cell viability and increased apoptosis upon the combined treatment. However, DZNep treatment alone resulted in 49 deregulated cell-

cycle related genes, and the difference of additional cell-cycle related genes for the DZNep + *EP300* knockdown condition to DZNep alone was only 28 % compared to 69 % for the DZNep + *CHD8* knockdown condition.

Taken together, overlapping deregulated genes upon the knockdown of *CHD8* or *EP300* further corroborated the finding of overlapping *CHD8* and *EP300*-related gene sets in the TCGA data. The functional annotation of cell-cycle related genes comprised decreased cell viability and increased apoptosis rates in cells that were treated with DZNep in conjunction with a knockdown of *CHD8* or *EP300*, thereby underlining the synthetically lethal impact of the combined treatment.

3.5.1.2 Analysis of canonical signaling pathways

To elucidate canonical signaling pathway implications of the deregulated genes, the dataset was analyzed using the Ingenuity Pathways Analysis (IPA) software suite. First, canonical pathways were identified in DZNep-treated H2030 cells compared to *CHD8*- or *EP300* depleted cells (Figure 43).

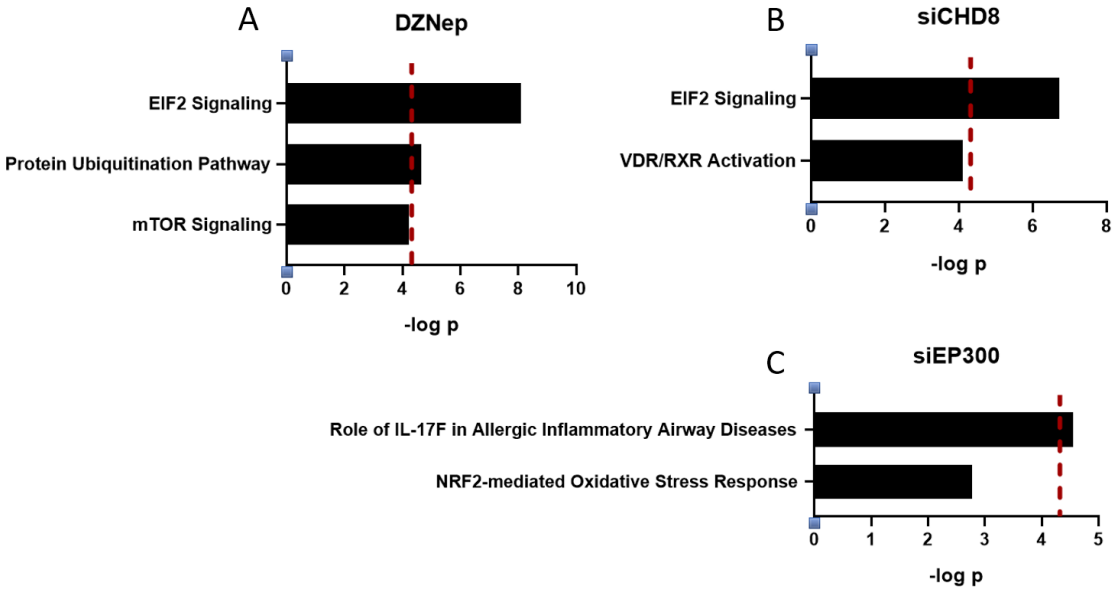


Figure 43 | Ingenuity Pathway Analysis of deregulated genes after DZNep treatment or knockdown of *CHD8* or *EP300*.

Deregulated canonical pathways in H2030_Cas9 cells that were treated with DZNep or a knockdown of either *CHD8* or *EP300* were annotated using Ingenuity Pathway Analysis. -log p-values larger than 4.32 (red line, equivalent to p = 0.05) indicate a significant gene enrichment.

As seen in Figure 43, genes that were deregulated by both DZNep (A) and siCHD8 (B) treatment were significantly enriched for EIF2 signaling. Interestingly, eIF2β, a subunit

of translation-initiation factor EIF2, has been identified as a potential therapeutic target in the *KRAS*-mutant lung cancer cell lines H358 (*KRAS* G12C) and H460 (*KRAS* Q61H)[167]. However, no effects on cell viability by DZNep treatment or knockdown of *CHD8* could be observed in the present study using H2030 cells. Other affected canonical pathways included `Protein Ubiquitination´ and `mTOR Signaling´ in DZNep-treated cells, and activation of the vitamin D receptor/retinoid X Receptor complex (VDR/RXR) in CHD8-diminished cells, with the latter two showing a trend towards enrichment.

Deregulated genes in cells in which *EP300* expression levels were knocked down (C) showed a significant enrichment of `IL-17F in Allergic Inflammatory Airway Diseases´ and a trend towards enrichment of `NRF2-mediated Oxidative Stress Response´. Of note, EP300 has been described as a key member of the KEAP1-NRF2-ARE survival pathway[168], suggesting a functional link of deregulated genes upon a knockdown of *EP300* to this canonical pathway. Of note, H2030 cells carry a mutation in *KEAP1*[143], and a subsequent *in silico* analysis of *KEAP1* mutations in 32 NSCLC cell lines revealed a significant correlation of *KEAP1* and *KRAS* mutations (Supplementary figure 6 and Supplementary table 3).

Next, affected canonical pathways were identified in DZNep-treated H2030 cells in conjunction with a knockdown of either *CHD8* or *EP300* (Figure 44).

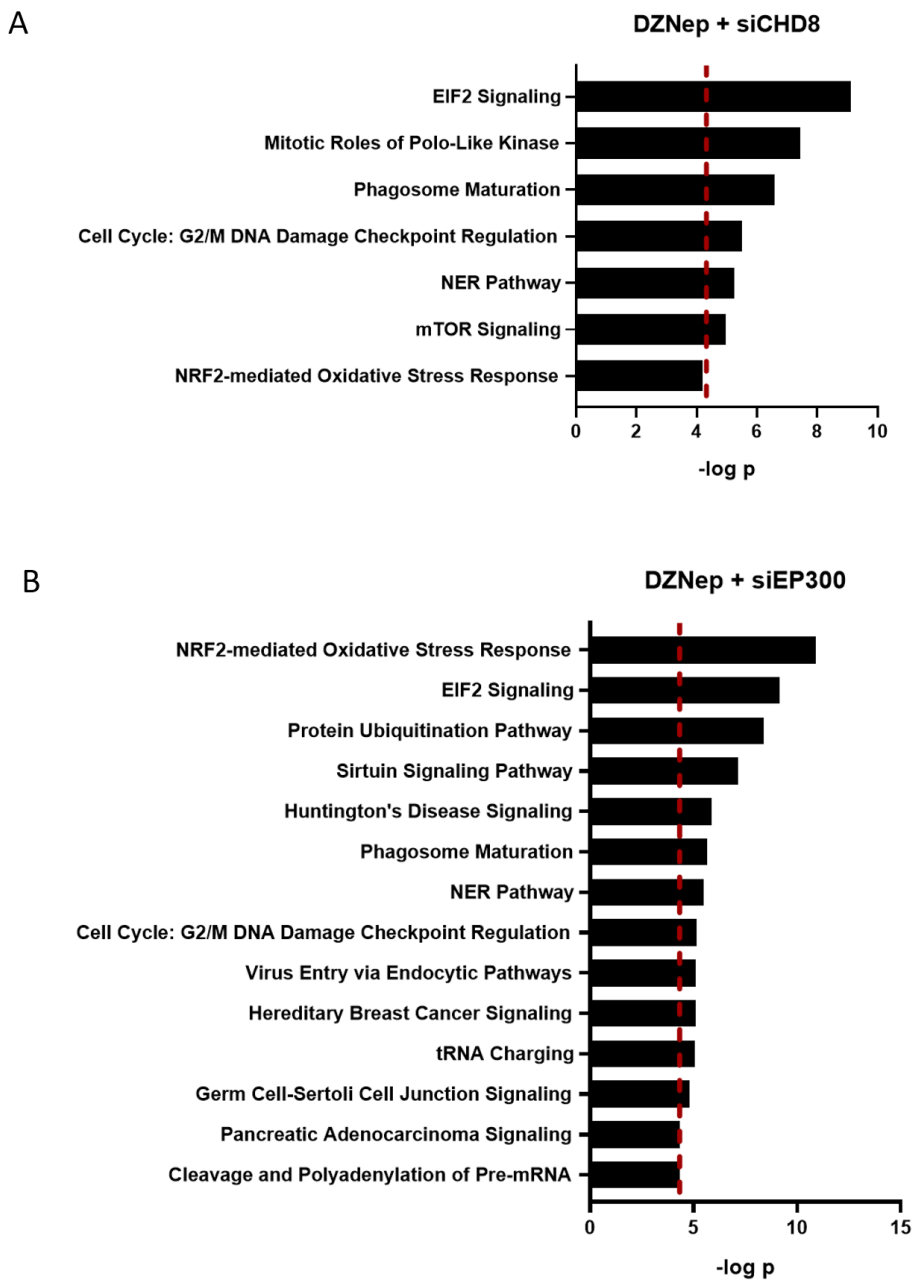


Figure 44 | Ingenuity Pathway Analysis of deregulated genes after DZNep treatment in conjunction with a siRNA-mediated knockdown of *CHD8* or *EP300*.

Deregulated canonical pathways in H2030_Cas9 cells that were treated with DZNep in conjunction with a knockdown of *CHD8* or *EP300* were annotated using Ingenuity Pathway Analysis. -log p-values larger than 4.32 (red line, equivalent to p = 0.05) indicate a significant gene enrichment.

As seen in Figure 44, novel canonical pathways that were not seen in either DZNep-treated or *CHD8*-depleted cells were identified in H2030 cells that were treated with DZNep in conjunction with a knockdown of *CHD8*. Notably, genes involved in `Mitotic Roles of Polo-Like Kinase´ (PLK) were significantly enriched. PLK has been reported as a target in *KRAS*-mutant NSCLC[72]. Furthermore, the canonical `Cell Cycle: G2/M

DNA Damage Checkpoint Regulation Pathway´ and ´Nucleotide excision repair (NER) Pathway´ were significantly enriched among the deregulated genes, further corroborating the previously observed decreased cell viability. Finally, genes involved in mTOR signaling, which is downstream of KRAS, were identified as significantly deregulated. The canonical ´NRF2-mediated oxidative stress response´- pathway was enriched in DZNep/siCHD8-treated cells. This was most significantly deregulated in DZNep/si*EP300*-treated cells. The strongly increased level of significance compared to non-DZNep, but *EP300* depleted cells (Figure 43 C) indicates a potential role of EP300-dependend NRF2-mediated signaling in response to DZNep treatment in *KEAP1*-mutant H2030 cells. Similar to DZNep/si*CHD8*-treated cells, multiple canonical pathways were significantly enriched in DZNep/si*EP00*-treated cells that were not observed in either DZNep or *EP300*-depleted cells, including `Cell Cycle: G2/M DNA Damage Checkpoint Regulation Pathway´ and ´Nucleotide excision repair (NER) Pathway´.

In conclusion, the IPA analysis identified multiple concordantly deregulated canonical pathways in DZNep treated H2030 cells after knockdown of either *CHD8* or *EP300*. Furthermore, the IPA analysis indicated a potential relevance of the ´NRF2-mediated oxidative stress response´- pathway in response to DZNep.

3.5.2 Identification of *CCL20* as a potential mediator of DZNep resistance

A possible strategy for cancer cells to cope with drug treatment is to up- or downregulate genes to facilitate survival. Consequently, H2030 cells may facilitate survival by deregulating genes that counteract the impact of DZNep. As the knockdown of *CHD8* or *EP300* resulted in increased susceptibility to DZNep, it was reasonable to assume that *CHD8* or *EP300*, or a downstream mediator thereof, would be required for such a response. Consequently, a knockdown of *CHD8* or *EP300* could result in opposing regulation of genes that were deregulated by DZNep. To analyze this hypothesis, genes that were deregulated by DZNep (foldchange cutoff values of > 1.5 or < 0.66) were called from the microarray data and analyzed for the presence of opposingly regulated gene expression upon a knockdown of *CHD8*, *EP300*, or combined DZNep treatment and gene knockdown (Figure 45).

	DZNep	siCHD8	siCHD8 + DZNep	siEP300	siEP300 + DZNep
RPS24	0.64	4.50	5.31	0.67	0.79
KIAA1199	0.64	3.24	1.02	1.21	0.75
RPS15	0.59	2.82	2.76	0.89	0.78
TMEM166	0.57	2.24	1.05	3.83	1.60
RPS7	0.65	1.96	2.06	0.88	0.66
SGCE	0.46	1.36	1.10	1.14	0.67
CAV2	0.62	1.76	1.25	0.61	0.70
DDIT3	0.66	1.76	1.15	2.52	0.87
RPS21	0.63	1.59	1.86	0.85	0.88
ANTXR2	0.66	1.57	1.15	1.35	1.28
BOLA2	0.66	1.53	1.58	0.99	0.80
MAGEA1	1.50	0.71	0.92	1.02	0.94
SFRS6	1.53	0.71	0.70	0.66	1.15
CLEC2B	1.51	0.70	0.91	0.73	1.25
FGB	1.57	0.70	0.90	0.55	0.75
BMP4	1.55	0.69	0.84	1.02	1.10
SNRNP200	1.69	0.75	0.74	0.69	0.84
KIAA0319	1.70	0.71	1.01	1.02	1.01
H3F3B	1.73	0.69	0.93	0.81	1.86
FTH1	1.52	0.60	0.78	0.69	1.22
LGMN	1.66	0.58	0.87	0.84	2.78
SLC22A18	1.89	0.62	0.77	0.85	0.81
CCL20	2.55	0.26	0.38	0.23	0.41

Figure 45 | Shortlisted genes that are opposingly regulated upon DZNep treatment (DZN) and *CHD8* or *EP300* knockdown in H2030 cells

Genes that were deregulated by DZNep (foldchange cutoff values of > 1.5 or < 0.66) were called from the microarray data and analyzed for the presence of opposingly regulated gene expression upon a knockdown of *CHD8* or *EP300*, or combined DZNep treatment and gene knockdown to identify possible mediators of synthetic lethality of *CHD8* and *EP300* with DZNep. The C-C Motif Chemokine Ligand 20 (*CCL20*) was the strongest upregulated gene upon DZNep treatment, but downregulated upon knockdown of *CHD8* and *EP300*. Combined treatment with DZNep and gene knockdown resulted in downregulation of *CCL20*.

Among the genes that fulfilled the above-mentioned criteria, the *C-C Motif Chemokine Ligand 20* (*CCL20*) was the strongest upregulated gene upon DZNep treatment (FC 2.55). Furthermore, a knockdown of *CHD8* or *EP300* resulted in the downregulation of *CCL20* to 26 % and 23 % of the expression level of the control group, respectively. Importantly, when *CHD8* knockdown was combined with DZNep treatment, *CCL20* mRNA levels remained low (38 % of the control group), conforming with 41 % for combined knockdown of *EP300* and DZNep treatment.

CCL20 is a member of the C-C motif chemokine subfamily, and involved in chemoattraction of lymphocytes and dendritic cells by binding to the chemokine receptor CCR6 [169]. *CCL20* has been reported to be triggered by chemotherapy and to

mediate resistance to taxane-containing drugs in triple-negative breast cancer cells[170]. Moreover, *CCL20* might be connected to lung cancer, as it has been demonstrated that tobacco smoke induces *CCL20* in this cancer entity[171]. Notably, other genes involved in inflammatory signaling, such as *CXCL1* or *CXCL2,* among others, showed similar responses to DZNep and *CHD8* or *EP300* knockdown (Figure 46).

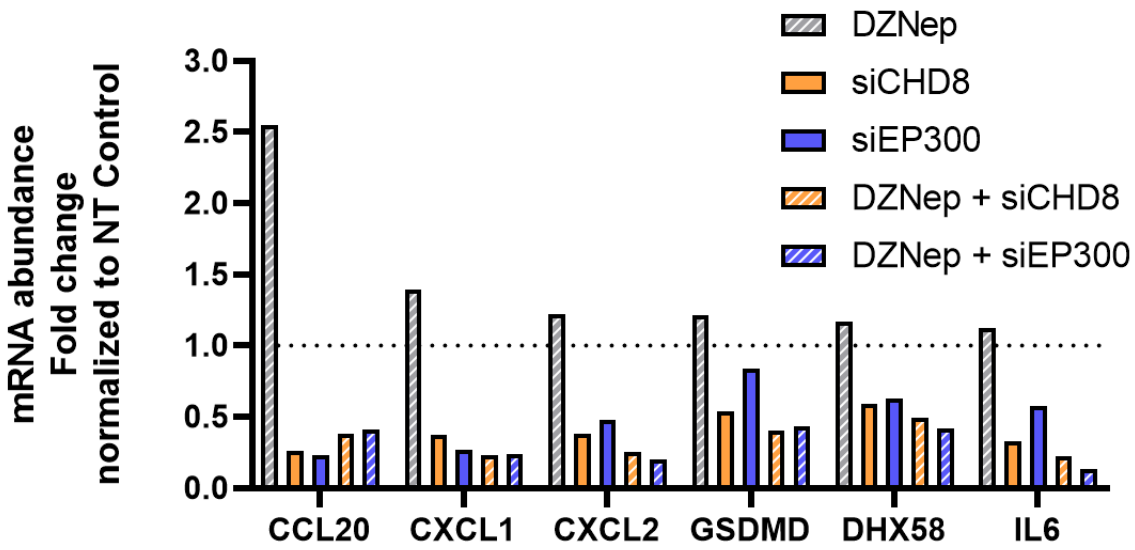


Figure 46 | mRNA levels of genes involved in inflammatory signaling upon DZNep treatment, knockdown of *CHD8* or *EP300* or combined treatment

Genes involved in inflammatory signaling were analyzed for deregulation upon DZNep treatment, knockdown of *CHD8* or *EP300* or combined treatment. All genes were downregulated by a knockdown of *CHD8* and *EP300*, and combined treatment of DZNep with knockdown of either *CHD8* or *EP300* did not restore gene expression.

Although no gene other than *CCL20* was upregulated > FC 1.5 following DZNep treatment, *CXCL1*, *CXCL2*, *GSDMD*, *DHDX58,* and *IL6* share the characteristic of being slightly upregulated upon DZNep treatment, while being downregulated upon knockdown of *CHD8* or *EP300*. As for *CCL20*, the downregulation remained stable in the presence of DZNep. This data may indicate a functional relevance of inflammatory signaling in conferring resistance to DZNep.

Summarizing the microarray results, gene expression profiling of H2030 cells revealed overlapping deregulated genes upon knockdown of *CHD8* or *EP300*, which were in concordance with results in the TCGA dataset. Functional annotation using the DAVID annotation tool showed enrichment of cell cycle, cell division, and mitosis annotation terms in cells that were treated with DZNep in conjunction with a knockdown of *CHD8* or *EP300*. The canonical ´NRF2-mediated oxidative stress response´- pathway was

identified as a possible effector pathway of synthetic lethality of EP300 and possibly CHD8 with DZNep by Ingenuity Pathway analysis. Furthermore, *CCL20* was identified as a potential mediator involved in maintaining H2030 cell viability upon DZNep treatment, making it an interesting target for further experiments.

3.6 Quantitative real-time PCR of *CCL20* mRNA levels following DZNep, *CHD8* knockdown or combined treatment

In order to validate the microarray findings of upregulated *CCL20* upon DZNep treatment and to analyze if this effect would also be seen in other cell lines, H2030, H1975 and H1944 cells were incubated with 0 µM, 1 µM, 2 µM and 5 µM DZNep for 4 days. Afterwards, the mRNA levels of *CCL20* were determined by RT-qPCR (Figure 47).

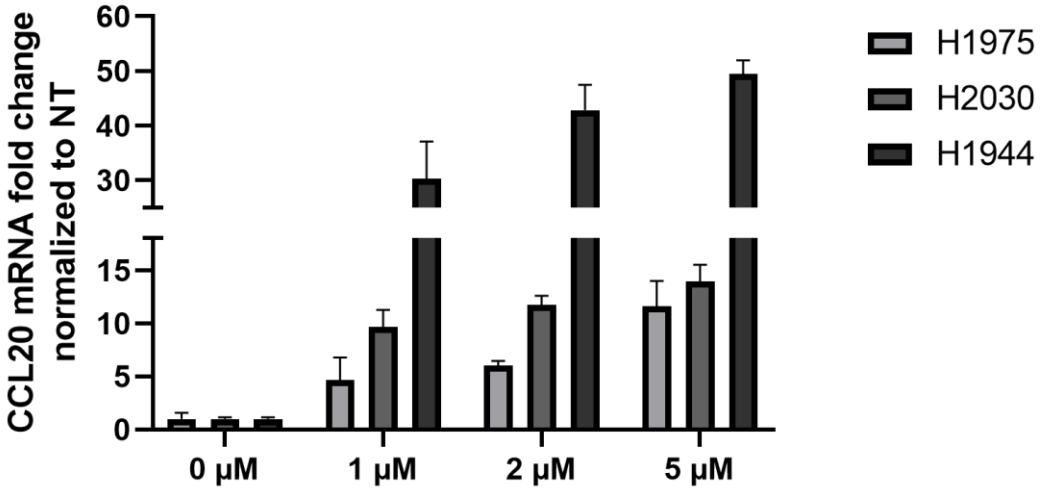


Figure 47 | *CCL20* mRNA levels in H1975, H2030, and H1944 cells as determined by qPCR after four days with increasing concentrations of DZNep

H1975, H2030, and H1944 cells were treated with 1 µM, 2 µM, and 5 µM DZNep treatment for four days. *CCL20* mRNA levels increased in a concentration-depended fashion in all cell lines.

In accordance with the microarray results, DZNep treatment upregulated *CCL20* mRNA in H2030 cells. In 1944 cells, the strongest upregulation was observed. Remarkably, *CCL20* levels were also upregulated in *EGFR* mutant H1975 cells upon DZNep treatment, although to a lesser extent compared to the other cell lines at doses of 1µM and 2 µM. As H1975 cells did not exhibit synthetic lethality of DZNep treatment and knockdown of *CHD8* or *EP300*, it remains elusive whether the upregulation of *CCL20* implicates a biological function in H1975.

To validate the microarray finding that a depletion of *CHD8* or *EP300* impaired the upregulation of *CCL20* upon DZNep treatment, *CHD8* and *EP300* were knocked down in H2030 cells using siRNA, and the cells were subsequently treated with DZNep as described above. This experiment was exploratory and performed in technical triplicates. The *CCL20* mRNA expression was determined by RT-qPCR and normalized to non-target siRNA treated cells (Figure 48).

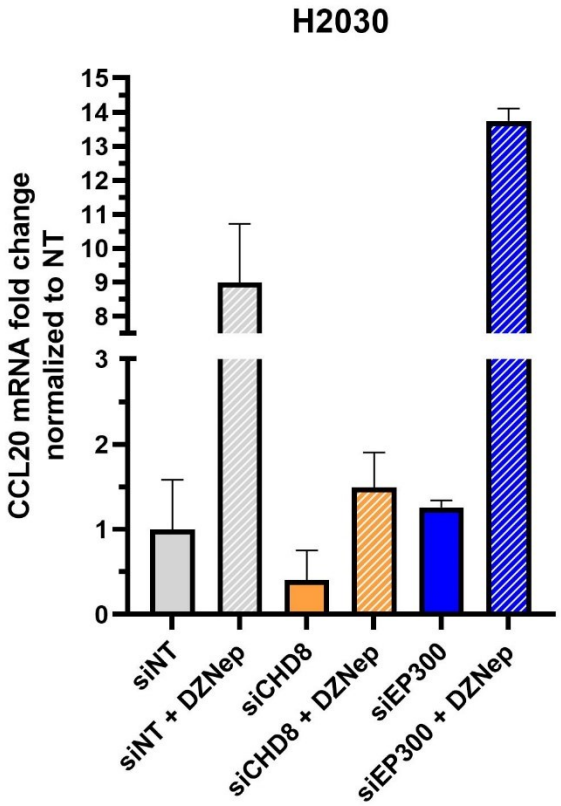


Figure 48 | Analysis of *CCL20* mRNA expression by qPCR in H2030 cells after 4 days of DZNep treatment, knockdown of *CHD8* or combined treatment

The abundance of *CCL20* mRNA was assessed following DZNep treatment, knockdown of *CHD8* or *EP300*, or DZNep treatment in conjunction with a knockdown of *CHD8* or *EP300*. *CCL20* mRNA levels are upregulated in DZNep treated cells, but knockdown of *CHD8* counteracts this upregulation. A knockdown of *EP300* did not affect *CCL20* mRNA levels upon DZNep treatment. All bars are normalized to the non-target (NT) sample.

As shown in Figure 48, the treatment of H2030 cells with DZNep induced the expression of *CCL20,* as seen before in the microarray results. A knockdown of *CHD8* led to downregulation of *CCL20*. Although 1 µM DZNep induced a 9-fold upregulation of *CCL20*, the *CCL20* mRNA levels increased only 1.5-fold when *CHD8* was depleted before DZNep treatment. This preliminary result confirms the microarray data and suggests that *CHD8* may be involved in transcriptional regulation of *CCL20* in H2030

cells. However, the microarray result could not be validated for *EP300*, as the knockdown of *EP300* did not decrease *CCL20* levels and did not impair the upregulation of *CCL20* upon DZNep treatment. This finding will need to be confirmed in future experiments.

3.7 Cell viability analysis after knockdown of *CCL20* and DZNep treatment

To corroborate the hypothesis that upregulation of *CCL20* might be an effector response to mediate DZNep resistance, *CCL20* was depleted from H2030 cells using RNAi, and the cells were subsequently treated with DZNep. This exploratory experiment was performed in technical triplicates. The cells were seeded at a density of 50,000 cells per well in a 12-well format and treated with 50 nM siRNA the following day. The cells were incubated with siRNA for 48 hours. Afterwards, the cells were divided into two arms and cells in the DZNep arm were treated with 1 µM DZNep immediately after splitting. After additional 72 hours of incubation, cell viability was assessed by staining with CellTiter Blue and normalized to untreated cells (Figure 49).

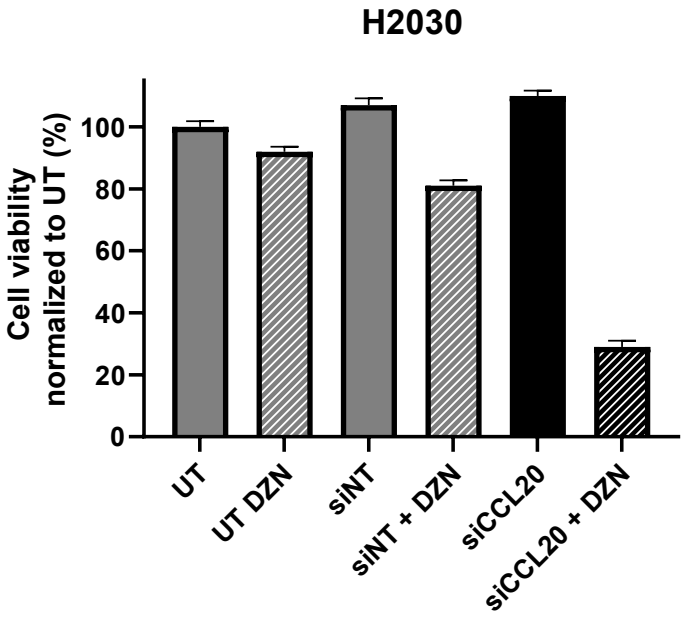


Figure 49 | Measurement of cell viability with CellTiter Blue in H2030 cells after DZNep treatment and siRNA mediated knockdown of *CCL20*.

H2030 cells were incubated with siRNA targeting *CCL20* for 48 hours and split into a DZNep arm and no-DZNep arm. After 72 hours the cell viability was measured by CTB and the data was normalized to the UT sample. A knockdown of *CCL20* or DZNep treatment did not impair cell viability in H2030 cells, but DZNep treatment in conjunction with a knockdown of *CCL20* led to decreased cell viability.

As seen in Figure 49, a knockdown of *CCL20* alone does not impair cell viability in H2030 cells. Importantly, when *CCL20* is depleted in DZNep treated cells, cell viability is reduced to 35 % of the DZNep only treated cells. This result indicates that *CCL20* might be involved in the facilitation of cell viability upon DZNep treatment in H2030 cells. In future experiments, additional cell lines need to be analyzed to elucidate a possible differential effect in *KRAS* and *EGFR*-mutant backgrounds.

Taken together, the results of sections 3.6 and 3.7 showed that *CCL20* is upregulated in a concentration-dependent manner in H2030, H1944, and H1975 cells. Further, an exploratory quantitative real-time PCR experiment showed that *CCL20* is upregulated upon DZNep treatment in H2030 cells and that *CHD8* is essential for the upregulation of *CCL20* upon DZNep treatment, similar to the gene expression profiling results observed before. However, this finding could not be reproduced for *EP300*. Finally, a knockdown of *CCL20* leads to decreased cell viability when performed in conjunction with DZNep treatment. Combined with the result that *CHD8* is essential for the upregulation of *CCL20*, this preliminary finding paves the way for a follow up study on the role of *CCL20* as a *CHD8*-controlled mediator of DZNep resistance in *KRAS*-mutant NSCLC cell lines.

4 Discussion

4.1 Summary of the results

This project successfully established a pooled lentiviral CRISPR/Cas9 screening platform for the identification of essential viability genes and synthetically lethal genes with epigenetic drugs. Using this platform, previously described and novel essential viability genes were identified in *KRAS*-mutant H2030 and *EGFR*-mutant H1975 NSCLC cell lines. Moreover, genes that act synthetically lethal with the epigenetic drugs DZNep and Entinostat were uncovered. Among the hits identified from the screens, two major target genes, *CHD8* and *EP300*, whose knockout was synthetically lethal with 1 µM DZNep treatment in H2030 cells, were selected for further characterization and functional studies.

In silico analysis of *CHD8* and *EP300* in the TCGA NSCLC dataset revealed strongly correlating expression levels and further indicated functional overlaps of both genes in NSCLC patients. The synthetically lethal effect of *CHD8* or *EP300* depletion with DZNep treatment was confirmed in H2030 cells and an additional *KRAS*-mutant cell line (H1944) by the quantification of the metabolic capacity of the cells. Flow cytometry analysis confirmed that the reduced cell viability was due to elevated apoptosis rates in cells that were treated with the combination of gene knockdown and DZNep.

The mechanism of synthetic lethality of the target genes with DZNep treatment was elucidated by gene expression profiling. The functional annotation of deregulated genes verified a high number of affected cell cycle and cell viability-related genes, and the NRF2-mediated Oxidative Stress Response Pathway was identified as potentially involved in mediating DZNep resistance in an *EP300*-, and possibly also *CHD8*-dependent fashion.

Further, the cytokine *CCL20* was identified as upregulated by DZNep, but downregulated upon knockdown of *CHD8* or *EP300*, suggesting a role in conferring *CHD8* or *EP300*-mediated resistance to DZNep treatment in H2030 cells. It was shown that *CCL20* is upregulated in a concentration-dependent fashion in *KRAS*-mutant H2030 and H1944-, but also in *EGFR*-mutant H1975 cells. Finally, initial proof of concept

experiments showed that *CHD8* is critical for the upregulation of *CCL20* upon DZNep treatment in H2030 cells and that a knockdown of *CCL20* mirrors the synthetically lethal effect of *CHD8* depletion with DZNep treatment.

4.2 Establishment of pooled lentiviral CRISPR/Cas9 screens

The first milestone of this work was to establish a platform for pooled CRISPR/Cas9 screens. It has been shown that the CRISPR/Cas9 technology-based screening approaches exhibit a higher target specificity, less false-negative hits, and higher rates of target validation compared to RNAi-based screens[139,140], and was therefore chosen for this work.

The setup of the CRISPR/Cas9 screens involved the production of lentiviral Cas9 particles, followed by Cas9 stable transfection of H2030 and H1975 cells and confirmation of functionally active Cas9. As a functional validation, 86 % knockout efficiency was shown in H2030_Cas9 cells by transfection of a guide RNA targeting the essential viability gene *PCNA*, and 78 % for H1975_Cas9 cells, whereas the transduction with a non-target control sgRNA did not affect cell viabillty. Both values are in good accordance with published data for the efficacy of CRISPR/Cas9-mediated knockouts and therefore were suitable for subsequent gene knockout screens[172].

The sgRNA library used for this study covered 6,500 genes with 6 to 8 sgRNAs targeting each gene. Previously published CRISPR/Cas9 screens mostly consisted of 4 to 6 sgRNAs per gene[173]; consequently, the high count of sgRNAs per gene was a solid foundation for producing reliable results. To prevent false-positive dropouts from the pool of sgRNAs due to a cell harboring multiple sgRNA constructs, 2000 cells were transduced per sgRNA, and a minimum coverage of 1,000 cells per sgRNA was maintained throughout the screen, even after cell splitting. As a result, the screens performed in this study exceeded the coverage of previously reported screens 2 to 3-fold[174]. The high number of sgRNAs per gene and the high coverage of cells per sgRNA resulted in more than 2000 reads per gene after high-throughput sequencing, which allowed a detailed analysis with a high signal over background noise ratio.

The reliability of the screens conducted in this study was validated by an analysis of 88 positive and 172 negative (non-target) control sgRNAs. The abundances of the control sgRNAs in the no-drug-condition of each screen was compared to their abundances in the vector library. As expected, almost all positive controls, whose knockout would kill

a cell, were found in the group of significantly depleted sgRNAs, proving efficient Cas9-mediated knockout. In contrast, negative control sgRNAs were randomly distributed, proving the absence of unwanted off-target effects.

Taken together, the data presented in this study showed that stably Cas9 expressing cell lines were successfully established and that the risk of biased results of pooled lentiviral CRISPR/Cas9 screens were mitigated at all steps of the process. Finally, the analysis of control sgRNAs proved high robustness of the sgRNA readcounts.

4.3 Identification of essential viability genes in H2030_Cas9 and H1975_Cas9 cell lines

Oncogenic mutations in genes, such as *EGFR* or *KRAS*, may cause dependencies on other genes coding for proteins in signaling pathways, metabolic pathways, or transcriptional regulators[72]. Therefore, the identification of genes whose depletion selectively impairs the viability of a cell harboring an oncogenic mutation is a promising approach to deepen the understanding of the processes within cancer cells. In this study, 242 essential viability genes in *KRAS*-mutant H2030_Cas9- and 290 genes in *EGFR*-mutant H1975_Cas9 cells were identified.

In H2030_Cas9 cells, previously published targets genes in *KRAS* mutant NSCLC cells were confirmed in this work, e.g., *PLK1*[175], *TAK1/MEKK7/MAP3K7*[176], and *AURKA*[177], all of which are linked to aberrant *KRAS* signaling. In this study, these genes were identified as viability genes in the H2030_Cas9 only, but not in the EGFR-mutant H1975, cell lines, which is a strong indication that H2030 cells depend on oncogenic *KRAS* signaling, despite contradictory reports[74,178]. The identification of several previously published targets served as a control for the screen quality and indicated a robust and reliable screening platform. Since the main scope of this work was to identify targets for synthetic lethality with epigenetic drugs, the list of identified target genes will serve as a good starting point for further studies on novel targets.

As for H2030_Cas9 cells, previously reported viability genes were found among the significantly depleted genes in H1975_Cas9 cells, e.g., *Phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha* (*PIK3CA*). H1975 cells harbor an activating mutation of *PIK3CA*[179], and the gene has been described as an oncogenic driver in lung cancer[180]. In conclusion, the CRISPR/Cas9 screening platform was functionally validated by confirming previously published essential viability genes.

Of note, the depletion of driver gene sgRNA readcounts was stronger in H1975_Cas9 cells than in H2030_Cas9 cells. In particular, the knockout of *EGFR* resulted in a residual cell viability of 15 %, as measured by the reduction of *EGFR*-targeting sgRNA readcounts in the no-drug arm compared to the vector library. Compared to this result, the knockout of *KRAS* resulted in 48 % of the reference readcounts in H2030_Cas9 cells, suggesting a stronger dependence on *EGFR* as an oncogenic driver in H1975_Cas9 cells compared to *KRAS* in H2030_Cas9 cells.

As expected, the functional annotation of the identified viability genes reported an enrichment of cell cycle-related pathways in H2030_Cas9 cells. However, in H1975_Cas9 cells, the impaired viability could not be directly mapped to cell cycle-related pathways, but rather to tRNA biosynthesis, ribosomal function, or metabolic pathways. Based on this analysis, the underlying mechanisms that lead to decreased viability may differ in both cell lines.

Overall, the results demonstrate that the established screening platform was able to identify and confirm *EGFR* and *KRAS* as essential viability genes, thereby validating the screening platform to be reliable. Furthermore, previously reported cell viability genes, as well as novel potential target genes, were identified. These novel target genes may serve as a starting point for further independent studies.

4.4 Identification of synthetically lethal genes with epigenetic drugs

This work aimed to identify synthetically lethal genes with epigenetic drugs, with a focus on DZNep. DZNep acts as a direct inhibitor of AHCY, which leads to the accumulation of S-Adenosyl-L-homocysteine and thereby inhibits EZH2 through a negative feedback loop, causing the induction of apoptosis[121,122,181]. As shown in this study, the viability of H2030 cells was only slightly impaired by DZNep, even at high concentrations of 10 µM. While no published data concerning DZNep treatment of H2030 cells exist to date, other NSCLC cell lines have been shown to be sensitive to DZNep treatment[122]. H2030 is among the NSCLC cell lines with the lowest expression of *EZH2*[182], giving a rationale for the insensitivity to DZNep in this cell line. In contrast, H1944 cells showed a moderate response to DZNep in this study, where treatment with 1 µM DZNep resulted in a decrease of cell viability to 75 % compared to the control group.

Interestingly, EZH2 levels and sensitivity to DZNep in H1944 cells have been shown to be in the same range as in the *EZH2*-upregulated H1299 NSCLC cell line[120,183], suggesting a link to the slightly elevated DZNep sensitivity of H1944 cells compared to H2030 cells. Another possible cause for DZNep resistance has been shown in a B-cell lymphoma model, where resistance to DZNep could be tracked to copy number gains of AHCY[184]. However, no records for *AHCY* copy number variations or upregulated expression are filed within the COSMIC Cell Lines Project Database for H2030 cells.

Four possible mediators of DZNep resistance in H2030 cells were identified as candidates for further characterization in this study, in particular *PRMT1*, *KDM5C*, *CHD8*, and *EP300*. The protein arginine methyltransferase 1 (*PRMT1*) is involved in the regulation of many essential cell functions through the modification of proteins, including histone H2, H3, and H4, leading to aberrant gene expression.[185]. An upregulation of *PRMT1* has been found in many types of cancer, including NSCLC[186]. It has been shown that PRMT1 regulates cell proliferation and epithelial-mesenchymal-transition, and its depletion resulted in decreased cell growth[187].

Interestingly, PRMT1 features a SAM-dependent methyltransferase PRMT-type domain[187], providing a link to DZNep, which inhibits the conversion of SAH to SAM. Indeed, depletion of *PRMT1* was found to be synthetically lethal with DZNep treatment in H2030_Cas9 cells in this study; moreover, it exhibited the highest synthetic lethality with DZNep of all four identified target genes. However, it was also observed that a knockout of *PRMT1* even without additional DZNep treatment significantly decreased the viability of H2030_Cas9 cells in the pooled CRISPR/Cas9 screen. Since this work focused on targets that would increase the sensitivity to DZNep without being essential viability genes themselves, *PRMT1* was not taken into consideration for further characterization.

In contrast to PRTM1, the protein demethylase KDM5C catalyzes the demethylation of arginine residues of H3 histone-, but also non-histone proteins. KDM5C has been suggested as a tumor suppressor, and mutations in *KDM5C* have been identified in clear cell renal cell carcinoma, pancreatic cancer, and human papillomavirus-associated cancer[188]. However, a knockdown of *KDM6A*, which is, as *KDM5C*, classified into the subgroup of JmjC domain-containing demethylases, has been reported as an epigenetic mechanism contributing to lung tumorgenesis[187]. A knockout of *KDM6A* increased EZH2 protein levels in an *in vivo* lung cancer model and sensitized the tumor cells to DZNep treatment[189]. Since no comparable findings have been published for *KDM5C*,

the identification of synthetic lethality of *KDM5C* with DZNep in this study could be a first indication that *KDM5C* exhibits similar effects to *KDM6A*. *KDM5C* did not only show up as a synthetically lethal gene with DZNep but also with Entinostat in H2030_Cas9 cells; therefore, *KDM5C* was excluded from the shortlisted candidate genes due to the risk of an unspecific mechanism.

In conclusion, four synthetically lethal genes have been identified by a pooled CIRSPR/Cas9 screen in *KRAS*-mutant H2030 NSCLC cells. *PRMT1* was ruled out as a synthetically lethal candidate gene due to its characteristic of being an essential viability gene. *KDM5C* was ruled out due to the fact that it also appeared as a synthetically lethal hit with Entinostat in the same cell line, pointing to a rather unspecific mechanism. Consequently, *CHD8* and *EP300* were selected for further analysis, and both genes will be discussed in the following section.

4.5 CHD8 and EP300 as targets for exploiting synthetic lethality with DZNep

CHD8 and *EP300* were identified as significantly depleted from the DZNep-treated cell pool with median sgRNA foldchange values of 0.56 (*CHD8*) and 0.70 (*EP300*).

CHD8 belongs to the family of ATP-dependent chromodomain helicase DNA-binding (CHD) proteins. The *CHD* family consists of nine genes, which are grouped into three subfamilies. *CHD8* belongs to subfamily III, whose members are characterized by the presence of Bromo- and Kismet domains. CHD proteins affect chromatin compaction and are involved in fundamental biological processes, including transcription, cellular proliferation, and DNA damage repair[190]. EP300 is a lysine acetyltransferase (KAT) that is involved in transcriptional regulation through acetylation of histones and transcription factors, such as P53, HIF-1α, or STAT1[191]. EP300 acts as a cofactor with a large number of transcription factors, thereby regulating the expression of thousands of genes. Similar to CHD8, several fundamental biological processes are regulated by EP300, including proliferation, cell cycle, cell differentiation, and DNA damage response[161]. Of note, neither a knockout nor a knockdown of *CHD8* or *EP300* affected the viability of H2030 and H1944 cells in this study, suggesting that both cell lines do not depend on either of the proteins for the above-mentioned biological processes.

Interestingly, the expression of *CHD8* and *EP300* significantly correlated in NSCLC patients in the TCGA dataset, and genes whose expression positively correlated with the

expression levels of *CHD8* and *EP300* showed considerable overlap. Furthermore, the functional annotation of *CHD8*-correlating genes and *EP300*-correlating genes in NSCLC patients revealed that `S-adenosyl-methionine (SAM)´ and `methyltransferase activity´ are related in both genesets, suggesting a potential link to DZNep, which causes alterations in the ratio of SAM and SAH. In accordance with the TCGA data, a microarray analysis of siRNA-mediated knockdown of either *CHD8* or *EP300* showed that almost half of all genes that were deregulated upon *CHD8* knockdown were also deregulated upon *EP300* knockdown and about one-third of genes that were deregulated vice versa. However, the microarray analysis could not confirm the TCGA data of overlapping functional annotations for both genes, and only a few genes related to SAM and methyltransferase activity were found among the deregulated genes upon knockdown of *CHD8* or *EP300*. As the TCGA dataset is comprised of data from over 500 patients, it might be difficult to recapitulate these results in one clonal cell line. To elucidate overlapping effects of *CHD8* and *EP300* on SAM, measurements of intracellular SAM and SAH levels could be performed in follow-up studies. Also, a pulldown of CHD8 and EP300 in NSCLC cell lines may be feasible to identify a direct interaction of CHD8 and EP300, e.g., in complex formation, giving further insights into possible overlapping functions.

CHD8 and *EP300* have been implicated in cancer development and progression. *CHD8* is frequently mutated in breast cancer[154], deleted in gastric cancers and colorectal cancers[155], and silenced by promoter methylation in prostate cancer[156]. Further, CHD8 is required to facilitate a proliferative effect of estrogen in breast cancer cells[192]. In colorectal cancers, mutations of *CHD8* correlate with mutations in *BRAF*, *KRAS*, and *TP53*[193]. Finally, the expression of *CHD8* has been reported to correlate with decreased survival and increased metastasis in prostate cancer patients[156]. However, as some of these reports suggest tumor-suppressing properties of *CHD8* while others suggest proto-oncogenic properties, as of now the precise role of *CHD8* in cancer is not clear and needs further elucidation. Similarly, *EP300* has been reported to have tumor-suppressive, but also proto-oncogenic properties. For instance, EP300 has been described to act as a tumor suppressor in breast cancer, colorectal cancer, and gastric cancer[194,195]. On the other hand, over-expression of *EP300* is a poor prognostic factor in breast cancer, prostate cancer, hepatocellular carcinoma, and esophageal squamous cell carcinoma[196–199], and low *EP300* expression was found to be an independent prognostic marker of favorable clinical outcomes in operable NSCLC patients[200]. Another study

showed an opposing finding, where the inhibition of EP300 degradation decreased the proliferation and metastasis activity of lung cancer cells, indicating that EP300 acts as a tumor suppressor in lung cancer[201].

In conclusion, the role of *EP300* remains unclear and might depend on the cancer entity and additional factors that need to be elucidated. In this study, using shRNA-mediated knockdown and subsequent CTB cell viability analysis, the synthetically lethal action with DZNep treatment was confirmed for *CHD8* and *EP300*, and it was shown that the observed effects were independent of CRISPR/Cas9. The finding of synthetic lethality could also be validated by elevated apoptosis rates in DZNep treated cells that additionally received a knockdown of either *CHD8* or *EP300*. Both findings of decreased cell viability and enhanced apoptosis were reproduced in *KRAS*-mutant H1944, but not in *EGFR*-mutant H1975 cells. Taken together, these results demonstrate that CHD8 and EP300 are involved in maintaining DZNep resistance in two *KRAS*-mutant NSCLC cell lines, indicating their roles as proto-oncogenes .

The Ingenuity Pathway Analysis of deregulated canonical pathways upon DZNep treatment in conjunction with a knockdown of *EP300* indicated a possible role of the ´NRF2-mediated oxidative stress response´- pathway in response to the drug. Importantly, EP300 has been identified as a key molecule in this pathway, which is a primary regulator of cellular processes that aid in resisting the action of chemotherapy drugs[202]. Of note, the ´NRF2-mediated oxidative stress response´-pathway showed a strong tendency towards enrichment ($p = 0.054$) in DZNep treated cells also in conjunction with a knockdown of *CHD8*, but no implications of CHD8 in the NRF2 stress response pathway have been reported yet. In this context, it is important to mention that both H2030 and H1944 cells harbor a mutation in the NRF2 suppressor *KEAP1*, suggesting a link of EP300, and possibly CHD8, to the NRF2 stress response pathway as a mediator of DZNep resistance in these cell lines. To further look into this, the deregulation of NRF2 pathway annotated genes in additional *KRAS*-mutant, but *KEAP1*wt cell lines should be assessed. Furthermore, a knockdown of *NRF2* in conjunction with DZNep treatment could be utilized to elucidate the role of this pathway in mediating DZNep resistance.

In summary, *CHD8* and *EP300* have been identified as synthetically lethal genes with DZNep treatment in *KRAS*-mutant H2030 cells, and the effects were confirmed in *KRAS*-mutant H1944, but not in *EGFR*-mutant H1975 cells. On the functional level, the canonical `NRF2-mediated oxidative stress response´- pathway was identified as

deregulated in DZNep/siEP300-treated cells and showed a trend in DZNep/siCHD8-treated cells, making it a suitable target for further elucidation in subsequent studies.

4.6 CCL20 as a possible effector of CHD8 or EP300-mediated synthetic lethality with DZNep treatment

In addition to the identification of canonical pathways, this work aimed at identifying candidate genes that might be involved in facilitating DZNep resistance in a CHD8/EP300-dependend fashion. Following the rationale that possible mediators would be upregulated upon DZNep treatment to facilitate cell survival, but downregulated upon the depletion of *CHD8* or *EP300* from the cell using siRNA, the chemokine CCL20 was identified. CCL20 is a member of the C-C motif chemokine subfamily and is involved in the chemoattraction of lymphocytes and dendritic cells by binding the chemokine CCR6 receptor[169]. However, implications of CCL20 have also been described beyond this function, and a possible role of CCL20 in different cancer entities has been suggested. For instance, it has been reported that tobacco smoke induces the expression of *CCL20* in lung cancer, thereby promoting disease progression[171], and *CCL20* expression was shown to be triggered by taxane chemotherapy in breast cancer cells, leading to chemoresistance via activation of NFkB[170]. Of note, other genes involved in inflammatory signaling, such as *CXCL1*, *CXCL2,* and *IL6* were found in this study to act similarly to *CCL20*, that is, being downregulated by either *CHD8* or *EP300* depletion and being upregulated upon DZNep treatment to varying degrees. In this regard, it is worth mentioning that chemokines have emerged as important regulators in NSCLC[203]. Serresi et al. observed an upregulation of *IL6* upon PRC2 inhibition[119] and also reported that an inflammatory response is triggered upon inhibition of the PRC2-member EZH2, which subsequently confers resistance to the EZH2 inhibitor GSK126[127].

Of note, the upregulation of *CCL20* upon DZNep treatment could also be due to *CCL20* being a target of PRC2, and *CCL20* being upregulated by the inhibition of EZH2. In favor of this argument is that *CCL20* is also upregulated in DZNep-treated H1975 cells, which did not exhibit a synthetically lethal action of DZNep and si*CHD8*/si*EP300*. On the other hand, an initial experiment showed that a knockdown of *CCL20* is synthetically lethal with DZNep treatment, pointing to functional relevance. However, this finding needs to be reproduced in H2030 cells. Furthermore, H1944 and H1975

cells need to be probed for synthetic lethality of *CCL20* depletion and DZNep treatment. Additionally, inhibitors of CCL20 and the CCR6 receptor[204] could be utilized to elucidate the role of CCL20 in DZNep resistance.

Taken together, CCL20 was identified as a possible CHD8/EP300-dependend mediator of DZNep resistance in H2030 cells, and literature reports are in favor of a possible role of CCL20 in drug resistance in other cancer entities. These findings will serve as a good starting point for further studies to elucidate the role of the CHD8/EP300/CCL20-axis in resistance to DZNep.

4.7 Outlook

This work established a pooled CRISPR/Cas9 screen and identified *CHD8* and *EP300* as synthetically lethal target genes with DZNep treatment in *KRAS*-mutant H2030 cells. Gene expression profiling using Ingenuity Pathway Analysis identified the NRF2-mediated oxidative stress response – pathway as significantly altered upon DZNep treatment in conjunction with a knockdown of *EP300.* Also, DZNep treatment in conjunction with a knockdown of *CHD8* showed a strong trend towards significance for this pathway. Further, CCL20 was identified as a possible mediator of CHD8/EP300 – dependent resistance to DZNep in H2030 cells. In follow up studies to this project, it would be of high value to elucidate if the CHD8/EP300 – dependent upregulation of *CCL20* confers resistance to DZNep in other NSCLC cell lines by combined DZNep treatment and *CCL20* knockdown experiments and by blocking the CCR6 receptor. Ultimately, the dependence of such effects on the *KRAS* and *KEAP1* mutation status of the cells should be studied, thereby integrating this study's findings of deregulated NRF2 signaling and *CCL20* deregulation.

5 References

1. The Global Cancer Observatory. *Cancer fact sheet*. *World Health Organization* (2019) doi:L11\n10.1051/0004-6361/201016331.

2. Thomas, A., Liu, S. V., Subramaniam, D. S. & Giaccone, G. Refining the treatment of NSCLC according to histological and molecular subtypes. *Nat. Rev. Clin. Oncol.* 1–16 (2015) doi:10.1038/nrclinonc.2015.90.

3. Siegel, R. L. & Miller, et al. Cancer Statistics, 2018. *Ca Cancer J Clin* (2018) doi:10.3322/caac.21387.

4. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, C. K. SEER Cancer Statistics Review, 1975-2016, National Cancer Institute.

5. Cooley, M. E. Symptoms in adults with lung cancer: A systematic research review. *J. Pain Symptom Manage.* **19**, 137–153 (2000).

6. Walser, T. *et al.* Smoking and lung cancer: The role of inflammation. *Proc. Am. Thorac. Soc.* **5**, 811–815 (2008).

7. Beckerman, S. C. The Health Consequences of Smoking: A Report of the Surgeon General. *The Journal of the Maine Medical Association* (2004) doi:10.2307/1973644.

8. de Groot, P. M., Wu, C. C., Carter, B. W. & Munden, R. F. The epidemiology of lung cancer. *Transl. Lung Cancer Res.* **7**, 220–233 (2018).

9. Wood, S. L., Pernemalm, M., Crosbie, P. A. & Whetton, A. D. Molecular histology of lung cancer: From targets to treatments. *Cancer Treat. Rev.* **41**, 361–375 (2015).

10. Noone AM, Howlader N, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, C. K. Cancer Statistics Review, 1975-2015 - SEER Statistics. *SEER Cancer Stat. Rev.* (2017).

11. Molina, J. R., Yang, P., Cassivi, S. D., Schild, S. E. & Adjei, A. A. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin. Proc.* **83**, 584–94 (2008).

12. Dresler, C. M. *et al.* Phase III intergroup study of talc poudrage vs talc slurry sclerosis for malignant pleural effusion. *Chest* (2005) doi:10.1378/chest.127.3.909.

13. Inamura, K. Lung cancer: understanding its molecular pathology and the 2015 wHO classification. *Front. Oncol.* **7**, 1–7 (2017).

14. Travis, W. D. *et al.* International association for the study of lung cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma. *J. Thorac. Oncol.* **6**, 244–285 (2011).

15. Leeman, J. E. *et al.* Histologic Subtype in Core Lung Biopsies of Early-Stage Lung Adenocarcinoma is a Prognostic Factor for Treatment Response and Failure Patterns After Stereotactic Body Radiation Therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **97**, 138–145 (2017).

16. Jameson JL; Fausi AS et al. *Harrison's Principles of Internal Medicine, 20e*. *McGraw-Hill* (2018).

17. Collisson, E. A. *et al.* Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550 (2014).

18. Skoulidis, F. *et al.* Co-occurring genomic alterations define major subsets of KRAS-mutant lung adenocarcinoma with distinct biology, immune profiles, and therapeutic vulnerabilities. *Cancer Discov.* **5**, 861–878 (2015).

19. Malumbres, M. & Barbacid, M. RAS oncogenes: The first 30 years. *Nat. Rev. Cancer* **3**, 459–465 (2003).

20. Santos, E. *et al.* Malignant activation of a K-ras oncogene in lung carcinoma but not in normal tissue of the same patient. *Science (80-.).* (1984) doi:10.1126/science.6695174.

21. Simanshu, D. K., Nissley, D. V. & McCormick, F. RAS Proteins and Their Regulators in Human Disease. *Cell* **170**, 17–33 (2017).

22. Cherfils, J. & Zeghouf, M. Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiological Reviews* (2013) doi:10.1152/physrev.00003.2012.

23. Gysin, S., Salt, M., Young, A. & McCormick, F. Therapeutic Strategies for Targeting Ras Proteins. *Genes Cancer* **2**, 359–372 (2011).

24. Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* (2001) doi:10.1126/science.1062023.

25. Hunter, J. C. *et al.* Biochemical and structural analysis of common cancer-associated KRAS mutations. *Mol. Cancer Res.* (2015) doi:10.1158/1541-7786.MCR-15-0203.

26. Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D. & Kuriyan, J. The structural basis of the activation of Ras by Sos. *Nature* (1998) doi:10.1038/28548.

27. Gorfe, A. A., Grant, B. J. & McCammon, J. A. Mapping the Nucleotide and Isoform-Dependent Structural and Dynamical Features of Ras Proteins. *Structure* (2008) doi:10.1016/j.str.2008.03.009.

28. Deshpande, N. *et al.* The RCSB Protein Databa Bank: A redesigned query system and relational database based on the mmCIF schema. *Nucleic Acids Res.* (2005) doi:10.1093/nar/gki057.

29. Hancock, J. F. & Parton, R. G. Ras plasma membrane signalling platforms. *Biochemical Journal* (2005) doi:10.1042/BJ20050231.

30. Rotblat, B. *et al.* Three Separable Domains Regulate GTP-Dependent Association of H-ras with the Plasma Membrane. *Mol. Cell. Biol.* (2004) doi:10.1128/mcb.24.15.6799-6810.2004.

31. McCormick, F. How receptors turn Ras on. *Nature* **363**, 15–16 (1993).

32. Ferrer, I. *et al.* KRAS-Mutant non-small cell lung cancer: From biology to therapy. *Lung Cancer* **124**, 53–64 (2018).

33. Khosravi-Far, R., Solski, P. A., Clark, G. J., Kinch, M. S. & Der, C. J. Activation of Rac1, RhoA, and mitogen-activated protein kinases is required for Ras transformation. *Mol. Cell. Biol.* (1995) doi:10.1128/mcb.15.11.6443.

34. Leevers, S. J., Paterson, H. F. & Marshall, C. J. Requirement for Ras in Raf activation is overcome by targeting Raf to the plasma membrane. *Nature* (1994) doi:10.1038/369411a0.

35. Gupta, S. *et al.* Binding of Ras to Phosphoinositide 3-Kinase p110α Is Required for Ras- Driven Tumorigenesis in Mice. *Cell* (2007) doi:10.1016/j.cell.2007.03.051.

36. Rodriguez-Viciana, P. *et al.* Role of phosphoinositide 3-OH kinase in cell

transformation and control of the actin cytoskeleton by Ras. *Cell* (1997) doi:10.1016/S0092-8674(00)80226-3.

37. Gibbs, J. B., Sigal, I. S., Poe, M. & Scolnick, E. M. Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules. *Proc. Natl. Acad. Sci. U. S. A.* (1984) doi:10.1073/pnas.81.18.5704.

38. McGrath, J. P., Capon, D. J., Goeddel, D. V. & Levinson, A. D. Comparative biochemical properties of normal and activated human ras p21 protein. *Nature* (1984) doi:10.1038/310644a0.

39. Sweet, R. W. *et al.* The product of ras is a GTPase and the T24 oncogenic mutant is deficient in this activity. *Nature* (1984) doi:10.1038/311273a0.

40. Scheffzek, K., Lautwein, A., Kabscht, W., Ahmadian, M. R. & Wittinghofer, A. Crystal structure of the GTPase-activating domain of human p120GAP and implications for the interaction with Ras. *Nature* (1996) doi:10.1038/384591a0.

41. Arbour, K. C. *et al.* Effects of co-occurring genomic alterations on outcomes in patients with KRAS-mutant non-small cell lung cancer. *Clin. Cancer Res.* **24**, 334–340 (2018).

42. Dogan, S. *et al.* Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: Higher susceptibility of women to smoking-related KRAS-mutant cancers. *Clin. Cancer Res.* (2012) doi:10.1158/1078-0432.CCR-11-3265.

43. Dearden, S., Stevens, J., Wu, Y. L. & Blowers, D. Mutation incidence and coincidence in non small-cell lung cancer: Meta-analyses by ethnicity and histology (mutMap). *Ann. Oncol.* (2013) doi:10.1093/annonc/mdt205.

44. Redig, A. J. *et al.* Genomic complexity in KRAS mutant non-small cell lung cancer (NSCLC) from never/light-smokers v smokers. . *J. Clin. Oncol.* (2016) doi:10.1200/jco.2016.34.15_suppl.9087.

45. Mascaux, C. *et al.* The role of RAS oncogene in survival of patients with lung cancer: A systematic review of the literature with meta-analysis. *British Journal of Cancer* (2005) doi:10.1038/sj.bjc.6602258.

46. Johnson, M. L. *et al.* Association of KRAS and EGFR mutations with survival in patients with advanced lung adenocarcinomas. *Cancer* (2013)

doi:10.1002/cncr.27730.

47. Shepherd, F. A. *et al.* Pooled analysis of the prognostic and predictive effects of KRAS mutation status and KRAS mutation subtype in early-stage resected non–small-cell lung cancer in four trials of adjuvant chemotherapy. *J. Clin. Oncol.* (2013) doi:10.1200/JCO.2012.48.1390.

48. Macerelli, M. *et al.* Does KRAS mutational status predict chemoresistance in advanced non-small cell lung cancer (NSCLC)? *Lung Cancer* (2014) doi:10.1016/j.lungcan.2013.12.013.

49. Khambata-Ford, S. *et al.* Analysis of potential predictive markers of cetuximab benefit in BMS099, a phase III study of cetuximab and first-line taxane/carboplatin in advanced non-small-cell lung cancer. *J. Clin. Oncol.* (2010) doi:10.1200/JCO.2009.25.2890.

50. O'Byrne, K. J. *et al.* Molecular biomarkers in non-small-cell lung cancer: a retrospective analysis of data from the phase 3 FLEX study. *Lancet Oncol.* (2011) doi:10.1016/S1470-2045(11)70189-9.

51. Yang, H., Liang, S. Q., Schmid, R. A. & Peng, R. W. New horizons in KRAS-mutant lung cancer: Dawn after darkness. *Front. Oncol.* **9**, 1–13 (2019).

52. Ihle, N. T. *et al.* Effect of KRAS oncogene substitutions on protein behavior: Implications for signaling and clinical outcome. *J. Natl. Cancer Inst.* (2012) doi:10.1093/jnci/djr523.

53. Li, S. *et al.* Assessing therapeutic efficacy of MEK inhibition in a KrasG12C-driven mouse model of lung cancer. *Clin. Cancer Res.* **24**, 4854–4864 (2018).

54. Nadal, E. *et al.* KRAS-G12C mutation is associated with poor outcome in surgically resected lung adenocarcinoma. *J. Thorac. Oncol.* (2014) doi:10.1097/JTO.0000000000000305.

55. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* (2013) doi:10.1038/nature12796.

56. Janes, M. R. *et al.* Targeting KRAS Mutant Cancers with a Covalent G12C-Specific Inhibitor. *Cell* **172**, 578-589.e17 (2018).

57. Canon, J. *et al.* The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour

immunity. *Nature* (2019) doi:10.1038/s41586-019-1694-1.

58. Corral de la Fuente, E., Olmedo Garcia, M. E., Gomez Rueda, A., Lage, Y. & Garrido, P. Targeting KRAS in Non-Small Cell Lung Cancer. *Front. Oncol.* **11**, 1–11 (2022).

59. Blasco, R. B. *et al.* C-Raf, but Not B-Raf, Is Essential for Development of K-Ras Oncogene-Driven Non-Small Cell Lung Carcinoma. *Cancer Cell* (2011) doi:10.1016/j.ccr.2011.04.002.

60. Blumenschein, G. R. *et al.* A randomized phase II study of the MEK1/MEK2 inhibitor trametinib (GSK1120212) compared with docetaxel in KRAS-mutant advanced non-small-cell lung cancer (NSCLC). *Ann. Oncol.* (2015) doi:10.1093/annonc/mdv072.

61. Carter, C. A. *et al.* Selumetinib with and without erlotinib in KRAS mutant and KRAS wild-type advanced nonsmall-cell lung cancer. *Ann. Oncol.* (2016) doi:10.1093/annonc/mdw008.

62. D.E., G. *et al.* Phase ii study of defactinib, VS-6063, a focal adhesion kinase (FAK) inhibitor, in patients with KRAS mutant non-small cell lung cancer (NSCLC). *J. Thorac. Oncol.* (2015).

63. Riely, G. J. *et al.* A randomized discontinuation phase II trial of ridaforolimus in non-small cell lung cancer (NSCLC) patients with KRAS mutations. *J. Clin. Oncol.* (2012) doi:10.1200/jco.2012.30.15_suppl.7531.

64. Jänne, P. A. *et al.* Selumetinib plus docetaxel for KRAS-mutant advanced non-small-cell lung cancer: A randomised, multicentre, placebo-controlled, phase 2 study. *Lancet Oncol.* (2013) doi:10.1016/S1470-2045(12)70489-8.

65. Jänne, P. A. *et al.* Selumetinib plus docetaxel compared with docetaxel alone and progression-free survival in patients with KRAS-mutant advanced non-small cell lung cancer: The SELECT-1 randomized clinical trial. *JAMA - J. Am. Med. Assoc.* (2017) doi:10.1001/jama.2017.3438.

66. Engelman, J. A. *et al.* Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. *Nat. Med.* (2008) doi:10.1038/nm.1890.

67. Sos, M. L. *et al.* Identifying genotype-dependent efficacy of single and combined

PI3K- and MAPK-pathway inhibition in cancer. *Proc. Natl. Acad. Sci. U. S. A.* (2009) doi:10.1073/pnas.0907325106.

68. Kadara, H. *et al.* Whole-exome sequencing and immune profiling of early-stage lung adenocarcinoma with fully annotated clinical follow-up. *Ann. Oncol.* (2017) doi:10.1093/annonc/mdw436.

69. Calles, A. *et al.* Expression of PD-1 and Its Ligands, PD-L1 and PD-L2, in Smokers and Never Smokers with KRAS-Mutant Lung Cancer. in *Journal of Thoracic Oncology* (2015). doi:10.1097/JTO.0000000000000687.

70. Borghaei, H. *et al.* Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N. Engl. J. Med.* (2015) doi:10.1056/NEJMoa1507643.

71. Manchado, E. *et al.* A combinatorial strategy for treating KRAS-mutant lung cancer. *Nature* (2016) doi:10.1038/nature18600.

72. Luo, J. *et al.* A Genome-wide RNAi Screen Identifies Multiple Synthetic Lethal Interactions with the Ras Oncogene. *Cell* **137**, 835–848 (2009).

73. Scholl, C. *et al.* Synthetic Lethal Interaction between Oncogenic KRAS Dependency and STK33 Suppression in Human Cancer Cells. *Cell* (2009) doi:10.1016/j.cell.2009.03.017.

74. Barbie, D. A. *et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108–112 (2009).

75. Román, M. *et al.* KRAS oncogene in non-small cell lung cancer: Clinical perspectives on the treatment of an old target. *Mol. Cancer* **17**, 1–14 (2018).

76. Holliday, R. The inheritance of epigenetic defects. *Science (80-.).* (1987) doi:10.1126/science.3310230.

77. Peschansky, V. J. & Wahlestedt, C. Non-coding RNAs as direct and indirect modulators of epigenetic regulation. *Epigenetics* (2014) doi:10.4161/epi.27473.

78. Shi, Y. X., Sheng, D. Q., Cheng, L. & Song, X. Y. Current Landscape of Epigenetics in Lung Cancer: Focus on the Mechanism and Application. *J. Oncol.* **2019**, (2019).

79. Goll, M. G. & Bestor, T. H. EUKARYOTIC CYTOSINE METHYLTRANSFERASES. *Annu. Rev. Biochem.* (2005) doi:10.1146/annurev.biochem.74.010904.153721.

80. Pasculli, B., Barbano, R. & Parrella, P. Epigenetics of breast cancer: Biology and clinical implication in the era of precision medicine. *Seminars in Cancer Biology* (2018) doi:10.1016/j.semcancer.2018.01.007.

81. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* (1997) doi:10.1038/38444.

82. Albert, L. L. & Lehninger, A. L. *Lehninger principles of biochemistry / David L. Nelson, Michael M. Cox*. *Principles of biochemistry* (2005).

83. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* (2000) doi:10.1038/47412.

84. Rodríguez-Paredes, M. & Esteller, M. Cancer epigenetics reaches mainstream oncology. *Nat. Med.* **17**, 330–339 (2011).

85. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* (1997) doi:10.1038/38664.

86. Plass, C. *et al.* Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat. Rev. Genet.* **14**, 765–80 (2013).

87. Deaton, A. M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* (2011) doi:10.1101/gad.2037511.

88. Rauch, T. A. *et al.* High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. *Proc. Natl. Acad. Sci. U. S. A.* (2008) doi:10.1073/pnas.0710735105.

89. Daskalos, A. *et al.* Hypomethylation of retrotransposable elements correlates with genomic instability in non-small cell lung cancer. *Int. J. Cancer* (2009) doi:10.1002/ijc.23849.

90. Mehta, A., Dobersch, S., Romero-Olmedo, A. J. & Barreto, G. Epigenetics in lung cancer diagnosis and therapy. *Cancer Metastasis Rev.* (2015) doi:10.1007/s10555-015-9563-3.

91. Zöchbauer-Müller, S. *et al.* Aberrant promoter methylation of multiple genes in non-small cell lung cancers. *Cancer Res.* (2001).

92. Fraga, M. F. *et al.* Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* (2005)

doi:10.1038/ng1531.

93. Halkidou, K. *et al.* Upregulation and Nuclear Recruitment of HDACl in Hormone Refractory Prostate Cancer. *Prostate* (2004) doi:10.1002/pros.20022.

94. Ropero, S. & Esteller, M. The role of histone deacetylases (HDACs) in human cancer. *Mol. Oncol.* **1**, 19–25 (2007).

95. Mai, A. *et al.* Histone deacetylation in epigenetics: An attractive target for anticancer therapy. *Medicinal Research Reviews* (2005) doi:10.1002/med.20024.

96. Damiani, L. A. *et al.* Carcinogen-induced gene promoter hypermethylation is mediated by DNMT1 and causal for transformation of immortalized bronchial epithelial cells. *Cancer Res.* (2008) doi:10.1158/0008-5472.CAN-08-1276.

97. Sterlacci, W. *et al.* A comprehensive analysis of p16 expression, gene status, and promoter hypermethylation in surgically resected non-small cell lung carcinomas. *J. Thorac. Oncol.* (2011) doi:10.1097/JTO.0b013e3182295745.

98. Brock, M., Hooker, C., Ota-Machida, E., Han, Y. & Guo, M. DNA methylation markers and early recurrence in stage I lung cancer. *N Engl J Med* (2008).

99. Lin, R. K. *et al.* The tobacco-specific carcinogen NNK induces DNA methyltransferase 1 accumulation and tumor suppressor gene hypermethylation in mice and lung cancer patients. *J. Clin. Invest.* (2010) doi:10.1172/JCI40706.

100. O'Hagan, H. M. *et al.* Oxidative Damage Targets Complexes Containing DNA Methyltransferases, SIRT1, and Polycomb Members to Promoter CpG Islands. *Cancer Cell* (2011) doi:10.1016/j.ccr.2011.09.012.

101. Tew, B. Y. *et al.* Genome-wide DNA methylation analysis of KRAS mutant cell lines. *Sci. Rep.* **10**, 1–16 (2020).

102. Van Den Broeck, A. *et al.* Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. *Clin. Cancer Res.* (2008) doi:10.1158/1078-0432.CCR-08-0869.

103. Sasaki, H. *et al.* Histone deacetylase 1 mRNA expression in lung cancer. *Lung Cancer* (2004) doi:10.1016/j.lungcan.2004.03.021.

104. Minamiya, Y. *et al.* Expression of histone deacetylase 1 correlates with a poor prognosis in patients with adenocarcinoma of the lung. *Lung Cancer* (2011) doi:10.1016/j.lungcan.2011.02.019.

105. Minamiya, Y. *et al.* Strong expression of HDAC3 correlates with a poor prognosis in patients with adenocarcinoma of the lung. *Tumour Biol.* (2010) doi:10.1007/s13277-010-0066-0.

106. Stojanović, D., Nikić, D. & Lazarević, K. The level of nickel in smoker's blood and urine. *Cent. Eur. J. Public Health* (2004).

107. Zhou, X., Li, Q., Arita, A., Sun, H. & Costa, M. Effects of nickel, chromate, and arsenite on histone 3 lysine methylation. *Toxicol. Appl. Pharmacol.* (2009) doi:10.1016/j.taap.2009.01.009.

108. Cantone, L. *et al.* Inhalable metal-rich air particles and histone H3K4 dimethylation and H3K9 Acetylation in a Cross-sectional Study of Steel Workers. *Environ. Health Perspect.* (2011) doi:10.1289/ehp.1002955.

109. Ansari, J., Shackelford, R. E. & El-Osta1, H. Epigenetics in non-small cell lung cancer: From basics to therapeutics. *Transl. Lung Cancer Res.* **5**, 155–171 (2016).

110. Hatzimichael, E. & Crook, T. Cancer Epigenetics: New Therapies and New Challenges. *J. Drug Deliv.* (2013) doi:10.1155/2013/529312.

111. Tang, M., Xu, W., Wang, Q., Xiao, W. & Xu, R. Potential of DNMT and its Epigenetic Regulation for Lung Cancer Therapy. *Curr. Genomics* (2009) doi:10.2174/138920209788920994.

112. Jones, P. A. & Taylor, S. M. Cellular differentiation, cytidine analogs and DNA methylation. *Cell* (1980) doi:10.1016/0092-8674(80)90237-8.

113. Miyanaga, A. *et al.* Antitumor activity of histone deacetylase inhibitors in non-small cell lung cancer cells: Development of a molecular predictive model. *Mol. Cancer Ther.* (2008) doi:10.1158/1535-7163.MCT-07-2140.

114. Kim, H. R. *et al.* Trichostatin A induces apoptosis in lung cancer cells via simultaneous activation of the death receptor-mediated and mitochondrial pathway. *Exp. Mol. Med.* (2006) doi:10.1038/emm.2006.73.

115. Traynor, A. M. *et al.* Vorinostat (NSC# 701852) in Patients with relapsed non-small cell lung cancer; A wisconsin oncology network phase II study. *J. Thorac. Oncol.* (2009) doi:10.1097/JTO.0b013e3181952478.

116. Shieh, J. M. *et al.* Mitochondrial apoptosis and FAK signaling disruption by a

novel histone deacetylase inhibitor, HTPB, in antitumor and antimetastatic mouse models. *PLoS One* (2012) doi:10.1371/journal.pone.0030240.

117. Juergens, R. A. *et al.* Combination epigenetic therapy has efficacy in patients with refractory advanced non-small cell lung cancer. *Cancer Discov.* **1**, 598–607 (2011).

118. Witta, S. E. *et al.* Randomized phase II trial of erlotinib with and without entinostat in patients with advanced non-small-cell lung cancer who progressed on prior chemotherapy. *J. Clin. Oncol.* **30**, 2248–55 (2012).

119. Serresi, M. *et al.* Polycomb Repressive Complex 2 Is a Barrier to KRAS-Driven Inflammation and Epithelial-Mesenchymal Transition in Non-Small-Cell Lung Cancer. *Cancer Cell* **29**, 17–31 (2016).

120. Cao, W. *et al.* EZH2 Promotes Malignant Behaviors via Cell Cycle Dysregulation and Its mRNA Level Associates with Prognosis of Patient with Non-Small Cell Lung Cancer. *PLoS One* **7**, 1–10 (2012).

121. Tan, J. *et al.* Pharmacologic disruption of polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. *Genes Dev.* **21**, 1050–1063 (2007).

122. Kikuchi, J. *et al.* Epigenetic therapy with 3-deazaneplanocin A, an inhibitor of the histone methyltransferase EZH2, inhibits growth of non-small cell lung cancer cells. *Lung Cancer* **78**, 138–143 (2012).

123. Kondo, Y. *et al.* Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nat. Genet.* (2008) doi:10.1038/ng.159.

124. Fillmore, C. M. *et al.* EZH2 inhibition sensitizes BRG1 and EGFR mutant lung tumours to TopoII inhibitors. *Nature* **520**, 239–42 (2015).

125. Forde, P. M., Brahmer, J. R. & Kelly, R. J. New Strategies in Lung Cancer: Epigenetic Therapy for Non– Small Cell Lung Cancer. *Clin Cancer Res.* **20**, 2244–2248 (2014).

126. Huffman, K. & Martinez, E. D. Pre-Clinical Studies of Epigenetic Therapies Targeting Histone Modifiers in Lung Cancer. *Front. Oncol.* **3**, 235 (2013).

127. Serresi, M. *et al.* Ezh2 inhibition in Kras-driven lung cancer amplifies

inflammation and associated vulnerabilities. *J. Exp. Med.* **215**, jem.20180801 (2018).

128. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science (80-.).* (2007) doi:10.1126/science.1138140.

129. Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* (2015) doi:10.1038/nrmicro3569.

130. Shmakov, S. *et al.* Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* (2017) doi:10.1038/nrmicro.2016.184.

131. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology* (2017) doi:10.1016/j.mib.2017.05.008.

132. Pickar-Oliver, A. & Gersbach, C. A. The next generation of CRISPR–Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* **20**, 490–507 (2019).

133. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (80-.).* (2012) doi:10.1126/science.1225829.

134. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science (80-.).* (2013) doi:10.1126/science.1231143.

135. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science (80-.).* (2013) doi:10.1126/science.1232033.

136. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* (2009) doi:10.1099/mic.0.023960-0.

137. Wyman, C. & Kanaar, R. DNA Double-Strand Break Repair: All's Well that Ends Well. *Annu. Rev. Genet.* (2006) doi:10.1146/annurev.genet.40.110405.090451.

138. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* (2013) doi:10.1038/nprot.2013.143.

139. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science (80-.).* (2014) doi:10.1126/science.1247005.

140. Hart, T. *et al.* High-Resolution CRISPR Screens Reveal Fitness Genes and

Genotype-Specific Cancer Liabilities. *Cell* **163**, 1515–1526 (2015).

141. Sharma, S. & Petsalaki, E. Application of CRISPR-Cas9 based genome-wide screening approaches to study cellular signalling mechanisms. *Int. J. Mol. Sci.* **19**, (2018).

142. Ting, P. Y. *et al.* Guide Swap enables genome-scale pooled CRISPR–Cas9 screening in human primary cells. *Nat. Methods* (2018) doi:10.1038/s41592-018-0149-1.

143. Forbes, S. A. *et al.* COSMIC: Somatic cancer genetics at high-resolution. *Nucleic Acids Res.* (2017) doi:10.1093/nar/gkw1121.

144. KABIR, J., LOBO, M. & ZACHARY, I. Staurosporine induces endothelial cell apoptosis via focal adhesion kinase dephosphorylation and focal adhesion disassembly independent of focal adhesion kinase proteolysis. *Biochem. J.* (2002) doi:10.1042/bj20020665.

145. Winter, J. *et al.* CRISPRAnalyzeR: Interactive analysis, annotation and documentation of pooled CRISPR screens. *bioRxiv* (2017) doi:10.1101/109967.

146. dos Santos, E. O., Carneiro-Lobo, T. C., Aoki, M. N., Levantini, E. & Bassères, D. S. Aurora kinase targeting in lung cancer reduces KRAS-induced transformation. *Mol. Cancer* **15**, 1–14 (2016).

147. Mahoney, C. L. *et al.* LKB1/KRAS mutant lung cancers constitute a genetic subset of NSCLC with increased sensitivity to MAPK and mTOR signalling inhibition. *Br. J. Cancer* **100**, 370–375 (2009).

148. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* (2009) doi:10.1038/nprot.2008.211.

149. Xinjun, L. I. & Hemminki, K. Inherited predisposition to early onset lung cancer according to histological type. *Int. J. Cancer* (2004) doi:10.1002/ijc.20436.

150. Jiang, J. & Gu, Y. Homozygous deletion of PTEN and poor outcomes in EGFR mutant NSCLC patients with gefitinib therapy. in (2015). doi:10.1183/13993003.congress-2015.pa4841.

151. Pathway, W. S. Histone H1 Recruitment by CHD8 Is Essential for Suppression of the Wnt-beta-Catenin Signaling Pathway. *Mol. Cell. Biol.* 501–512 (2012)

doi:10.1128/MCB.06409-11.

152. Kita, Y. *et al.* The Autism-Related Protein CHD8 Cooperates with C/EBPβ to Regulate Adipogenesis. *Cell Rep.* **23**, 1988–2000 (2018).

153. Ho, L. & Crabtree, G. R. Chromatin remodelling during development. *Nature* (2010) doi:10.1038/nature08911.

154. Pongor, L. *et al.* A genome-wide approach to link genotype to clinical outcome by utilizing next generation sequencing and gene chip data of 6,697 breast cancer patients. *Genome Med.* (2015) doi:10.1186/s13073-015-0228-1.

155. Kim, M. S., Chung, N. G., Kang, M. R., Yoo, N. J. & Lee, S. H. Genetic and expressional alterations of CHD genes in gastric and colorectal cancers. *Histopathology* (2011) doi:10.1111/j.1365-2559.2011.03819.x.

156. Damaschke, N. A. *et al.* Frequent Disruption of Chromodomain Helicase DNA-Binding Protein 8 (CHD8) and Functionally Associated Chromatin Regulators in Prostate Cancer. *Neoplasia (United States)* (2014) doi:10.1016/j.neo.2014.10.003.

157. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* (1996) doi:10.1016/S0092-8674(00)82001-2.

158. Tropberger, P. *et al.* Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell* (2013) doi:10.1016/j.cell.2013.01.032.

159. Hatzi, K. *et al.* A hybrid mechanism of action for BCL6 in B cells defined by formation of functionally distinct complexes at enhancers and promoters. *Cell Rep.* (2013) doi:10.1016/j.celrep.2013.06.016.

160. Sweeney, S. M. *et al.* AACR project genie: Powering precision medicine through an international consortium. *Cancer Discov.* (2017) doi:10.1158/2159-8290.CD-17-0151.

161. Attar, N. & Kurdistani, S. K. Exploitation of EP300 and CREBBP Lysine Acetyltransferases by Cancer. (2017).

162. Yang, W. S. *et al.* Regulation of ferroptotic cancer cell death by GPX4. *Cell* (2014) doi:10.1016/j.cell.2013.12.010.

163. Tumber, A. *et al.* Potent and Selective KDM5 Inhibitor Stops Cellular Demethylation of H3K4me3 at Transcription Start Sites and Proliferation of MM1S Myeloma Cells. *Cell Chem. Biol.* **24**, 371–380 (2017).

164. Belužić, L. *et al.* Knock-down of AHCY and depletion of adenosine induces DNA damage and cell cycle arrest. *Sci. Rep.* **8**, 1–16 (2018).

165. Lee, H. O. *et al.* S-adenosylhomocysteine hydrolase over-expression does not alter S-adenosylmethionine or S-adenosylhomocysteine levels in CBS deficient mice. *Mol. Genet. Metab. Reports* **15**, 15–21 (2018).

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

167. Tanaka, I. *et al.* eIF2β, a subunit of translation-initiation factor EIF2, is a potential therapeutic target for non-small cell lung cancer. *Cancer Sci.* (2018) doi:10.1111/cas.13602.

168. Aljohani, H. M. *et al.* Genetic mutations associated with lung cancer metastasis to the brain. *Mutagenesis* (2018) doi:10.1093/mutage/gey003.

169. Hieshima, K. *et al.* Molecular cloning of a novel human CC chemokine liver and activation- regulated chemokine (LARC) expressed in liver. Chemotactic activity for lymphocytes and gene localization on chromosome 2. *J. Biol. Chem.* (1997) doi:10.1074/jbc.272.9.5846.

170. Chen, W. *et al.* CCL20 triggered by chemotherapy hinders the therapeutic efficacy of breast cancer. *PLoS Biol.* **16**, 1–27 (2018).

171. Wang, G. Z. *et al.* Tobacco smoke induces production of chemokine CCL20 to promote lung cancer. *Cancer Lett.* **363**, 60–70 (2015).

172. Dang, Y. *et al.* Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol.* (2015) doi:10.1186/s13059-015-0846-3.

173. Shalem, O., Sanjana, E. N., Hartenian, E. & Zhang, F. Genome-Scale CRISPR-Cas9 Knockout. *Science (80-.).* **343**, 84–88 (2014).

174. Yau, E. H. & Rana, T. M. Next-generation sequencing of genome-wide CRISPR screens. in *Methods in Molecular Biology* (2018). doi:10.1007/978-1-4939-7514-3_13.

175. Wang, J. *et al.* Suppression of KRas-mutant cancer through the combined inhibition of KRAS with PLK1 and ROCK. *Nat. Commun.* **7**, 1–13 (2016).

176. Singh, A. *et al.* TAK1 inhibition promotes apoptosis in KRAS-dependent colon cancers. *Cell* (2012) doi:10.1016/j.cell.2011.12.033.

177. dos Santos, E. O., Carneiro-Lobo, T. C., Aoki, M. N., Levantini, E. & Bassères, D. S. Aurora kinase targeting in lung cancer reduces KRAS-induced transformation. *Mol. Cancer* (2016) doi:10.1186/s12943-016-0494-6.

178. McDonald, E. R. *et al.* Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening. *Cell* (2017) doi:10.1016/j.cell.2017.07.005.

179. Heavey, S. *et al.* In pursuit of synergy: An investigation of the PI3K/mTOR/MEK co-targeted inhibition strategy in NSCLC. *Oncotarget* (2016) doi:10.18632/oncotarget.12755.

180. Spoerke, J. M. *et al.* Phosphoinositide 3-kinase (PI3K) pathway alterations are associated with histologic subtypes and are predictive of sensitivity to PI3K inhibitors in lung cancer preclinical models. *Clin. Cancer Res.* (2012) doi:10.1158/1078-0432.CCR-12-2347.

181. Miranda, T. B. *et al.* DZNep is a global histone methylation inhibitor that reactivates developmental genes not silenced by DNA methylation. *Mol. Cancer Ther.* **8**, 1579–1588 (2009).

182. Shi, B. *et al.* Oncogenic enhancer of zeste homolog 2 is an actionable target in patients with non-small cell lung cancer. *Cancer Med.* **8**, 6383–6392 (2019).

183. Lee, J. K. & Kim, K. C. DZNep, inhibitor of S-adenosylhomocysteine hydrolase, down-regulates expression of SETDB1 H3K9me3 HMTase in human lung cancer cells. *Biochem. Biophys. Res. Commun.* **438**, 647–652 (2013).

184. Akpa, C. A. *et al.* Acquired resistance to DZNep-mediated apoptosis is associated with copy number gains of AHCY in a B-cell lymphoma model. *BMC Cancer* **20**, 1–12 (2020).

185. Huang, S., Litt, M. & Felsenfeld, G. Methylation of histone H4 by arginine methyltransferase PRMT1 is essential in vivo for many subsequent histone modifications. *Genes Dev.* (2005) doi:10.1101/gad.1333905.

186. Avasarala, S. *et al.* PRMT1 Is a novel regulator of epithelial-mesenchymal-transition in non-small cell lung cancer. *J. Biol. Chem.* **290**, 13479–13489 (2015).

187. Chen, Y. *et al.* Lung Cancer Therapy Targeting Histone Methylation: Opportunities and Challenges. *Comput. Struct. Biotechnol. J.* **16**, 211–223 (2018).

188. Chang, S., Yim, S. & Park, H. The cancer driver genes IDH1/2, JARID1C/ KDM5C, and UTX/ KDM6A: crosstalk between histone demethylation and hypoxic reprogramming in cancer metabolism. *Exp. Mol. Med.* **51**, (2019).

189. Wu, Q. *et al.* In vivo CRISPR screening unveils histone demethylase UTX as an important epigenetic regulator in lung tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E3978–E3986 (2018).

190. Mills, A. A. The CHD chromatin remodelers: Family traits that protect and promote cancer. 1–20 (2017) doi:10.1101/cshperspect.a026450.The.

191. Hou, X. *et al.* P300 Promotes Proliferation, Migration, and Invasion Via Inducing Epithelial-Mesenchymal Transition in Non-Small Cell Lung Cancer Cells. *BMC Cancer* **18**, 1–8 (2018).

192. Caldon, C. E. *et al.* Estrogen Regulation of Cyclin E2 Requires Cyclin D1 but Not c-Myc. *Mol. Cell. Biol.* (2009) doi:10.1128/mcb.00269-09.

193. Tahara, T. *et al.* Fusobacterium in colonic flora and molecular features of colorectal carcinoma. *Cancer Res.* (2014) doi:10.1158/0008-5472.CAN-13-1865.

194. Koshiishi, N. *et al.* p300 gene alterations in intestinal and diffuse types of gastric carcinoma. *Gastric Cancer* (2004) doi:10.1007/s10120-004-0273-8.

195. Gayther, S. A. *et al.* Mutations truncating the EP300 acetylase in human cancers. *Nat. Genet.* (2000) doi:10.1038/73536.

196. Zhang, C. *et al.* p300 expression repression by hypermethylation associated with tumour invasion and metastasis in oesophageal squamous cell carcinoma. *J. Clin. Pathol.* (2007) doi:10.1136/jcp.2006.044099.

197. Li, Y. *et al.* High expression of p300 has an unfavorable impact on survival in resectable esophageal squamous cell carcinoma. *Ann. Thorac. Surg.* (2011) doi:10.1016/j.athoracsur.2010.12.012.

198. Li, M. *et al.* High expression of transcriptional coactivator p300 correlates with

aggressive features and poor prognosis of hepatocellular carcinoma. *J. Transl. Med.* (2011) doi:10.1186/1479-5876-9-5.

199. Debes, J. D. *et al.* p300 in Prostate Cancer Proliferation and Progression. *Cancer Res.* (2003).

200. Hou, X. *et al.* High expression of the transcriptional co-activator p300 predicts poor survival in resectable non-small cell lung cancers. *Eur. J. Surg. Oncol.* (2012) doi:10.1016/j.ejso.2012.02.180.

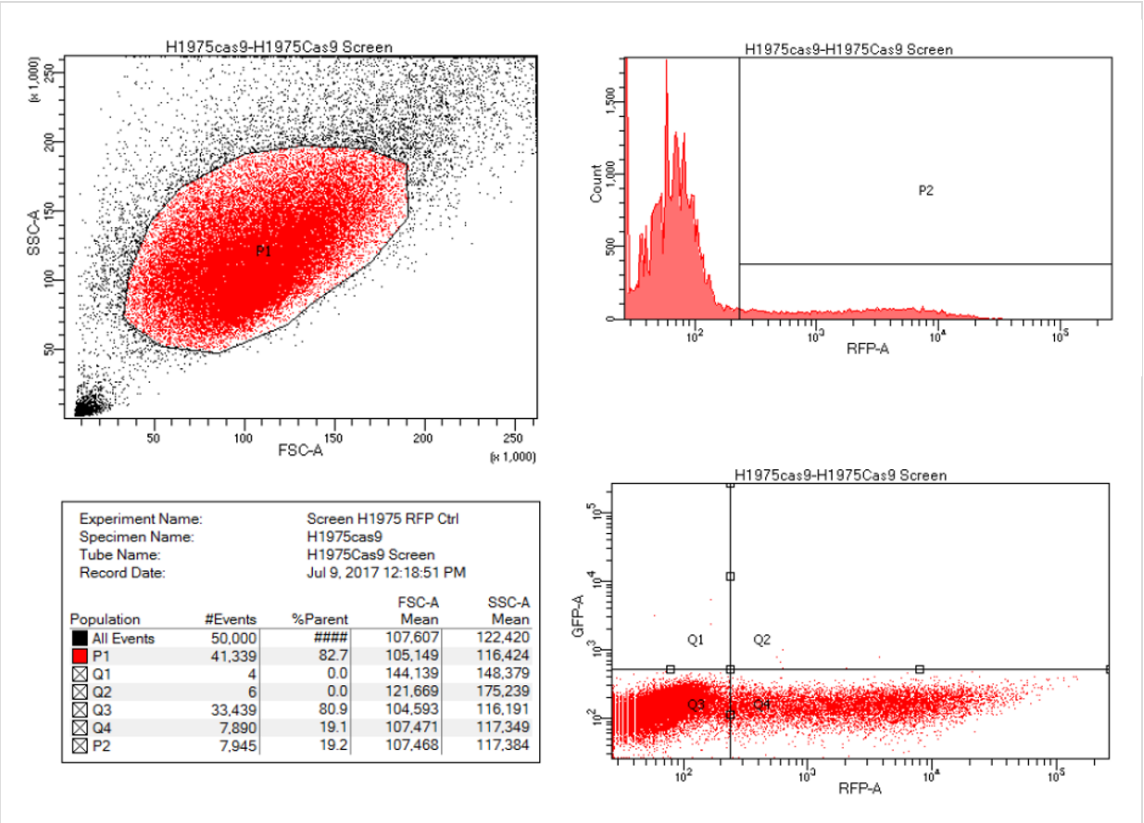
201. Wang, S. A. *et al.* Phosphorylation of p300 increases its protein degradation to enhance the lung cancer progression. *Biochim. Biophys. Acta - Mol. Cell Res.* (2014) doi:10.1016/j.bbamcr.2014.02.001.

202. Barrera-Rodriguez, R. Importance of the Keap1-Nrf2 pathway in NSCLC: Is it a possible biomarker? (Review). *Biomed. Reports* 375–382 (2018) doi:10.3892/br.2018.1143.

203. Rivas-Fuentes, S. *et al.* Role of chemokines in non-small cell lung cancer: Angiogenesis and inflammation. *J. Cancer* **6**, 938–952 (2015).

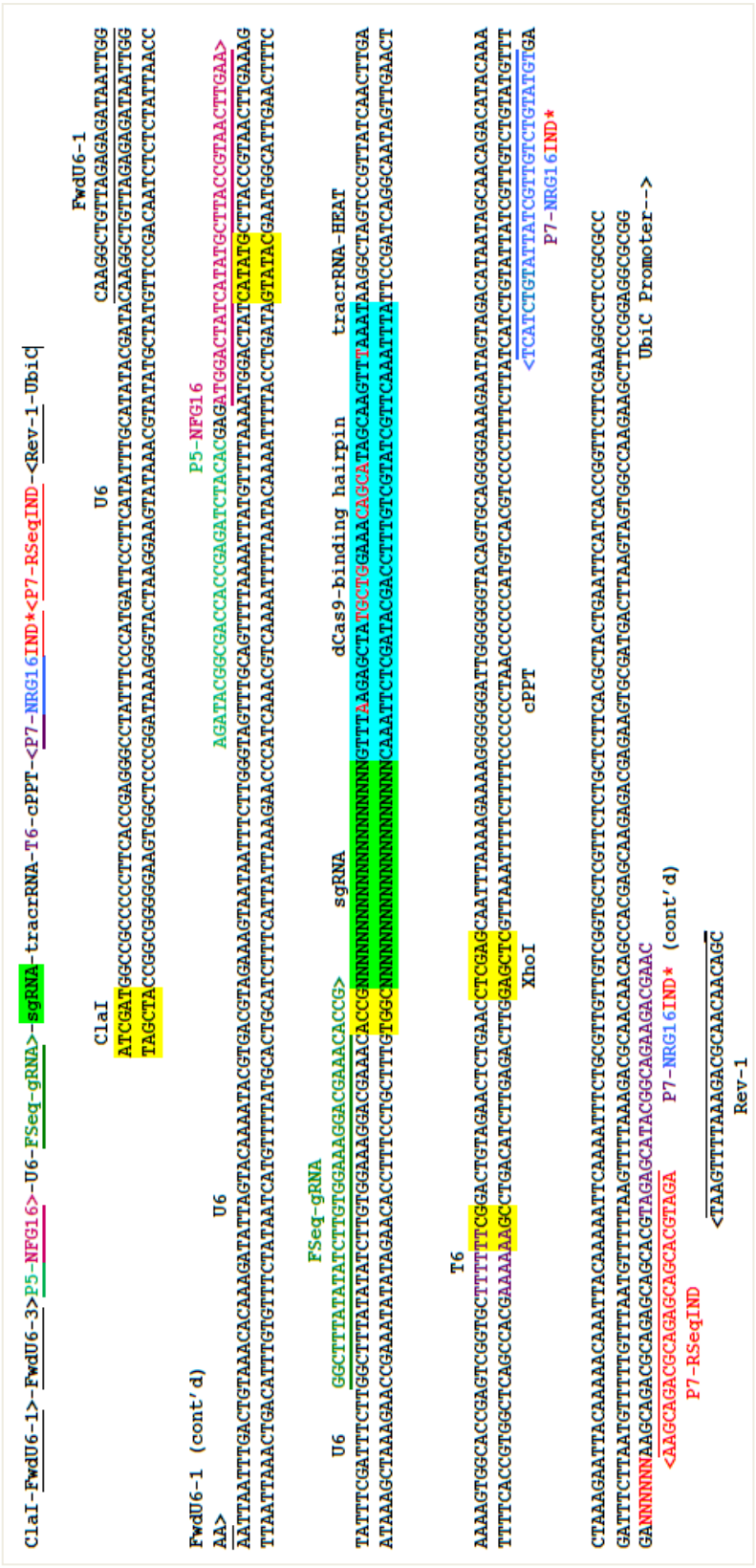
204. Ranasinghe, R. & Eri, R. Modulation of the CCR6-CCL20 Axis: A Potential Therapeutic Target in Inflammation and Cancer. *Medicina (B. Aires).* **54**, 88 (2018).

6 Supplementary data

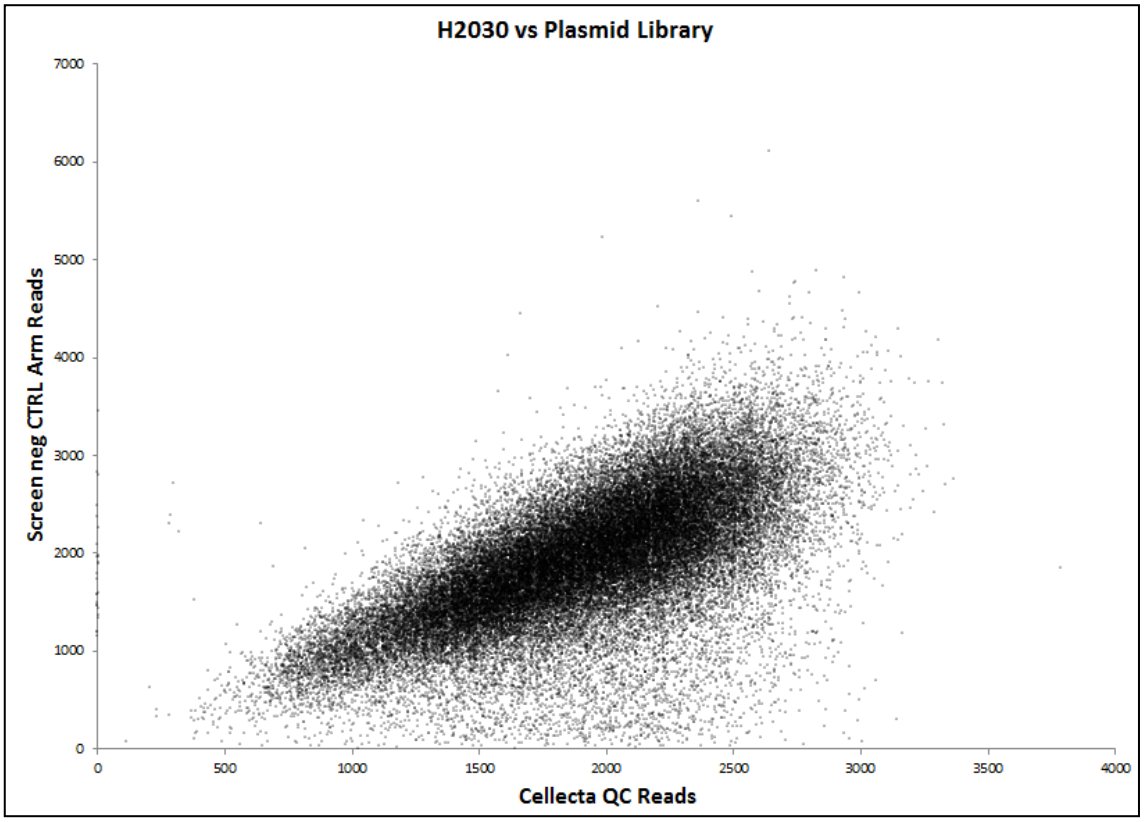


Supplementary figure 1 | Determination of the fraction of transduced H1975_Cas9 cells by flow cytometry

The fraction of RFP(+) cells was assessed 72 hours post-transduction with a library of pooled lentiviral particles.

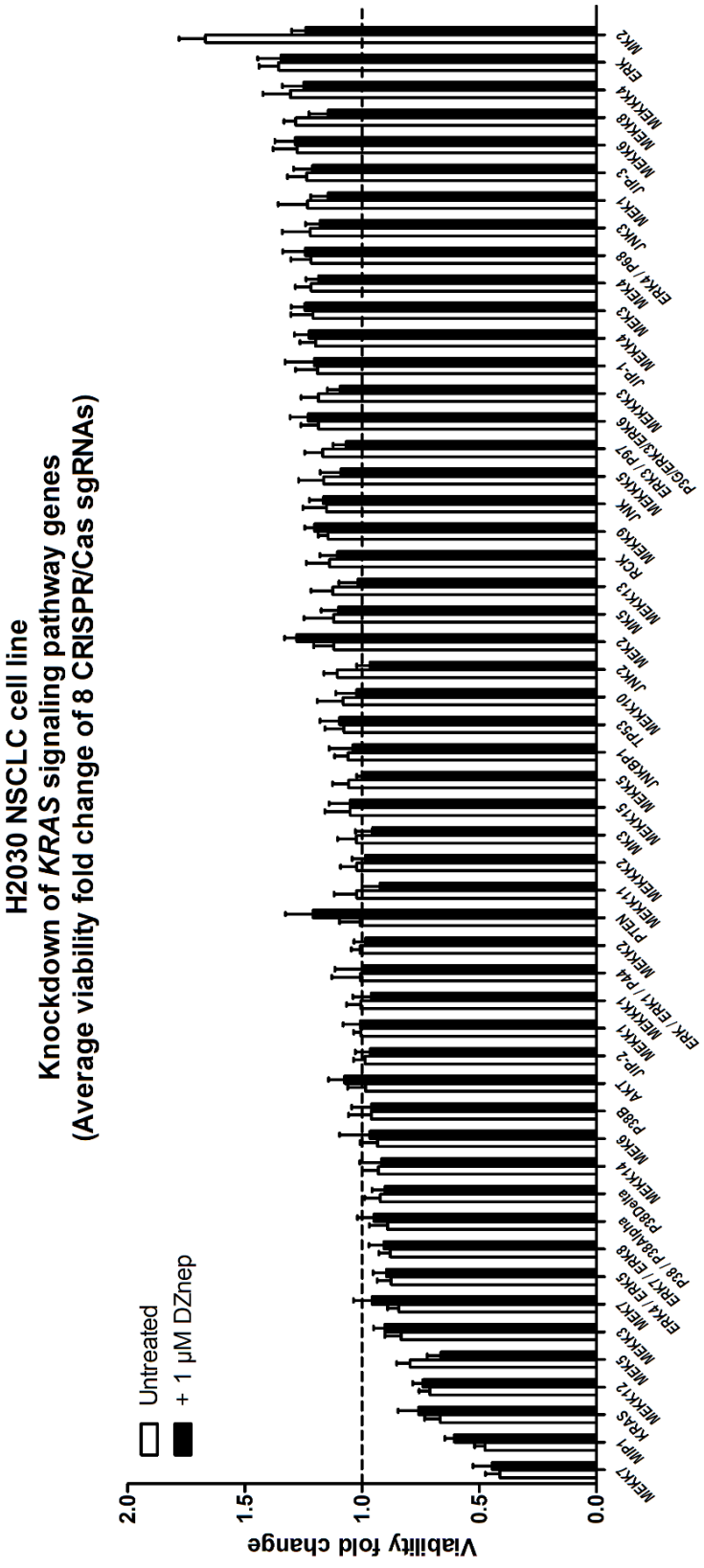


Supplementary figure 2 | Sequences of sgRNA flanking primers



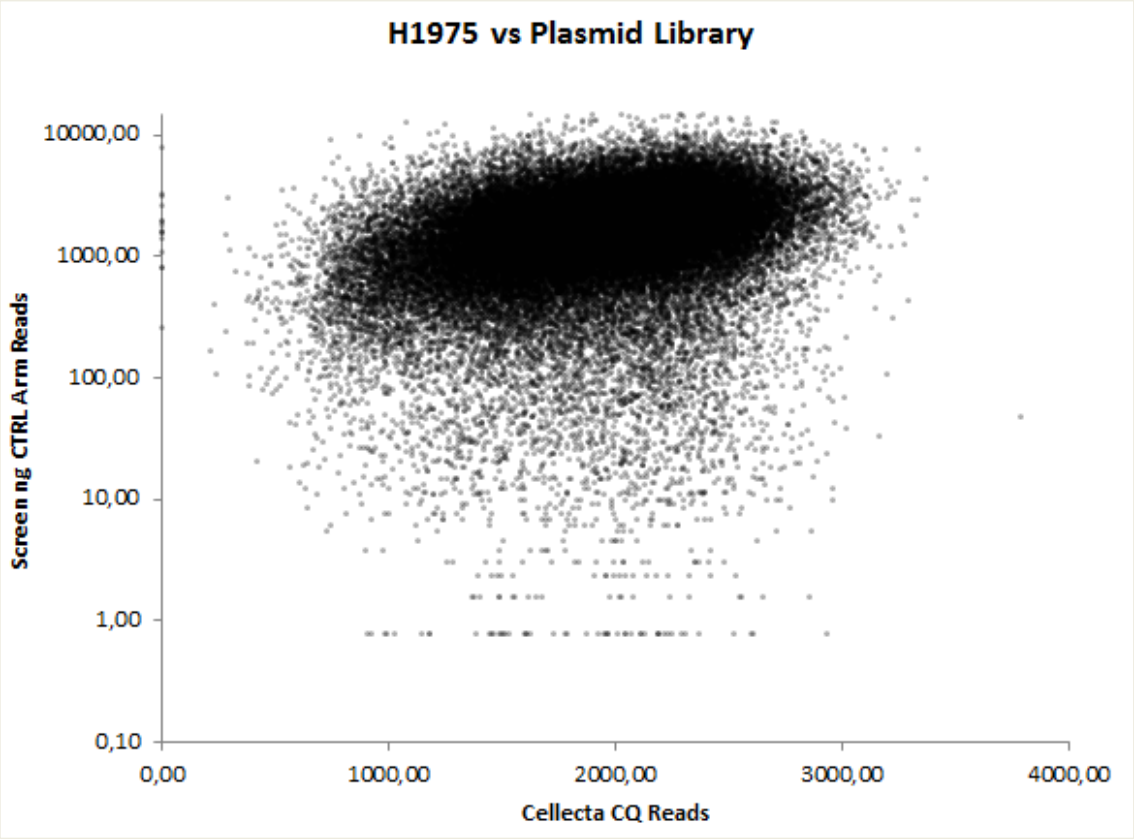
Supplementary figure 3 | Scatterplot representation of readcounts of the control / no-drug arm of H2030 cells vs. the lentiviral library

Both samples were normalized to a sum of 100 million reads per sample.



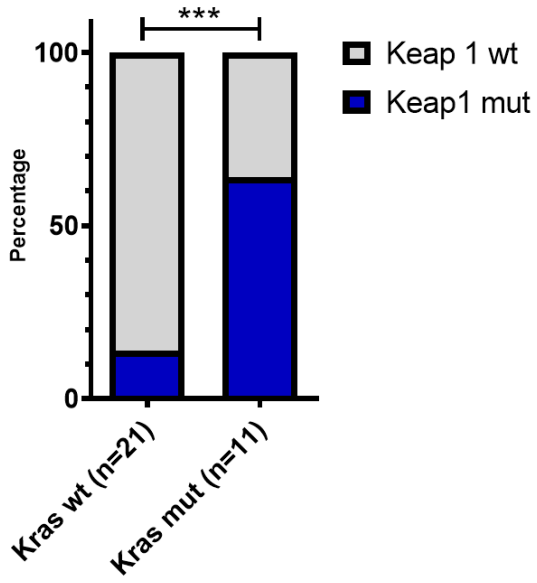
Supplementary figure 4 | Cell viability upon knockout of genes associated with KRAS signaling in H2030_Cas9 cells

Genes were knocked out using a pooled lentiviral CRISPR/Cas9 library and split into two arms. Cells of one arm were treated without adding a drug, the cells of the other arm were treated with 1µM DZNep for 14 days. Cell viability was calculated based on the abundance of sgRNAs for each gene.



Supplementary figure 5 | Scatterplot representation of readcounts of the control / no-drug arm of H1975 cells vs. the lentiviral library

Both samples were normalized to 100 million reads.



Supplementary figure 6 | *In silico* analysis of NSCLC cell lines for KEAP1 mutations

The *KRAS* and *KEAP1* mutation status of NSCLC 32 cell lines was assessed in the COSMIC database[143], and the frequency of *KEAP1* mutations was correlated with the *KRAS* mutation status using a student's t-test, showing a significant correlation of *KEAP1* mutations with the *KRAS* mutation status. The panel included the following cell lines: H2030, H1944, H1299, H1975 ,HCC827, H2228, H522, H1563, H1650, A549, H358, H2122, H3122, H1395, H1355, H1568, H1755, H1993, H23, H647, H838, H1435, H1651, H1693, H1703, H1792, H1793, H1838, H2009, H2291, H2342, and H2405.

Supplementary table 1 | 242 genes that were identified as essential to cell viability in H2030_Cas9 cells through comparison of the normalized screen control arm to the normalized vector library readcounts. The values underneath the different algorithms indicate the respective p-values.

Gene	Ratio	Wilcox	DESeq2	MAGeCK	sgRSEA	EdgeR
CDC16	0.073685734	0.000689651		4.1E-05	0.000144927	0.00821
PCNA	0.076267528	0.000689651		4.1E-05	0.000144927	0.00821
KPNB1	0.08331604	0.000933042		4.1E-05	0.000144927	0.00821
PSMB7	0.083687462	0.000689651		4.1E-05	0.000144927	0.00821
PLK1	0.124384996	0.000689651		4.1E-05	0.000144927	0.00821
DUT	0.14286629	0.000689651		4.1E-05	0.000144927	0.00821
CDK1	0.154280357	0.000689651		4.1E-05	0.000144927	0.00821
HSPA5	0.15513905	0.000689651		4.1E-05	0.000144927	0.00821
PSMA5	0.155531825	0.000689651		4.1E-05	0.000144927	0.00821
HSPE1	0.156314758	0.000689651		4.1E-05	0.000144927	0.00821
RPL34	0.164253129	0.000689651		4.1E-05	0.000144927	0.00821
CCNA2	0.168490464	0.000689651		4.1E-05	0.000144927	0.00821
PSMA6	0.182241219	0.000689651		4.1E-05	0.000144927	0.00821
SARS	0.182722662	0.000689651		4.1E-05	0.000144927	0.00821
ATP6V0C	0.183031514	0.000689651		4.1E-05	0.000144927	0.015539432
POLR2I	0.195317893	0.000689651		4.1E-05	0.000144927	0.00821
SF3B3	0.198729758	0.000689651		4.1E-05	0.000144927	0.019927184
PSMA3	0.200379477	0.000689651		4.1E-05	0.000144927	0.00821
CDC7	0.201325138	0.000689651		4.1E-05	0.000144927	0.00821
IARS	0.202915812	0.000689651		4.1E-05	0.000144927	0.00821
RPS15A	0.20508254	0.000689651		4.1E-05	0.000144927	0.00821
RPL5	0.207821872	0.001219509		4.1E-05	0.000144927	0.00821
AURKA	0.212507278	0.000689651		4.1E-05	0.000144927	0.00821
MED14	0.213356912	0.000689651		4.1E-05	0.000144927	0.00821
NFS1	0.213903656	0.000689651		4.1E-05	0.000144927	0.00821
HMGCS1	0.214319195	0.000689651		4.1E-05	0.000144927	0.00821
PSMB3	0.215992985	0.000689651		4.1E-05	0.000144927	0.019927184
YARS	0.216136915	0.000689651		4.1E-05	0.000144927	0.00821
TUBB	0.228040127	0.009088263		4.1E-05	0.000144927	0.00821
PSMD1	0.237980473	0.000689651		4.1E-05	0.000144927	0.015539432
RPL35A	0.238415973	0.000689651		4.1E-05	0.000144927	0.00821
POLR2D	0.239848527	0.000689651		4.1E-05	0.000144927	0.00821
DHDDS	0.248712855	0.00541155		4.1E-05	0.000144927	0.00821
SOD1	0.256107446	0.000933042		4.1E-05	0.000144927	0.00821
PAFAH1B1	0.256827408	0.00190903		4.1E-05	0.000144927	0.015539432
PMPCB	0.258761708	0.000689651		4.1E-05	0.000144927	0.015539432
UBE2M	0.264524637	0.000689651		4.1E-05	0.000144927	0.00821
ALDOA	0.264552429	0.000689651		4.1E-05	0.000144927	0.00821
HSPA9	0.268099554	0.006454656		4.1E-05	0.000144927	0.00821
POLR2L	0.272028733	0.002730776		4.1E-05	0.000144927	0.00821
RPS14	0.273191533	0.001219509		4.1E-05	0.000144927	0.00821
ANAPC2	0.274389089	0.006253342		4.1E-05	0.000144927	0.00821

MCM2	0.275764408	0.000689651		4.1E-05	0.000144927	0.019927184
PMPCA	0.27858626	0.000911987		4.1E-05	0.000144927	0.00821
EIF2S1	0.281100019	0.000933042		4.1E-05	0.000144927	0.00821
RPL9	0.285992248	0.003788718		4.1E-05	0.000144927	0.00821
GAPDH	0.287716898	0.001402758		4.1E-05	0.000144927	0.00821
MAT2A	0.288711365	0.001049654		4.1E-05	0.000144927	0.00821
UBE2I	0.292364291	0.001212299		4.1E-05	0.000144927	0.00821
GNB2L1	0.295864243	0.00146918		4.1E-05	0.000144927	0.00821
PSMA1	0.296402626	0.002756171		4.1E-05	0.000144927	0.00821
TPX2	0.297762873	0.000689651		4.1E-05	0.000144927	0.015539432
FARSA	0.297834711	0.000740109		4.1E-05	0.000144927	0.00821
LONP1	0.301465545	0.001723888		4.1E-05	0.000144927	0.00821
TRMT112	0.305803169	0.000689651		4.1E-05	0.000144927	0.015539432
SNRPE	0.307510777	0.00167248		4.1E-05	0.000144927	0.00821
RRM1	0.311746097	0.009993407		4.1E-05	0.000144927	0.00821
ALG11	0.315843467	0.001603742		4.1E-05	0.000144927	0.00821
ATP6V1A	0.320663641	0.001271908		4.1E-05	0.000144927	0.015539432
GTF2B	0.322519987	0.001299305		4.1E-05	0.000144927	0.00821
PSMB4	0.322729209	0.000985737		4.1E-05	0.000144927	0.00821
CENPE	0.32370667	0.001461736		4.1E-05	0.000144927	0.00821
NAA10	0.325716932	0.000933042		4.1E-05	0.000144927	0.00821
CCT8	0.329095814	0.001312346		4.1E-05	0.000144927	0.00821
PHB2	0.334971052	0.012402517		4.1E-05	0.000144927	0.015539432
ABCE1	0.341360014	0.005434518		4.1E-05	0.000144927	0.00821
ATP6V0B	0.343666679	0.002411219		4.1E-05	0.000144927	0.015539432
ATP2A2	0.344353742	0.004270504		4.1E-05	0.000144927	0.00821
ACOX1	0.344562926	0.00209373		4.1E-05	0.000144927	0.00821
LOC100996337	0.344821934	0.002202567		4.1E-05	0.000144927	0.00821
CRKL	0.346535337	0.001366408		4.1E-05	0.000144927	0.00821
RBM14	0.347068033	0.006524006		4.1E-05	0.000144927	0.015539432
GOSR2	0.348754316	0.006253342		4.1E-05	0.000144927	0.00821
POLR2H	0.350775537	0.002077754		4.1E-05	0.000144927	0.00821
ACADL	0.352150918	0.000874356		4.1E-05	0.000144927	0.00821
HGS	0.355568873	0.000898804		4.1E-05	0.000144927	0.00821
ATP1A1	0.35608501	0.007150393		4.1E-05	0.000144927	0.00821
COPB2	0.360035931	0.002077754		0.00012	0.000144927	0.00821
NUDT21	0.362859847	0.004185776		4.1E-05	0.000144927	0.00821
PSMC4	0.368606279	0.001995644	4.5199E-287	4.1E-05	0.000144927	0.00821
SBDS	0.369854092	0.00557259		4.1E-05	0.000144927	0.015539432
PSMA7	0.372398435	0.010184941		4.1E-05	0.000144927	0.015539432
WWTR1	0.374216231	0.000898804		4.1E-05	0.000144927	0.00821
NDUFAB1	0.37756625	0.002077754		4.1E-05	0.000144927	0.00821
ITGAV	0.377805077	0.00642046		4.1E-05	0.000144927	0.015539432
AATK	0.38157726	0.000933042		4.1E-05	0.000144927	0.015539432
GGPS1	0.382032131	0.00120914	6.0642E-289	0.00027	0.000144927	0.00821
SF3A3	0.385114204	0.017344371		4.1E-05	0.000144927	0.00821

PSMB2	0.385204605	0.006253342		4.1E-05	0.000144927	0.00821
HSCB	0.385494794	0.001163288	1.0995E-247	4.1E-05	0.000144927	0.00821
EIF3B	0.394067757	0.002411219		4.1E-05	0.000144927	0.00821
TARS	0.394790293	0.001312346	8.4508E-301	4.1E-05	0.000144927	0.00821
SLC25A10	0.398161038	0.001163288		4.1E-05	0.000144927	0.00821
PYCR1	0.398829378	0.001040527	5.6118E-303	4.1E-05	0.000144927	0.00821
PRPF31	0.401399253	0.006063409		0.00062	0.000144927	0.00821
QARS	0.401691665	0.00239183	2.0777E-298	4.1E-05	0.000144927	0.00821
POLR2B	0.401988529	0.001312346	4.8303E-294	0.00068	0.000144927	0.015539432
KIF11	0.40246405	0.003555569	9.944E-249	0.00712	0.000578033	0.00821
LARS	0.402697436	0.028692044	3.8259E-268	0.00078	0.000578033	0.015539432
ATP6V1E1	0.403041341	0.012589609	1.9822E-264	0.00084	0.000144927	0.00821
CDK9	0.408140832	0.002077754		0.00177	0.000144927	0.00821
NARS	0.408299308	0.002609183	5.4667E-244	4.1E-05	0.000144927	0.015539432
P4HB	0.413493448	0.001040527		0.00062	0.000144927	0.00821
ATP6V1B2	0.414388311	0.00871596	5.2272E-276	0.00049	0.000144927	0.015539432
CHEK1	0.415843209	0.00288345	2.81E-291	0.00078	0.000144927	0.015539432
TLN1	0.417553099	0.001415008	3.7487E-285	0.00056	0.000144927	0.00821
PHB	0.417916229	0.006092675	2.585E-285	4.1E-05	0.000144927	0.00821
MTOR	0.418037338	0.001484858	8.39E-297	4.1E-05	0.000533332	0.00821
GTF3C4	0.418397258	0.001995644	8.0615E-270	4.1E-05	0.000144927	0.00821
GTF2H4	0.420037222	0.00557259		0.00035	0.000144927	0.00821
PYROXD1	0.421765937	0.002059717	1.9445E-240	0.00817	0.000533332	0.019927184
KAT8	0.422557628	0.002681071	1.7287E-250	0.00472	0.000144927	0.00821
HYOU1	0.423123035	0.004780237	2.6758E-267	4.1E-05	0.000144927	0.015539432
HSD17B4	0.423260432	0.00120914	1.2912E-279	0.00152	0.000144927	0.00821
ECT2	0.42440699	0.001212299	4.2245E-292	0.00166	0.000578033	0.00821
GPS1	0.424603955	0.0022955	1.6215E-286	0.00078	0.000144927	0.00821
SIRT7	0.425141204	0.012495455	4.7017E-288	4.1E-05	0.000144927	0.00821
PGAM1	0.425930945	0.001995644	1.4664E-263	0.0012	0.000144927	0.00821
CCNH	0.426082805	0.001896885	7.0591E-297	0.00084	0.000144927	0.00821
MVK	0.428259274	0.003098623	3.4447E-286	0.00115	0.000144927	0.00821
FXN	0.429549449	0.007150393	1.2699E-261	0.00149	0.000533332	0.00821
XPO1	0.430088199	0.011599919	6.0503E-210	0.00078	0.000144927	0.00821
MLLT6	0.433473395	0.001245314	7.7066E-240	0.00088	0.000144927	0.015539432
CDK12	0.433842468	0.015182812	2.6114E-267	0.00274	0.000144927	0.015539432
LAMTOR2	0.433956244	0.00810869	2.9153E-235	4.1E-05	0.000144927	0.019927184
PCYT2	0.437096009	0.002009476	1.2976E-237	4.1E-05	0.000144927	0.015539432

TAMM41	0.439912419	0.002354874	4.3296E-223	0.00027	0.000578033	0.019927184
CCNK	0.443177183	0.001461736	4.3098E-236	0.00337	0.000578033	0.00821
POLR2E	0.445891957	0.003092527	6.4482E-171	0.00152	0.000144927	0.023529171
TADA2A	0.447064827	0.001219509	1.4437E-268	0.00165	0.000144927	0.00821
KRT17	0.449105457	0.001371752	4.7505E-250	4.1E-05	0.000144927	0.00821
PSMD2	0.449619783	0.01231079	4.7874E-213	4.1E-05	0.000144927	0.00821
PDE6G	0.450638442	0.001421544	2.1954E-249	0.00165	0.000533332	0.00821
EPRS	0.451936325	0.011599919	4.0084E-191	4.1E-05	0.000533332	0.015539432
ING3	0.452433929	0.002730776	3.6932E-261	0.00166	0.000283687	0.00821
TRMT5	0.453542622	0.00249809	2.0649E-214	0.0052	0.000578033	0.00821
ALG14	0.45623734	0.006997447	3.7102E-246	0.00093	0.000144927	0.00821
DNAJA3	0.456449891	0.003624105	2.4806E-184	0.00312	0.000144927	0.00821
CLP1	0.457303978	0.016731707	4.2429E-226	4.1E-05	0.000144927	0.00821
NOP2	0.459597236	0.001415008	1.4493E-239	0.00049	0.000533332	0.00821
CDC6	0.462965272	0.002077754	5.2253E-214	0.00595	0.000681816	0.00821
GTF2H2	0.463751269	0.004227869	1.1732E-192	4.1E-05	0.000851061	0.00821
CDK2	0.465733132	0.001548072	1.4851E-208	0.01798	0.002690575	0.00821
ATP6V1G1	0.466596469	0.006352561	2.0695E-196	4.1E-05	0.000144927	0.00821
RRM2	0.468212426	0.010184941	1.4485E-199	4.1E-05	0.000144927	0.00821
SMN1	0.468213597	0.005722799	3.0041E-217	0.00088	0.000144927	0.00821
WEE1	0.470067736	0.015182812	4.6233E-231	0.0056	0.000533332	0.00821
NARFL	0.471237323	0.006049076	1.4065E-186	0.01604	0.000851061	0.00821
MRP63	0.472160221	0.003968725	1.2258E-206	0.00166	0.00163461	0.015539432
RPTOR	0.474503836	0.0452222	5.6844E-206	4.1E-05	0.000144927	0.00821
ALDOC	0.475129539	0.001603742	4.6762E-184	0.01399	0.002037031	0.00821
AMY1A	0.475546214	0.006324598	7.8533E-203	0.00078	0.001293528	0.00821
KRAS	0.47766351	0.003968725	3.7236E-172	0.00149	0.000681816	0.015539432
DAD1	0.478309783	0.03743749	1.6386E-211	4.1E-05	0.000144927	0.00821
DDX12P	0.479017515	0.021728503	3.1493E-220	0.00088	0.000144927	0.00821
SACM1L	0.480358099	0.013126063	6.0204E-169	0.00182	0.000578033	0.00821
ATR	0.480999935	0.003784099	1.6678E-202	0.0023	0.000578033	0.015539432
VARS	0.486410944	0.035274807	5.9068E-194	0.0014	0.000283687	0.00821
ALG2	0.486536307	0.017434139	5.7882E-209	0.00078	0.000144927	0.00821
HUWE1	0.487014766	0.015428473	2.7124E-184	0.00169	0.000851061	0.00821
TIMM23	0.487560529	0.003011782	3.975E-172	0.00166	0.000851061	0.00821
IARS2	0.487853206	0.017434139	9.6687E-	0.00219	0.000578033	0.019927184

			193			
PIGW	0.488733441	0.001603742	3.3235E-190	0.01455	0.002956513	0.015539432
DHRS11	0.488946077	0.002074372	1.3069E-164	0.02916	0.002690575	0.015539432
FARSB	0.490034435	0.00239183	7.8125E-189	0.00168	0.002956513	0.00821
DARS	0.490414129	0.005040322	1.6223E-157	0.00219	0.001030925	0.00821
AK6	0.491334268	0.008406178	2.887E-185	4.1E-05	0.002037031	0.015539432
RPN1	0.491976895	0.013816803	3.7221E-146	0.02768	0.000578033	0.019927184
NDUFB3	0.492620013	0.010372201	1.4503E-173	0.0014	0.001818176	0.015539432
FDX1L	0.49367544	0.005490773	1.193E-196	0.00203	0.001477828	0.00821
FTSJ3	0.493851843	0.006559659	3.0089E-191	4.1E-05	0.001386134	0.00821
KAT5	0.493982298	0.011307083	1.3425E-180	0.00249	0.000578033	0.019927184
PKMYT1	0.494224789	0.003968725	1.8474E-158	0.00152	0.000851061	0.015539432
TRIAP1	0.495660098	0.015182812	7.3956E-187	0.02328	0.002920345	0.00821
MAPKAP1	0.497723699	0.002411219	5.8311E-190	0.00579	0.000851061	0.00821
CDK6	0.498036084	0.009970377	1.625E-188	4.1E-05	0.000578033	0.026132626
PREB	0.501419351	0.013646904	1.6504E-184	0.00472	0.000578033	0.00821
MRPL45	0.501690345	0.016461589	5.287E-171	0.00177	0.001293528	0.015539432
CPS1	0.503165352	0.00802045	1.4164E-193	0.00035	0.000578033	0.00821
MMP28	0.503464229	0.005195476	7.4726E-199	0.00078	0.001030925	0.00821
RPS27A	0.50366828	0.018494998	1.6055E-194	0.00437	0.001030925	0.00821
CCNB1	0.504811118	0.007234981	1.0282E-161	0.00137	0.00163461	0.015539432
ASH2L	0.505022602	0.029194746	1.6087E-194	0.00056	0.000578033	0.00821
DUSP14	0.506259841	0.002277193	3.9053E-159	0.00817	0.000947366	0.023529171
EXOSC10	0.507273659	0.006269293	4.4897E-186	0.00434	0.00163461	0.00821
RNF20	0.510163924	0.002110003	4.4079E-170	0.01538	0.003404245	0.00821
POLR2G	0.513555625	0.032012466	1.3922E-185	0.00212	0.000578033	0.00821
CXXC1	0.515380106	0.005434518	4.2167E-152	0.00115	0.000851061	0.00821
PRELID1	0.515794686	0.004075841	2.9067E-156	0.01885	0.003749989	0.023529171
NPIPA5	0.516823289	0.012828754	1.672E-156	4.1E-05	0.000578033	0.015539432
DMAP1	0.51862805	0.013816803	4.2432E-156	0.00115	0.00121827	0.00821
EZH1	0.519308054	0.01945338	1.6467E-167	0.0276	0.008920836	0.00821
PIGS	0.519346018	0.002681071	1.5076E-175	0.0023	0.002110085	0.00821
ENO1	0.519901386	0.002411219	1.7254E-158	4.1E-05	0.001030925	0.00821
GTF2H3	0.522790367	0.00784701	3.6765E-144	0.03232	0.002956513	0.023529171
CACNB1	0.523153176	0.009904949	6.7039E-158	0.00093	0.000851061	0.015539432
LYZL6	0.524896273	0.002411219	3.4381E-129	0.01782	0.013198613	0.023529171
ITGB5	0.525706685	0.009088263	2.9422E-116	0.00145	0.003934414	0.019927184

TRRAP	0.527256125	0.011862459	1.3401E-168	0.00219	0.002690575	0.00821
IKBKAP	0.5275351	0.022492586	2.0056E-136	0.002	0.001122446	0.019927184
ORC1	0.527649018	0.00802045	7.6329E-131	0.0012	0.003474566	0.00821
ASNA1	0.527890047	0.002219971	1.1387E-174	0.0143	0.00163461	0.00821
DTYMK	0.5285842	0.032012466	2.8153E-139	0.00391	0.002956513	0.00821
ROMO1	0.530285326	0.005091396	1.7242E-146	0.00149	0.001293528	0.023529171
WBSCR22	0.535704742	0.042533804	6.6147E-156	0.00137	0.002037031	0.00821
UBA52	0.537968156	0.035274807	4.8991E-152	0.00203	0.000578033	0.00821
IK	0.542619999	0.013126063	9.1757E-123	0.02653	0.004096373	0.015539432
GJD3	0.544726905	0.013180437	1.081E-144	0.01934	0.001122446	0.00821
ERBB2	0.545317898	0.010488836	5.4099E-122	0.002	0.002454538	0.026132626
MRPL12	0.545413406	0.012220266	2.0751E-131	0.01009	0.007732318	0.015539432
MARS	0.545447301	0.006559659	3.7142E-132	0.00345	0.000851061	0.015539432
WARS	0.54683542	0.009108749	4.9247E-102	0.0263	0.005496166	0.023529171
SGK494	0.549167059	0.002678532	1.1432E-130	0.02403	0.008058584	0.00821
PGD	0.5528574	0.029865953	4.2625E-126	0.01109	0.003376613	0.015539432
GUK1	0.555872926	0.007314229	2.914E-144	0.01385	0.002755547	0.00821
GTPBP4	0.560066875	0.012130941	6.7998E-121	0.00186	0.002037031	0.00821
FBL	0.560444317	0.042250428	1.90235E-93	0.00479	0.002037031	0.023529171
KARS	0.560838109	0.022492586	2.4314E-127	0.02937	0.003801641	0.015539432
PDPK1	0.56576839	0.029748773	1.6537E-129	0.02565	0.004302776	0.00821
TOMM40	0.566261616	0.015182812	1.1387E-134	0.01047	0.002755547	0.00821
COASY	0.569431801	0.013587844	1.5972E-128	0.00169	0.01900923	0.00821
SARS2	0.569689248	0.003966193	4.4577E-118	0.01144	0.004669246	0.023529171
TTK	0.570699995	0.036064095	5.72286E-86	0.01277	0.008044256	0.015539432
NAE1	0.573387722	0.025777011	4.43837E-97	0.00706	0.001293528	0.023529171
NAGLU	0.574317966	0.021821122	1.9198E-119	0.01132	0.000851061	0.00821
TRNT1	0.575100711	0.024216871	1.1506E-106	0.02388	0.012499962	0.019927184
NAT10	0.57632692	0.013126063	1.3074E-103	0.0014	0.004603161	0.00821
RNMT	0.577110213	0.018133987	1.5382E-118	0.0032	0.007245261	0.00821
PLXDC1	0.579947451	0.018133987	1.408E-112	0.00493	0.003404245	0.00821
CCL13	0.581659429	0.02629588	1.14515E-73	0.01067	0.007518774	0.030713669
SMARCA5	0.583976767	0.00862055	5.5297E-99	0.02094	0.020611558	0.023529171
NAA30	0.586204474	0.025621379	1.469E-112	0.00068	0.002110085	0.00821
TBL3	0.587445988	0.029449821	5.3743E-101	0.00757	0.004645655	0.015539432
SYVN1	0.590223467	0.029449821	7.65506E-78	0.0239	0.013399959	0.034805486
PGS1	0.593936101	0.02719745	4.7102E-	0.01705	0.012499962	0.015539432

			112			
HCRT	0.596721425	0.021853233	5.0558E-89	0.01735	0.018695595	0.023529171
SDF2	0.599493923	0.013434724	5.5462E-110	0.01362	0.010459332	0.015539432
STARD7	0.601113951	0.020707032	2.02961E-88	0.01735	0.010596459	0.023529171
DHPS	0.603984484	0.020590233	1.95811E-79	0.02219	0.005931541	0.028579114
CRLS1	0.605185586	0.040431771	9.3819E-100	0.00811	0.003404245	0.015539432
ABCB7	0.606702714	0.020707032	1.32796E-78	0.01501	0.022925303	0.023529171
SLC39A10	0.635279932	0.04542162	5.85996E-79	0.01043	0.029770024	0.00821

Supplementary table 2 | 290 genes that were identified as essential to cell viability in H1975_Cas9 cells through comparison of the normalized screen control arm to the normalized vector library readcounts.

Gene	Ratio	Wilcox	DESeq2	MAGeCK	sgRSEA	EdgeR
GTF2H4	0.025333022	0.013115403		0.000066	0.000121212	0.002409794
YARS2	0.026188047	0.03691053		0.000625	0.000121212	0.002409794
SARS	0.02998773	0.037498553		0.002211	0.000121212	0.002409794
CDC7	0.03357427	0.044736326		0.009039	0.000319148	0.002409794
GTF3C4	0.035034151	0.002474613		0.000066	0.000121212	0.004923521
PSMD1	0.035984214	0.03691053		0.000669	0.000121212	0.004923521
NAA10	0.036252509	0.03691053		0.002395	0.000121212	0.002409794
ASH2L	0.039479431	0.005218926		0.000066	0.000121212	0.002409794
RARS2	0.040908586	0.002046963		0.000066	0.000121212	0.002409794
GGPS1	0.044712848	0.005376171		0.000066	0.000121212	0.002409794
AASDHPPT	0.04506988	0.002046963		0.000066	0.000121212	0.002409794
TRMT112	0.050579933	0.005218926		0.000066	0.000121212	0.002409794
ABCB7	0.054416893	0.038147279		0.003898	0.000319148	0.002409794
GNB2L1	0.056094633	0.002701283		0.000066	0.000121212	0.002409794
GOLGA8O	0.056112024	0.002701283		0.000066	0.000121212	0.004923521
HSD17B10	0.05677772	0.00194871		0.000066	0.000121212	0.002409794
PSMB7	0.057642567	0.040351457		0.008483	0.000121212	0.002409794
LONP1	0.058364136	0.006514115		0.000066	0.000121212	0.002409794
HSCB	0.058483578	0.002141871		0.000066	0.000121212	0.002409794
KIF11	0.062060012	0.015831842		0.010094	0.004266654	0.002409794
CDIPT	0.062642019	0.006420112		0.000524	0.000121212	0.002409794
POLR2D	0.06376318	0.002054921		0.000066	0.000121212	0.002409794
TPI1	0.066007408	0.00194871		0.000066	0.000121212	0.002409794
GAPDH	0.066011682	0.014130114		0.000625	0.000121212	0.002409794
PYROXD1	0.068428368	0.003177405		0.000066	0.000121212	0.004923521
ALG6	0.069861085	0.00194871		0.000066	0.000121212	0.002409794
RPL5	0.072074582	0.006579396		0.000066	0.000121212	0.002409794
RPL34	0.072496028	0.043347884		0.012168	0.000121212	0.004923521
TIMM23	0.07338709	0.003049682		0.000066	0.000121212	0.002409794
PPP2CA	0.073636841	0.00194871		0.000066	0.000121212	0.002409794
PSMC4	0.073757836	0.015880502		0.00142	0.000121212	0.004923521
ATP2A2	0.074453733	0.016756655		0.018521	0.000121212	0.002409794
SSBP1	0.074926904	0.00194871		0.000066	0.000121212	0.002409794
TRMT5	0.07573377	0.015831842		0.00119	0.000121212	0.002409794
AK6	0.076151929	0.00194871		0.000066	0.000121212	0.002409794
NFS1	0.077301045	0.045293057		0.01682	0.000121212	0.004923521
MAT2A	0.077825888	0.017218026		0.027586	0.000319148	0.002409794
PIK3CA	0.077873802	0.002054921		0.000066	0.000121212	0.002409794
MFN2	0.080368744	0.00194871		0.000066	0.000121212	0.002409794
CCT8	0.081157102	0.042944418		0.009572	0.000121212	0.006769756
CRKL	0.081482878	0.002054921		0.000066	0.000121212	0.004923521
PPA1	0.083156382	0.009366624		0.006893	0.000841119	0.004923521
SNRPE	0.083165431	0.048405669		0.031441	0.000121212	0.002409794

ATP5B	0.084309382	0.00194871		0.000066	0.000121212	0.002409794
ATP5A1	0.085252101	0.002167559		0.000066	0.000121212	0.002409794
MARS2	0.088749225	0.00194871		0.000066	0.000121212	0.002409794
RPN2	0.089743826	0.003314142		0.000066	0.000121212	0.002409794
MRPS6	0.090845028	0.003177405		0.000066	0.000121212	0.002409794
QARS	0.091417322	0.019758155		0.032145	0.000121212	0.002409794
PGAM1	0.091608135	0.003374547		0.000982	0.000121212	0.002409794
GTF2H2	0.092476199	0.006713533		0.000066	0.000121212	0.002409794
PSMA1	0.092731659	0.008003722		0.001778	0.000121212	0.002409794
MRP63	0.095374774	0.00194871		0.000066	0.000121212	0.002409794
IARS2	0.096132076	0.009745661		0.000669	0.000121212	0.002409794
LARS	0.09620075	0.00963462		0.004703	0.001789878	0.002409794
SF3B3	0.096828994	0.002088651		0.000066	0.000121212	0.002409794
VARS2	0.099442446	0.0037208		0.000669	0.000121212	0.002409794
CHEK1	0.09963249	0.007643425		0.006076	0.000121212	0.002409794
PLK1	0.10041578	0.003392552		0.000719	0.000121212	0.002409794
RRM1	0.101522782	0.025323952		0.040505	0.000597013	0.002409794
MARS	0.10226141	0.006920678		0.000186	0.000121212	0.002409794
CCNK	0.103537174	0.00194871		0.000066	0.000121212	0.004923521
CLP1	0.103643558	0.0037208		0.000625	0.000121212	0.002409794
MRPL3	0.103718768	0.00194871		0.000066	0.000121212	0.002409794
MCM2	0.105072079	0.004696551	2.2516E-206	0.000066	0.000121212	0.002409794
POLR2I	0.105881035	0.016312252		0.000719	0.000121212	0.002409794
CDK9	0.107092788	0.003763772		0.000625	0.000121212	0.002409794
ALDOA	0.10794138	0.002141871		0.000066	0.000121212	0.002409794
KDSR	0.110916514	0.002237753		0.000066	0.000121212	0.002409794
ATP6V1A	0.111152745	0.002775282		0.000066	0.000121212	0.002409794
CCNA2	0.111302939	0.008549049		0.00537	0.000121212	0.002409794
KCTD10	0.113707182	0.002237753		0.000066	0.000121212	0.004923521
SLC25A26	0.115899442	0.004578533		0.008759	0.000121212	0.002409794
NRD1	0.116914737	0.004727105		0.000669	0.000121212	0.002409794
MED14	0.118011609	0.002167559		0.000066	0.000229884	0.004923521
IKBKAP	0.118409327	0.002046963		0.000413	0.000121212	0.002409794
LRPPRC	0.119389539	0.001960174		0.000066	0.000121212	0.002409794
MRPL53	0.121820507	0.002853895		0.001482	0.000121212	0.002409794
NARS2	0.123402646	0.002237753		0.000066	0.000121212	0.002409794
KAT5	0.123448324	0.002167559		0.000859	0.000121212	0.002409794
RHOA	0.124248938	0.002237753		0.000066	0.000121212	0.002409794
TFB2M	0.124725099	0.002046963		0.001333	0.000121212	0.002409794
TKT	0.125263477	0.00194871		0.000066	0.000121212	0.002409794
SPTLC1	0.126389665	0.004283664		0.021245	0.000900898	0.002409794
FDXR	0.126683159	0.003925858		0.002282	0.000597013	0.002409794
SHC1	0.127190546	0.002046963		0.000066	0.000121212	0.004923521
MRPL13	0.128127706	0.002046963		0.000625	0.000121212	0.002409794
WARS2	0.129409676	0.002054921		0.000066	0.000121212	0.002409794
MNAT1	0.130850066	0.002054921		0.001052	0.000121212	0.002409794
GFM1	0.130996042	0.004506252		0.000719	0.000121212	0.002409794

POLR2G	0.131575978	0.011634805		0.00142	0.000121212	0.002409794
CENPE	0.131932898	0.004849887		0.000719	0.000121212	0.004923521
ATP5D	0.131947898	0.002054921		0.000066	0.000121212	0.002409794
DNAJA3	0.132046273	0.004079167		0.000066	0.000121212	0.002409794
DOLPP1	0.132293527	0.00194871		0.000066	0.000121212	0.002409794
NOP2	0.132316318	0.00984132		0.006878	0.000121212	0.004923521
TTK	0.132542105	0.002701283		0.000719	0.000121212	0.002409794
NDUFV1	0.132577087	0.002701283		0.000066	0.000121212	0.002409794
NDUFAB1	0.134714291	0.027095831		0.015613	0.001078835	0.002409794
LARS2	0.135240589	0.004727105		0.000066	0.000121212	0.002409794
DAP3	0.138829493	0.002046963		0.000625	0.000121212	0.002409794
DHPS	0.139106398	0.00349704		0.000066	0.000121212	0.002409794
HDAC3	0.140371653	0.015258506		0.000669	0.000507613	0.002409794
DPM1	0.141463325	0.002046963		0.000066	0.000121212	0.002409794
NDUFS1	0.141512568	0.002046963		0.000186	0.000121212	0.002409794
WEE1	0.143785775	0.005455274		0.012633	0.000229884	0.002409794
POLG	0.143984369	0.002167559		0.000066	0.000121212	0.002409794
MVD	0.144048611	0.002141871		0.000784	0.000121212	0.002409794
HYOU1	0.146287938	0.015258506		0.000625	0.000121212	0.004923521
NUP62	0.14870565	0.007290839		0.000413	0.000121212	0.002409794
NDUFAF3	0.149155688	0.002054921		0.000625	0.000121212	0.002409794
TOMM40	0.14966176	0.008045521		0.009617	0.000121212	0.002409794
ATP5E	0.150296092	0.002383915		0.004508	0.000121212	0.002409794
METTL1	0.150628071	0.002054921		0.000066	0.000121212	0.002409794
MTG2	0.151225576	0.002054921		0.000066	0.000121212	0.002409794
EGFR	0.151512952	0.029031772		0.038697	0.001078835	0.002409794
ATP5H	0.152395578	0.003177405		0.030693	0.000121212	0.004923521
FARS2	0.152450729	0.002167559		0.006163	0.000121212	0.002409794
RAPGEF1	0.153738059	0.002853895		0.002247	0.000121212	0.002409794
MRPS2	0.153754833	0.002054921		0.006017	0.000121212	0.002409794
PMPCA	0.153790053	0.00694306		0.004842	0.000121212	0.002409794
RPE	0.154454979	0.002054921		0.029379	0.000319148	0.002409794
SOD1	0.15465479	0.025647429		0.013507	0.000960696	0.002409794
CTPS1	0.154946796	0.003925858		0.000625	0.000121212	0.002409794
ALG5	0.155170705	0.006012345		0.000982	0.000121212	0.002409794
IARS	0.155742352	0.03008863		0.009039	0.000121212	0.002409794
MRPS34	0.155828488	0.003049682		0.021245	0.000319148	0.002409794
ATP6V1C1	0.155953942	0.003574594		0.000719	0.000121212	0.002409794
MRPL38	0.158736076	0.002218921		0.004508	0.000121212	0.002409794
TSFM	0.159216135	0.002141871		0.000066	0.000121212	0.004923521
SYVN1	0.162977632	0.004556915		0.002247	0.000121212	0.002409794
PKN2	0.163501213	0.002054921		0.000625	0.000229884	0.002409794
UBE2M	0.164286686	0.002141871		0.000066	0.000121212	0.002409794
ATP1A1	0.165002119	0.02786964		0.005137	0.000121212	0.002409794
EXOSC10	0.165637546	0.003925858		0.000784	0.000121212	0.002409794
SLC35B1	0.166312038	0.012403575		0.036161	0.000121212	0.004923521
MRPL41	0.166681771	0.002167559		0.005884	0.000121212	0.002409794

CDK2	0.168291406	0.002054921		0.006076	0.000900898	0.002409794
ELP3	0.169455459	0.004008767		0.000066	0.000121212	0.002409794
CSNK1A1	0.170161852	0.006533052		0.000066	0.000319148	0.002409794
ENO1	0.172700864	0.002167559		0.000186	0.000121212	0.002409794
ATP6V1E1	0.175053386	0.014059165		0.000066	0.001599995	0.004923521
POLG2	0.175072526	0.002167559		0.004842	0.000121212	0.002409794
ICMT	0.175213253	0.002141871		0.003598	0.000507613	0.008302827
PAFAH1B1	0.177470485	0.044165299		0.010723	0.000121212	0.002409794
ROMO1	0.178721643	0.002685306		0.000625	0.000121212	0.002409794
ISCA2	0.179563921	0.002701283		0.000066	0.000121212	0.002409794
ITGB5	0.180737185	0.003348917		0.007156	0.000319148	0.002409794
COX7C	0.181118306	0.002167559		0.000625	0.000121212	0.002409794
TIMM50	0.181276993	0.003177405		0.016005	0.000229884	0.004923521
CDK7	0.181405954	0.004447228		0.000066	0.000900898	0.002409794
MRPS12	0.182265859	0.002830278		0.003172	0.000121212	0.002409794
CARS2	0.182949575	0.002167559		0.006182	0.000676326	0.002409794
POLR2A	0.183357692	0.028546703		0.000784	0.000121212	0.002409794
FXN	0.185699779	0.004506252		0.002614	0.000121212	0.004923521
TARS	0.186419987	0.011901044		0.001398	0.000319148	0.004923521
PYURF	0.186830365	0.002167559		0.008764	0.000121212	0.002409794
ERCC3	0.187431454	0.012614948		0.018952	0.001818176	0.002409794
MTIF2	0.188234312	0.002218921		0.006511	0.001732278	0.002409794
COPB2	0.189045381	0.009318122		0.004732	0.000507613	0.002409794
DMAP1	0.189290346	0.003772414		0.000066	0.000121212	0.002409794
CCNH	0.193916223	0.008008017		0.009662	0.000676326	0.002409794
DARS2	0.195319066	0.002205713		0.0166	0.000229884	0.002409794
MPDU1	0.196403567	0.003015449		0.013654	0.000121212	0.002409794
PREB	0.196524494	0.009699801		0.02012	0.000121212	0.002409794
COQ4	0.197285472	0.002701283		0.000625	0.000121212	0.002409794
GTPBP4	0.197637654	0.004506252		0.000625	0.000121212	0.002409794
PMPCB	0.199295629	0.041751213		0.036656	0.001824812	0.002409794
SAE1	0.199318105	0.009701786		0.002282	0.001012655	0.002409794
ITGAV	0.200191665	0.010032717		0.000066	0.000121212	0.002409794
MTO1	0.203548494	0.002436718		0.020522	0.001599995	0.002409794
OXSM	0.203659635	0.002701283		0.004079	0.000121212	0.002409794
N6AMT1	0.205425097	0.002994833		0.000066	0.000121212	0.002409794
KDM8	0.205613911	0.008196747		0.000625	0.000121212	0.002409794
CMTR1	0.205975385	0.003245029		0.044794	0.000900898	0.002409794
MOCS3	0.207209462	0.003504884		0.022004	0.000121212	0.002409794
DPH5	0.208576172	0.005591032		0.017802	0.000229884	0.002409794
PARS2	0.20949808	0.010303357		0.004111	0.000121212	0.004923521
UQCRC2	0.212119703	0.003518354		0.010602	0.001599995	0.002409794
TIMM13	0.212316883	0.004008767		0.006641	0.000121212	0.002409794
MRPL18	0.214060882	0.004594427		0.04302	0.001012655	0.002409794
TRMT61A	0.215005061	0.006070658		0.018793	0.000229884	0.002409794
PGD	0.215297662	0.009699801		0.000625	0.000121212	0.002409794
AIFM1	0.215967154	0.003177405		0.003465	0.000121212	0.008302827

COASY	0.216707287	0.004065776		0.009043	0.002315782	0.002409794
SMARCA5	0.218478488	0.0032961		0.049782	0.000960696	0.002409794
MRPS16	0.219327891	0.005218926		0.001603	0.000960696	0.002409794
CARM1	0.222373875	0.003049682		0.006657	0.000121212	0.002409794
WARS	0.223638378	0.010643441		0.012168	0.000900898	0.004923521
SLC3A2	0.223986748	0.00349704		0.005977	0.000841119	0.002409794
ASNA1	0.226900394	0.006181359		0.031759	0.000121212	0.002409794
MRPL4	0.227256729	0.00596314		0.002522	0.000121212	0.002409794
ATP5F1	0.228535856	0.004849887		0.000918	0.000900898	0.002409794
CDK4	0.231332052	0.015725018		0.005215	0.001316868	0.002409794
GFER	0.231670902	0.0135764		0.002211	0.000597013	0.002409794
DAD1	0.233952169	0.009366624		0.007642	0.002315782	0.002409794
SUPV3L1	0.234792083	0.008241597		0.004924	0.000507613	0.002409794
MRPS11	0.236470774	0.004263093		0.023666	0.000121212	0.002409794
SARS2	0.237800376	0.019556378		0.000066	0.000121212	0.002409794
METTL17	0.238460368	0.004357653		0.024646	0.000121212	0.002409794
TPX2	0.238862449	0.007613021		0.002868	0.000960696	0.002409794
MRPS21	0.239067667	0.004577558		0.000066	0.000121212	0.002409794
MRPL20	0.240117844	0.009745661		0.000918	0.00280487	0.002409794
QRSL1	0.240462353	0.007345493		0.027803	0.003911834	0.004923521
MRPS7	0.241664746	0.004506252		0.003598	0.000121212	0.004923521
BLM	0.244127146	0.006089693		0.029686	0.000676326	0.002409794
NDUFA2	0.245678959	0.003772414		0.004079	0.000121212	0.002409794
MRPS22	0.246050323	0.012359089		0.028889	0.001012655	0.002409794
RPUSD4	0.246250982	0.013115403		0.022964	0.003458203	0.002409794
MRPL46	0.247563552	0.005591032		0.000186	0.000676326	0.002409794
ATR	0.249830012	0.008262091		0.020491	0.001818176	0.002409794
COX5A	0.252083398	0.004506252		0.037794	0.003095229	0.002409794
DPM3	0.252916287	0.008392085		0.000066	0.000121212	0.002409794
ADSL	0.255917377	0.006384181		0.006625	0.002364858	0.002409794
MRPL39	0.256718737	0.012136971		0.002211	0.000121212	0.004923521
BRCA2	0.25774026	0.006420112		0.029379	0.00280487	0.002409794
MRPL34	0.260720007	0.006384181		0.001987	0.000319148	0.002409794
MRPL15	0.26143514	0.012291749		0.001822	0.001557372	0.002409794
CAD	0.263768236	0.003925858		0.005393	0.001824812	0.002409794
MRPS18C	0.26489664	0.009366624		0.012826	0.000597013	0.002409794
HARS2	0.266533046	0.015258506		0.022323	0.00280487	0.002409794
STT3A	0.266535888	0.010736626		0.000066	0.001012655	0.002409794
MRPL45	0.267111761	0.008040489		0.000066	0.000507613	0.002409794
NDUFS5	0.268774124	0.006669476		0.048726	0.001599995	0.004923521
MRPL49	0.274632808	0.012291749		0.003465	0.001824812	0.002409794
MRPS35	0.275088413	0.011634805		0.000066	0.001078835	0.002409794
NSUN4	0.276139412	0.012634927		0.007642	0.001818176	0.002409794
DDOST	0.281106828	0.044785668		0.014348	0.004409435	0.002409794
IDH3A	0.281830789	0.013115403		0.000625	0.001012655	0.002409794
MRPL47	0.282112295	0.006181359		0.02903	0.00232876	0.002409794
SLC25A3	0.283816997	0.015880502		0.001122	0.001660074	0.002409794

CTU1	0.283978546	0.006420112		0.000066	0.001789878	0.004923521
METTL14	0.287259466	0.0256545		0.03079	0.001824812	0.002409794
RFT1	0.287449806	0.02786964		0.037794	0.00267973	0.002409794
MRPL44	0.289531931	0.026256302		0.007642	0.00232876	0.002409794
ATP5O	0.291986766	0.008059042		0.046801	0.010193517	0.002409794
MRPS10	0.292087164	0.012527075		0.00142	0.002364858	0.002409794
MRPS28	0.293776069	0.00659756		0.037794	0.009179573	0.002409794
MRPS18A	0.294333429	0.0219453		0.000066	0.000841119	0.002409794
NOB1	0.295103194	0.016704933		0.01682	0.001660074	0.002409794
UGCG	0.299587715	0.014020747		0.002703	0.001078835	0.002409794
PKM	0.300823248	0.020134449		0.000066	0.000319148	0.002409794
AURKA	0.301834788	0.012656	1.544E-302	0.000669	0.000960696	0.002409794
PGK1	0.305327525	0.017033204		0.005977	0.003274844	0.002409794
PISD	0.30679865	0.026500073		0.03615	0.002700956	0.002409794
NAA30	0.309481638	0.015813065		0.018952	0.00232876	0.004923521
UQCRQ	0.310454679	0.00963462		0.027914	0.020770579	0.002409794
CDC37	0.312490785	0.007162215		0.010052	0.000900898	0.002409794
TFB1M	0.31328516	0.009771225		0.00142	0.000841119	0.002409794
MRPL23	0.314490825	0.02037153		0.02012	0.003063054	0.006769756
MRPL10	0.317611638	0.025004059		0.034814	0.008194419	0.002409794
NDUFS2	0.318016639	0.016811897		0.003172	0.003235284	0.002409794
ING3	0.319064907	0.016811897		0.00142	0.000319148	0.002409794
NAT10	0.319454649	0.011682771		0.000413	0.006682672	0.004923521
MRPL51	0.322013641	0.016804476		0.006163	0.012860001	0.002409794
WNK1	0.323732313	0.016756655		0.029585	0.019031599	0.002409794
UFC1	0.328820862	0.017218026		0.000186	0.0024161	0.002409794
COX7B	0.329475942	0.017575019		0.03534	0.006146323	0.002409794
TRIT1	0.332645091	0.036421644		0.014348	0.003063054	0.002409794
DPM2	0.336155246	0.022051231		0.002703	0.003599989	0.002409794
CHORDC1	0.337075871	0.042613773		0.000918	0.004119229	0.002409794
MRPL28	0.3374438	0.0135764		0.01583	0.002277573	0.002409794
MRPL36	0.338512809	0.015855988	8.8334E-282	0.044971	0.00969694	0.002409794
MRPL21	0.344779273	0.019758155		0.031759	0.002700956	0.004923521
NDUFB11	0.353503307	0.015855988		0.034091	0.005808063	0.002409794
CINP	0.353520372	0.025769653		0.000625	0.001824812	0.002409794
PTCD3	0.364133766	0.02610169		0.016508	0.001012655	0.002409794
ATP6AP1	0.367202778	0.04651406	1.1746E-293	0.008289	0.003854737	0.002409794
KDM1A	0.369191316	0.015813065	4.149E-251	0.036956	0.009022195	0.004923521
AARS2	0.372413545	0.039030693	2.9068E-223	0.027113	0.00389971	0.004923521
HERC2	0.377843454	0.027053704	8.2658E-247	0.042819	0.021923708	0.004923521
EARS2	0.37993102	0.014696238	7.7141E-269	0.007049	0.001660074	0.002409794
MRPS5	0.379938912	0.025358628		0.029585	0.003030294	0.002409794
MRPS9	0.381750701	0.045282564	5.3491E-218	0.010052	0.013547053	0.002409794
TOMM22	0.389066813	0.03498193	2.2353E-215	0.008477	0.00267973	0.002409794
CXXC1	0.393479247	0.026256302	2.0432E-	0.049754	0.006198528	0.004923521

			247			
VRK1	0.405996899	0.044035554	2.913E-175	0.000669	0.012148723	0.006769756
MRPS25	0.4170093	0.04401203	1.5656E-238	0.003691	0.005116263	0.002409794
KHSRP	0.42160713	0.024224798	7.0698E-257	0.043189	0.000676326	0.002409794
NOA1	0.422593963	0.044035554	2.1068E-203	0.00333	0.00280487	0.002409794
TAF5	0.425211293	0.04401203	2.1872E-221	0.016297	0.009315645	0.004923521
PDPK1	0.443234581	0.040165386	1.7693E-195	0.036697	0.007405638	0.002409794
MAPKAP1	0.448461534	0.027053704	3.1898E-218	0.03344	0.001818176	0.002409794
PIK3R4	0.456518449	0.042613773	1.4974E-187	0.018286	0.00280487	0.002409794
DHFR	0.541642685	0.018263159	7.7856E-101	0.022988	0.003095229	0.002409794
MEPCE	0.595017838	0.043985807	3.8138E-78	0.025159	0.001824812	0.002409794

Supplementary table 3 | Overview driver mutations LKB1, KEAP1, CHD8, EP300 and EZH2 status in NSCLC cell lines

Cell line	Driver Mutations	KRAS mut details	TP53 mut details	LKB1 deficient?	KEAP1	CHD8	EP300	EZH2	Other
H2030	KRAS	G12C	p.G262V	yes	p.V568F	WT (COSMIC), p.G54A (CCLE)	WT	WT	n/a
H1944	KRAS	G13D	WT	yes	p.R272L	WT	WT	WT	n/a
H1299	NRAS, TP53Null	WT	Null	no	WT	p.R1337Q	down	WT	n/a
H1975	EGFR, CDKN2A, PIK3CA	WT	p.R273H	no	WT	WT	WT	WT	n/a
HCC827	EGFR	WT	?	no	WT	WT (COSMIC), p.H2494Y (CCLE)	WT	WT	n/a
H2228	EML4/ALK, RB1, TP53	WT	p.Q331*	no	WT	WT	WT	WT	n/a
H522	TP53	WT	p.P191fs*56	no	WT	WT	WT	WT	n/a
H1563	TP53	WT	p.P191fs, p.P59fs	yes	WT	WT	WT	WT	n/a
H1650	EGFR, TP53	WT	c.A673G	no	WT	down	WT	WT	n/a
A549	KRAS	G12S	WT	Yes	p.G333C	WT	WT	WT	n/a
H358	KRAS	G12C	WT	No	WT	p.R303Q, p.R24Q	WT	WT	n/a
H2122	KRAS, TP53	G12C	p.C176F, p.Q16L	yes	p.A170_R204del35	WT	WT	WT	n/a
H3122	EML4/ALK, TP53	WT	E285V	No	WT	WT	WT	WT	n/a
H1395	BRAF, CDKN2A, TP53	WT	WT	yes	WT	WT	WT	WT	n/a
H1355	KRAS	G13C	p.R175H	Yes	p.Q75*	WT	WT	WT	n/a
H1568	EGFR	WT	WT	Yes	WT	WT (COSMIC	WT	WT	crebbp mut

						), p.E72V (CCLE)			
H1755	CDKN2A, TP53, PTEN	WT	p.C242F	Yes	p.E582*	WT	WT	WT	n/a
H1993	Met upregulation	WT	WT	Yes	WT	WT	WT	WT	n/a
H23	KRAS	G12C	WT	Yes	p.Q193H	WT	WT	WT	n/a
H647	KRAS	G13D	WT	Yes	WT	WT	WT	WT	n/a
H838	CDKN2A, TP53	WT	p.E62*	Yes	p.E444*, down	WT	WT	WT	n/a
H1435	N/A	WT	WT	Yes	p.R413L	WT	WT	WT	n/a
H1651	CDKN2A, TP53	WT	p.C176Y	Yes	WT	WT	WT	WT	crebbp mut
H1693	TP53	WT	p.?	No	WT	WT	WT	WT	n/a
H1703	CDKN2A, TP53	WT	p.?	No	WT	WT	p.A1437V, p.L1428L	WT	crebb down
H1792	KRAS, TP53	p.G12C	p.?	No	p.G462W	WT	WT	WT	n/a
H1793	CDKN2A, TP53	WT	p.?	No	WT	WT	p.G1711G	WT	crebb down
H1838	CDKN2A, TP53	WT	p.R273L	No	WT	WT	WT	WT	n/a
H2009	KRAS, RB1, TP53	c.35G>C	p.R273L	No	WT	WT	WT	WT	n/a
H2291	KRAS, TP53	p.G12F	p.G154V	No	WT	WT	WT	WT	n/a
H2342	TP53	WT	p.Y220C	No	WT	p.P1868P	WT	WT	crebbp mut
H2405	TP53, CDKN2A, SMAD4, BRAF	WT	p.R273H	No	WT	WT	WT	WT	n/a

7 Acknowledgements

First and foremost, I would like to sincerely thank Prof. Dr. Holger Sültmann for giving me the opportunity to work on this project and be a member of his research group. I want to thank Prof. Sültmann especially for giving me a lot of freedom to conduct the research while always being open to supporting me and to steering the project with true leadership. Equally important, I want to thank him for supporting the start of my industry tenure and for continued support even after I had left the laboratory.

I would like to thank Prof. Dr. Ursula Klingmüller for agreeing to be a member of the thesis advisory committee and for acting as the faculty examiner. Further, I want to thank her for her highly valuable input during our discussions.

I want to thank Dr. Michael Meister for his very insightful commentary during the thesis advisory meeting and for his continued support of this work.

I also would like to thank all current and former members of the Division of Cancer Genome Research for their help in the laboratory, the valuable input at our division seminars, and for great coffee chats. My best wishes go to Sabine, Simon, Sabrina, Arlou, Simone, Anka, Florian, Anja, Steffen, Doreen, Sajo, Leonie, Julia, and Vladimir. I would like to point out my sincere gratitude for the excellent assistance by Simon and Sabrina in the laboratory and for administrative affairs; moreover, I would like to thank Sabine for her feedback on the raw manuscript.

This work would not have been finished without the help of my dear friend Julian. You are a real friend, and I look forward to more adventures to come for us. Rest assured that I will never forget what you did for me.

Finally, my deepest gratitude shall be expressed for my wife Agnes. I am humbled and honored to be your husband, and without you, I am nothing.

I'll close by thanking Sofia and Antonia. Thank you for having let daddy work on this thesis late hours and on the weekends, and for supporting my mental health with all those fantastic drawings on my office's whiteboard. I couldn't be any prouder of you!