Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M.Sc. Ksenia Myacheva

born in Makhachkala, Russia

Oral examination: 27th October 2021

[This page intentionally left blank]

Custom CRISPRi-based screen

for general and genotype-specific lung cancer viability factors

Referees: Prof. Dr. Georg Stoecklin

Prof. Dr. Sven Diederichs

[This page intentionally left blank]

Abstract

Lung cancer remains the most common cause of cancer death worldwide. Non-small cell lung cancer (NSCLC) is the most common type of lung cancer, and lung adenocarcinoma (LUAD) is the most frequent subtype of NSCLC. Discovery of the prevalence of distinct driver mutations in LUAD tumors has led to a booming development of targeted therapies blocking the resulting oncogene activation. However, long-term therapeutic outcomes of such treatments are poor, and novel therapeutic targets for mutant LUAD are urgently required.

I set out to investigate novel genotype-specific LUAD viability factors, and therefore developed and conducted a pooled CRISPR interference (CRISPRi) dropout screen in eight LUAD cell lines with a custom single guide RNA (sgRNA) library. To compile the screen targets, I analyzed the transcriptional profiles from LUAD patients and selected protein-coding genes and long non-coding RNAs (lncRNAs) differentially overexpressed between LUAD samples with different driver mutational signatures. LncRNAs overexpressed in LUAD compared to normal lung samples were also added to the candidate list to enrich the screen with novel targets, however, neither genotype-specific nor general lncRNA viability factors for LUAD were identified in the screen.

The screen revealed that *EGFR*-mutant LUAD cell lines exhibited resistance to the knockdown of the mitotic regulator *CENPE*. Conversely, *CENPE* expression was essential for the viability of *EGFR*-wildtype cells. Functional validation of the observed divergent responses revealed a direct link between the mutant EGFR activity and the resistance to the *CENPE* knockdown in LUAD cell lines.

Moreover, silencing of *CASP8AP2* resulted in the loss of cell viability in all screened LUAD and additional NSCLC cell lines, while, in contrast, non-transformed lung cells tolerated its depletion. I uncovered that expression levels of *JUNB*, encoding for a subunit of the AP-1 transcription factor, were lower in steady state as well as upon *CASP8AP2* silencing in the tolerant cell lines compared to the cancer cells sensitive to the *CASP8AP2* loss. Inhibition of AP-1 rescued the loss of viability phenotype in NSCLC cell lines, uncovering the *CASP8AP2* – AP-1 functional axis.

Overall, my custom CRISPRi screen identified a novel genotype-specific mechanism in *EGFR*-mutant LUAD cells that were able to overcome disruption of the mitotic regulation and discovered *CASP8AP2* as an essential NSCLC viability determinant and a promising candidate for therapeutic targeting acting in cooperation with the AP-1 transcription factor.

[This page intentionally left blank]

Zusammenfassung

Lungenkrebs ist weltweit nach wie vor die häufigste krebsbedingte Todesursache. Das nicht-kleinzellige Bronchialkarzinom (NSCLC) ist die häufigste Form von Lungenkrebs, und das Adenokarzinom der Lunge (LUAD) ist wiederum der häufigste Subtyp von NSCLC. Die Entdeckung verschiedener Treibermutationen in LUAD-Tumoren hat zur Entwicklung mehrerer zielgerichteter Therapien geführt, die die aus der Mutation resultierende Onkogenaktivierung blockieren. Die langfristigen Therapieergebnisse solcher Behandlungen sind jedoch aufgrund von Resistenzen nicht zufriedenstellend, so dass neue therapeutische Zielstrukturen benötigt werden.

Ziel meiner Arbeit war, neue genotyp-spezifische LUAD-Viabilitätsfaktoren zu identifizieren und zu charakterisieren. Dazu habe ich einen maßgeschneiderten, gepoolten CRISPR-Interferenz-(CRISPRi)-negativen Selektionsscreen in acht LUAD-Zelllinien entwickelt und durchgeführt. Um die Liste der Screening-Targets zu erstellen, analysierte ich zunächst die Transkriptionsprofile von LUAD-Patienten, um proteincodierende Gene und lange nicht-codierende RNAs (lncRNAs) auszuwählen, die zwischen LUAD-Tumoren mit unterschiedlichen Treiber-Mutationen differentiell induziert waren. Zusätzlich wurden lncRNAs aufgenommen, die in LUAD im Vergleich zu normalem Lungengewebe überexprimiert waren, um den Screen mit neuen Zielstrukturen anzureichern. Jedoch wurden im Screen weder genotyp-spezifische noch allgemeine lncRNA-Viabilitätsfaktoren für LUAD identifiziert.

In dem Screen entdeckte ich, dass *EGFR*-mutierte LUAD-Zelllinien eine Resistenz gegenüber dem Knockdown des mitotischen Regulators *CENPE* zeigten. Umgekehrt war die *CENPE*-Expression für die Lebensfähigkeit von *EGFR*-Wildtypzellen essentiell. Die funktionelle Validierung der beobachteten divergenten Reaktion zeigte einen direkten Zusammenhang zwischen der mutierten EGFR-Aktivität und der Resistenz gegen dem *CENPE*-Inhibition in LUAD-Zelllinien.

Darüber hinaus führte die Verringerung von *CASP8AP2* zum Verlust der Zellviabilität in allen getesten LUAD-Zelllinien im Screen, während im Gegensatz dazu nicht-transformierte Lungenzellen den Verlust von *CASP8AP2* tolerierten. Weiter entdeckte ich, dass die Expression von *JUNB*, das für eine Untereinheit des AP-1-Transkriptionsfaktors kodiert, sowohl im Gleichgewichtszustand als auch bei *CASP8AP2*-Inhibition in toleranten Zelllinien niedriger waren als bei Tumorzellen, die auf den Verlust von *CASP8AP2* empfindlich

reagierten. Die Hemmung von AP-1 konnte den Verlust der Zellviabilität induziert durch CASP8AP2-Inhibition ausgleichen und deckte so die funktionelle Achse CASP8AP2 – AP-1 auf.

Insgesamt identifizierte mein maßgeschneiderter CRISPRi-Screening-Ansatz einen neuen genotyp-spezifischen Mechanismus in *EGFR*-mutierten LUAD-Zellen, der zu einer Resistenz gegenüber Störung der mitotischen Regulation führte, und entdeckte *CASP8AP2* als wesentliche Determinante der NSCLC-Viabilität und vielversprechenden Kandidaten als therapeutische Zielstruktur, die gemeinsam mit dem Transkriptionsfaktor AP-1 agiert.

Table of contents

List of abbreviations

Abbreviation	Meaning
#	Number
%	Percent
°C	Degrees Celsius
aa	Amino acid
ADP	Adenosine diphosphate
AFAP1-AS1	AFAP1 antisense RNA 1
AKT	Akt/protein kinase B
ALK	Anaplastic lymphoma receptor tyrosine kinase
ASO	Antisense oligonucleotide
AP-1	Activator protein 1
APC/C	Anaphase-promoting complex/cyclosome
ATCC	American Type Culture Collection
AURKA	Aurora kinase A
AURKB	Aurora kinase B
Blast	Blasticidin
bp	Base pairs
BRAF	V-raf murine sarcoma viral oncogene homolog B
BRCA1	Breast cancer 1, early onset
BRCA2	Breast cancer 2, early onset
BrdU	5-Bromo-2'-deoxyuridine
CAGE	Cap-Analysis of Gene Expression
CASP3	Caspase-3
CASP7	Caspase 7
CASP8	Caspase-8
CASP8AP2	Caspase-8 associated protein-2
CASP9	Caspase 9
CDC20	Cell division cycle 20
CDCA3	Cell division cycle associated 3
CDKN1A	Cyclin-dependent kinase inhibitor 1A (p21)
CDKN2A	Cyclin-dependent kinase inhibitor 2A
cDNA	Complementary DNA
CENPE	Centromere protein E
CEP55	Centrosomal protein 55kDa
CHEK1	Checkpoint kinase 1
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRi	CRISPR interference

crRNA	CRISPR RNA
CTG	CellTiter-Glo
dCas9	Catalytically “dead” Cas9
DDX11-AS1	DDX11 antisense RNA 1
DKFZ	German Cancer Research Center
DMEM	Dulbecco's modified Eagle’s medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleoside triphosphate
ds	Double-stranded
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
EGFR	Epidermal growth factor receptor
EMT	Epithelial-to-mesenchymal transition
ERBB2	V-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2
ERBB3	V-erb-b2 avian erythroblastic leukemia viral oncogene homolog 3
ERBB4	V-erb-b2 avian erythroblastic leukemia viral oncogene homolog 4
ERK	Extracellular signal-regulated kinase
e.v.	Empty vector
FACS	Fluorescence activated cell sorting
FC	Fold change
FCS	Fetal calf serum
FDA	U.S. Food and Drug Administration
FDR	False discovery rate
FITC	Fluorescein isothiocyanate
FPKM	Fragments per kilobase per million
FSC-A	Forward scatter area
FSC-H	Forward scatter height
FSC-W	Forward scatter width
g	Gram
g	Gravitational force
GC content	Guanine-cytosine content
GDP	Guanosine diphosphate
GOF	Gain-of-function
GPCR	G-protein coupled receptor
GTP	Guanosine triphosphate
H3K9me3	Histone 3 lysine 9 trimethylation
hg19/GRCh37	Human genome assembly 19 from Genome Reference Consortium

hg38/GRCh38	Human genome assembly 38 from Genome Reference Consortium
HRAS	Harvey rat sarcoma viral oncogene homolog
HUGO	Human Genome Organization
JAK	Janus kinase
JUNB	Jun B proto-oncogene
kb	Kilobases
KD	Knockdown
KRAB	Krüppel-associated box
KRAS	Kirsten rat sarcoma viral oncogene homolog
L	Liter
LCLC	Large cell lung cancer
lncRNA	Long non-coding RNA
lincRNA	Long intergenic non-coding RNA
LOF	Loss-of-function
Log_2	Binary logarithm
Log_{10}	Decimal logarithm
LUAD	Lung adenocarcinoma
LUSC	Lung squamous carcinoma
MAPK	Mitogen-activated protein kinase
Max.	Maximal / maximum
MCC	Mitotic checkpoint complex
MCM2	Minichromosome maintenance complex component 2
MEK	Mitogen-activated protein kinase kinase
MEM	Minimal essential medium
MET	Met proto-oncogene
mg	Milligram
Min.	Minimal / minimum
miRNA	microRNA
mL	Milliliter
mRNA	Messenger RNA
MALAT1	Metastasis Associated Lung Adenocarcinoma Transcript 1
MOI	Multiplicity of infection
Mut	Mutant
n	Sample size
NC	Negative control
ncRNA	Non-coding RNA
NF1	Neurofibromin 1
ng	Nanogram
NGS	Next-generation sequencing

NHEJ	Non-homologous end joining
NI	Normalized intensities
nm	Nanometer
NOA	Non-oncogene addiction
No.	Number
NRAS	Neuroblastoma RAS viral (v-ras) oncogene homolog
n.s.	Not significant
NSCLC	Non-small cell lung cancer
nt	Nucleotides
OD600	Optical density at 600 nm
OE	Overexpression
Opti-MEM	Optimized reduced serum medium
ORF	Open reading frame
p	p-value
p_{adj}	Adjusted p-value
PAM	Protospacer-adjacent motif
PARP	Poly(ADP-ribose) polymerase
PBS	Phosphate Buffered Saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PFS	Progression-free survival
pH	Potential of hydrogen
PI	Propidium Iodide
PI3K	Phosphoinositide 3-kinase
PLK1	Polo-like kinase 1
PLK4	Polo-like kinase 4
PNK	Polynucleotide kinase
pre-miRNA	Precursor miRNA
Puro	Puromycin
PVDF	Polyvinylidene fluoride
pri-miRNA	Primary miRNA
RISC	RNA-induced silencing compex
RNA	Ribonucleid acid
RNA-Seq	RNA sequencing
RNAi	RNA interference
RNase	Ribonuclease
RNP	Ribonucleoprotein complex
ROS1	C-ros oncogene 1
RPKM	Reads per kilobase per million

RPL11	Ribosomal protein L11
rpm	Revolutions per minute
RPMI-1640	Roswell Park Memorial Institute 1640 medium
RTK	Receptor tyrosine kinase
RT	Reverse transcription
RT-qPCR	Real-time quantitative PCR
SAC	Spindle assembly checkpoint
SCLC	Small cell lung carcinoma
SDS	Sodium dodecyl sulphate
sgRNA	Single guide RNA
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
siPOOL	siRNA pool
SPAG5	Sperm associated antigen 5
ss	Single-stranded
SSC-A	Side scatter area
STAT	Signal transducers and activators of transcription
STAT5	Signal transducer and activator of transcription 5
STK33	Serine/threonine kinase 33
TANRIC	The Atlas of Non-coding RNA in Cancer
TE	Tris-EDTA
TCF3	Transcription factor 3
TCGA	The Cancer Genome Atlas
TKI	Tyrosine kinase inhibitor
TP53	Tumor protein p53 (p53)
tracrRNA	trans-activating CRISPR RNA
TSS	Transcription start site
v	Volume
vs.	Versus
VSV-G	Vesicular Stomatitis Virus Envelope Glycoprotein
w	Weight
WT	Wildtype
ZNF10	Zinc-Finger Protein 10
µg	Microgram
µg	Microgram
µL	Microliter
µm	Micrometer

[This page intentionally left blank]

List of figures

List of tables

[This page intentionally left blank]

1 Introduction

1.1 Lung cancer

1.1.1 Epidemiology and major risk factors

Despite continuous clinical and research efforts, lung cancer remains the most common cause of cancer-associated death worldwide, accounting for approximately 2.1 million new cases and 1.8 million deaths per year [1, 2]. The current global 5-year survival rate estimations for lung cancer fall between 10-20 % [1]. Strikingly, unlike for most of the other cancer types, the 5-year survival for lung cancer has only shown a minor improvement trend over the last five decades [3].

Lung cancers are categorized into two main histological types, namely non-small cell lung carcinomas (NSCLC) and small cell lung carcinomas (SCLC). The NSCLC account for nearly 85% of all lung cancers. Lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) are the most prevalent NSCLC subtypes, accounting for 50% and 23% of lung cancers, respectively [4–6].

Cigarette smoking is the major risk factor for NSCLC and is linked to 80-90% of lung cancer deaths [7]. However, approximately 25% of lung cancers, predominantly of the LUAD subtype, occur in non-smokers. When considered separately, lung cancer in never-smokers is the seventh leading cause of cancer-related death worldwide [8, 9].

1.1.2 Mutational landscape

Estimates suggest that more than half of the NSCLC patients harbor somatic mutations in cancer driver genes, occurring predominantly in a mutually exclusive manner [10–14]. LUAD and LUSC subtypes of NSCLC have distinct mutational profiles [14, 15]. In LUAD, the most common mutations are detected within *EGFR* and *KRAS* oncogenes, each occurring in approximately 30% of total mutated LUAD cases [4, 14, 16–18]. Other common genetic alterations in LUAD include mutations in *MET*, *ALK*, *BRAF* and *ERBB2* proto-oncogenes and *TP53*, *NF1*, *CDKN2A, KEAP1* and *STK11* (*LKB1*) tumor suppressors [4, 13–15, 17, 18]. In LUSC, the most mutations affect the *TP53* and *CDKN2A* tumor suppressors [4].

1.1.3 Clinical management

For early-stage NSCLC, surgical resection of the tumor is the most effective treatment option [19]. However, currently most NSCLC cases are diagnosed at advanced stages where complete surgical resection is not possible due to cancer spreading [20]. For years, the empiric platinum-based chemotherapy has been the gold standard treatment option for advanced-stage NSCLC [4, 21]. Yet, the use of the cytotoxic treatments has been marked by limited efficacy coupled with severe toxicity in NSCLC patients, which has led to an extensive exploration of alternative therapeutic approaches for NSCLC over the recent decades [4, 22, 23].

Current treatment options for advanced-stage NSCLC include the use of novel immune checkpoint inhibitors. However, despite unprecedented survival benefits observed in selected patients, a significant proportion of NSCLC patients exhibit a lack of response and/or eventual resistance to the novel immunotherapeutic approaches [4, 24]. Notably, the current immunotherapies fail to perform in patients with NSCLC harboring *EGFR* mutations and *ALK* gene rearrangements [25–27].

Robust characterization of distinct mutational landscapes of NSCLC, enabled by the advent of next-generation sequencing, fueled the recent booming development of targeted therapy approaches [4]. In the context of drug development, oncogene activation by gain-of-function (GOF) mutations is often utilized as a target for therapeutic inhibition, while development of strategies restoring the function of tumor suppressors inactivated by loss-of-function (LOF) mutations has proven challenging [28]. As such, several targeted therapies, most notably various tyrosine kinase inhibitors (TKIs), have been approved as first-line treatments for the respective oncogene-driven NSCLC. Nevertheless, the long-term therapeutic outcomes of such treatments have so far also been limited by the inevitably occurring drug resistance [4, 19].

1.2 Oncogene addiction and targeted therapies in NSCLC treatment

Lung cancers are heterogeneous tumors composed of distinct malignant cell populations with different molecular profiles, termed “subclones” [4]. Despite this intratumoral heterogeneity, growth and survival of cancer cells within one tumor often depend on the activation of a single oncogene or a single oncogenic pathway [29, 30] . This concept, termed “oncogene addiction”, implies that inhibition of such individual driver pathway in a susceptible tumor would be sufficient to eliminate the malignant phenotype, providing a rationale for targeted cancer therapy aimed at molecular traits driving cancer development and progression [29–31].

As previously mentioned, a number of putatively actionable gain-of-function (GOF) mutations in oncogenes have been identified and, with varying success, utilized for targeted cancer therapy in NSCLC. However, intrinsic or acquired resistance of individual tumor subclones to the inhibition of the driver oncogenic signal, resulting in the eventual cancer recurrence, has been a major limitation to all targeted NSCLC therapies tested up to now [4, 19, 32–35].

1.2.1 *EGFR*-activating mutations as therapeutic targets in NSCLC

1.2.1.1 EGFR biology

The *EGFR* gene encodes for Epidermal Growth Factor Receptor (EGFR), a transmembrane protein that belongs to the ERBB family of cell surface receptor tyrosine kinases (RTKs) [36] . Besides EGFR, also known as ERBB1/HER1, the ERBB family includes three additional RTKs ERBB2/HER2, ERBB3/HER3 and ERBB4/HER4 [36, 37]. Genes encoding ERBB receptors, most prominently *EGFR*, are proto-oncogenes that are often found overexpressed and/or mutated in human cancers, which correlates with poor prognosis, low survival rates, metastasis and drug resistance [37].

Under normal physiological conditions, EGFR forms homodimers in response to the binding of specific ligands to its extracellular region. The EGFR homodimerization activates its intracellular tyrosine kinase domain, resulting in the autophosphorylation of tyrosine residues within the cytoplasmic tail of the EGFR and ultimately leading to the downstream initiation of numerous signaling pathways promoting cell survival, proliferation, differentiation, and motility [36, 38]. Moreover, activated EGFR can translocate into the nucleus, where it acts as

a transcriptional coactivator enhancing the expression of several cell proliferation and survival factors [39]. Additionally, nuclear EGFR has been shown promote cell proliferation by phosphorylating and thus stabilizing the proliferating cell nuclear antigen (PCNA) and enhance the repair of DNA double strand breaks by interacting with DNA-dependent protein kinase (DNA-PK) [39].

1.2.1.2 The prevalence and landscape of EGFR mutations in NSCLC

In NSCLC, *EGFR* mutations lead to a constitutive receptor activation in the absence of ligand binding [36, 38]. Activating mutations in the *EGFR* gene are very common in LUAD, occurring in 40 – 60% of patients of South-Eastern descent and 10 – 20% of patients of Caucasian descent, and are rarely detected in LUSC [4, 40, 41]. Interestingly, activating *EGFR* mutations are detected more frequently in LUAD patients with no history of smoking [42, 43].

The majority of *EGFR* mutations detected in LUAD occur within exons 18-21 encoding its tyrosine kinase domain [40]. In LUAD, 90% of *EGFR* mutations correspond to short in-frame deletions in exon 19 and a missense point mutation in exon 21, resulting in a leucine-to-arginine substitution in codon 858 (L858R) [44–46]. These mutations promote malignant transformation and cancer cell growth through overactivation of a number of downstream signaling pathways, including PI3K/AKT, MEK/ERK (MAPK/ERK) and JAK/STAT, which results in ligand-independent cancer cell addiction to EGFR signaling [47–50].

1.2.1.3 EGFR-TKIs: clinical use and drug resistance mechanisms in NSCLC

Currently, five tyrosine kinase inhibitors (TKIs) are approved for the treatment of *EGFR*-mutant NSCLC: first-generation reversible TKIs gefitinib and erlotinib, second-generation irreversible TKIs afatinib and dacomitinib, and the latest third-generation irreversible TKI osimertinib [40, 51]. In line with the oncogene addiction model, activating *EGFR* deletions in exon 19 and the L858R mutation sensitize NSCLC to all EGFR-TKIs [52]. As such, the EGFR-TKI monotherapies demonstrate improved response rates and progression free-survival (PFS) compared to standard chemotherapy in patients with advanced *EGFR*-mutated NSCLC, with median PFS of 4.9 months reported for latter and 9.2 – 18.9 months for different first-line TKIs [51, 53]. However, despite their initial efficacy, virtually all patients inevitably develop resistance to TKI treatments resulting in a drug resistant cancer relapse [4,

40, 54–56]. Moreover, a fraction of NSCLC patients exhibit intrinsic resistance to EGFR-TKIs [57].

Both intrinsic and acquired resistance to EGFR-TKIs can be conferred via various EGFR-dependent and -independent mechanisms [55, 57, 58]. The latter include the deregulation of other receptor tyrosine kinases, e.g., amplification of the *ERBB2* (*HER2*) gene or activation of ERBB3 via the interaction with MET in the context of *MET* gene amplification, or aberrant activation of their downstream signaling mediators. Ultimately, these events result in EGFR-independent modulation of oncogenic signaling through overactivation of PI3K/AKT or MAPK/ERK molecular pathways, and thus compensate for the EGFR inhibition [58]. In the case of the former, acquisition of secondary mutations in the *EGFR* gene interferes with the interaction between the compound and the tyrosine kinase domain of the EGFR, resulting in the treatment failure [58]. Moreover, resistance to EGFR-TKIs has been associated with increased accumulation of nuclear EGFR, which can result in the enhanced transcription of oncogenes, DNA replication and DNA damage repair [39].

Majority of the cases of acquired resistance to first- and second-generation TKIs are linked to the occurrence of a secondary "gatekeeper" T790M mutation in exon 20 of the *EGFR* tyrosine kinase domain, that has been shown to sterically hinder the binding of TKIs to the EGFR protein [59, 60]. Third-generation TKI osimertinib has been developed to specifically target the T790M *EGFR*-mutant NSCLC resistant to first- and second-generation TKI treatments [61]. Moreover, osimertinib has recently been approved as a first-line therapeutic option for advanced *EGFR*-mutant NSCLC regardless of the T790M status due to its superior clinical performance compared to previous generation TKIs. As such, a Phase 3 clinical trial reported a median PFS of 18.9 months in *EGFR*-mutant NSCLC patients treated with osimertinib compared to 9.2 – 14.7 months reported for the first and second generation TKIs [51, 62, 63].

However, despite its initial high efficacy, drug resistance inevitably develops in response to osimertinib via a number of mechanisms, including additional *EGFR* mutations, the aforementioned *MET* and *ERBB2 (HER2)* gene amplifications and/or the alternative activation of MAPK/ERK and PI3K/AKT signaling [60, 64–66]. Thus, further research efforts are urgently required in order to identify novel therapeutic possibilities for *EGFR*-mutant LUAD treatment.

1.2.2 KRAS-activating mutations in NSCLC

1.2.2.1 KRAS biology

The *KRAS* gene together with two additional genes *HRAS* and *NRAS* encode for four highly homologous isoforms of the RAS protein, with two isoforms resulting from alternative splicing of the *KRAS* gene product [67]. RAS is a membrane-bound small GTPase that acts as an on/off molecular switch of the Ras/Raf/MEK/ERK (MAPK/ERK) and PI3K/AKT signaling pathways. Thus, RAS plays a key role in the control of a vast array of cellular processes, such as cell cycle progression, proliferation, survival, differentiation and migration [68] .

RAS signaling can be initiated by the activation of a number of cellular receptors, including receptor tyrosine kinases (RTKs), G-protein coupled receptors (GPCRs), and integrin family members, which results in the conversion of RAS from the GDP-bound "inactive" state to the GTP-bound "active" state [69]. In turn, the activated RAS promotes the activation of its downstream effectors, most prominently RAF and PI3K [68]. Under normal physiological conditions, RAS activation is tightly regulated and is turned off by the hydrolysis of GTP to GDP [68]. The product of *NF1* tumor suppressor gene, neurofibromin, acts as a negative regulator of RAS by increasing the rate of GTP hydrolysis, and is often found inactivated by mutations in many human cancers, including NSCLC [70, 71].

1.2.2.2 The prevalence and landscape of KRAS mutations in NSCLC

KRAS, *HRAS* and *NRAS* are proto-oncogenes frequently mutated across many cancers, which results in their constitutive overactivation that drives tumor progression. Overall, mutations in the *RAS* genes are the most common genetic aberrations found in human cancers, with different tumor types exhibiting preferential oncogenic activation of different *RAS* isoforms [67, 72]. In NSCLC, *KRAS* gain-of-function alterations account for approximately 25% of all oncogenic mutations detected in Western populations, making it the most frequently altered oncogene in NSCLC [73]. In LUAD patients, *KRAS* mutations occur in 20 – 40% of cases, with a prevalence that is higher in Western populations compared to Asian populations and in smokers compared to non-smokers [74].

The most frequent *KRAS* mutations in NSCLC occur within codons 12 and 13 and result in missense substitutions [74]. The most common G12C substitutions are detected in 41% of *KRAS*-mutant NSCLC and are strongly associated with smoking, while the G12D substitutions

are more commonly detected in patients with no prior history of smoking [73]. These mutations impair GTP hydrolysis, resulting in the constitutive activation of RAS, which, in turn, promotes uncontrolled cell proliferation and survival of *KRAS*-mutant cancer cells [73, 75–77].

Interestingly, not all *KRAS*-mutant NSCLC are addicted to KRAS signaling [73]. As such, *KRAS*-mutant NSCLC cells that are addicted to the oncogenic KRAS are associated with an epithelial gene expression signature, while *KRAS*-mutant NSCLC that do not exhibit dependence on KRAS are more mesenchymal, suggesting a possible link between epithelial-to-mesenchymal transition (EMT) and the loss of oncogene addiction [76, 78].

1.2.2.3 KRAS targeted therapies in NSCLC

Despite the prevalence of *KRAS* mutations in human cancers, KRAS has long been considered "undruggable" due to the absence of accessible drug-binding sites in its protein structure. Currently, the standard treatment of *KRAS*-mutant NSCLC relies on conventional methods approved for patients with NSCLC without actionable driver mutations [76]. However, several novel inhibitors selectively targeting the G12C mutation have recently entered clinical trials [79]. In May 2021, one of such inhibitors sotorasib has received the FDA approval for the treatment of patients with previously treated *KRAS* G12C-mutated NSCLC [80, 81]. Different inhibitors utilize different modes of action, ultimately resulting in the stabilization GDP-bound inactive state of the G12C-mutant KRAS protein [73, 82]

Additionally, several other KRAS-targeting approaches are currently undergoing clinical trials. As such, a genetic depletion strategy utilizes the antisense nucleotide AZD4785 against the *KRAS* mRNA resulting in a decrease of KRAS protein levels [83]. Another tested compound inactivates the KRAS protein indirectly by inhibiting its interaction partner SOS1 that normally acts as a guanine exchange factor. Blocking of SOS1 prevents the phosphorylation of KRAS-bound GDP and thus blocks the KRAS protein in its inactive form [79].

1.2.3 *TP53* loss-of-function vs. gain-of-function mutations in cancer

The *TP53* tumor suppressor gene encodes for the "guardian of the genome" p53 protein that acts as a central regulator of cell cycle, apoptosis, and DNA repair in response to DNA damage and other diverse types of stress stimuli [84]. The major mode of p53 action is

its activity as a transcription factor controlling the expression of a plethora of genes with distinct biological functions [85]. The expression of p53 is primarily controlled at a post-transcriptional level through regulation of its protein stability [86]. Loss of p53 expression has been characterized as a prerequisite for tumor initiation and/or progression in a number of human cancers [87].

Somatic *TP53* mutations occur in > 50% of human cancers and predominantly result in single amino acid substitutions within its DNA-binding domain [88]. In NSCLC, *TP53* mutations are frequent in both LUAD and LUSC with significantly higher prevalence in LUSC, detected in 46% and 90 % of cases, respectively [4, 15]. However, in addition to loss-of-function (LOF) mutations inactivating its tumor suppressor functionality, a number of gain-of-function (GOF) *TP53* mutations, rendering the p53 protein tumorigenic, have been identified and characterized [87–90]. GOF mutations enhance cancer development and progression by stabilizing p53 and altering both its protein interactions and transcription factor activity, resulting in the activation of oncogenic pathways [88, 90]. Therefore, GOF *TP53* mutations have been attracting attention as putatively actionable for therapeutic targeting [91].

1.2.4 Other actionable mutations in NSCLC

A vast array of studies carried out in the last two decades suggest clinical utility of targeted therapies, mainly various tyrosine kinase inhibitors, against additional oncogenes activated in NSCLC by genomic alterations. A number of them are used in the clinic as first-line treatments for oncogene-driven NSCLC and exhibit efficient initial therapeutic responses. However, all targeted therapies available to date are universally susceptible to the development of acquired resistance [4, 92, 93]. Below are examples of actionable NSCLC driver mutations targeted in the current clinical practice.

1.2.4.1 ALK and ROS1 gene rearrangements

The *ALK* and *ROS1* genes encode for receptor tyrosine kinases whose tyrosine kinase domains share a significant homology [94, 95]. Both *ALK* and *ROS1* genes are activated by genomic translocation events resulting in the emergence of tumorigenic fusion proteins in up to 8% and 2% of NSCLC patients, respectively [4, 13, 95, 96]. Although *ROS1* and *ALK* rearrangements are mutually exclusive, *ALK*- and *ROS1*-positive NSCLC share similar

clinicopathological features, such as prevalence towards LUAD histology, young age and low smoking index [94–96]. As such, ALK-TKIs have been approved for the treatment of both *ALK*-positive and *ROS1*-positive NSCLC [4]. However, similar to EGFR-TKI-based targeted therapy approaches, cancers treated with ALK-TKIs ultimately acquire therapeutic resistance leading to tumor relapse within several years [4, 97–99].

1.2.4.2 MET mutations or gene amplification

The *MET* proto-oncogene encodes for a tyrosine kinase receptor whose aberrant activation has been linked to poor prognosis in NSCLC [100–103]. *MET* is activated via gene amplification in 3 – 4 % of NCLC patients, and via mutations in exon 14 resulting in exon skipping in 1 – 6 % of cases [103]. The exon 14 skipping removes the site required for the recruitment of a ubiquitin ligase and thus allows MET to escape degradation [103]. In 2020, MET-TKI capmatinib has been approved for NSCLC harboring exon 14 skipping mutations in the *MET* gene [104]. Yet, it demonstrated limited performance in patients with *MET* gene amplification [103]. *MET* amplification is regarded as one of the major drivers of resistance to EGFR-TKI treatments, and thus, strategies for the MET inhibition in the context of *MET* amplification might benefit NSCLC patients with *EGFR*-mutant NSCLC [105]. However, resistance mechanisms to MET inhibition are also currently being elucidated [106] .

1.2.4.3 BRAF activating mutations

Mutations in the *BRAF* gene, encoding for the serine/threonine kinase BRAF, are detected in 1.5 – 3.5 % of NSCLC cases and result in the constitutive activation of the Ras/Raf/MEK/ERK (MAPK/ERK) signaling pathway [107]. As such, a combination of BRAF and MEK inhibitors has been approved for therapeutic management of NSCLC expressing *BRAF* V600E mutation, the most common variant of mutant *BRAF* [107–109]. Moreover, since activating *BRAF* mutations confer NSCLC resistance to EGFR-TKIs, BRAF and/or MEK inhibition is being considered as an approach to target the EGFR-TKI resistance in *EGFR*-mutant NSCLC [110]. However, several resistance mechanisms to the combination of BRAF and MEK inhibitors, resulting in the restoration of the MAPK/ERK pathway, have been described in NSCLC [111, 112] .

Overall, despite their initial efficacy, the insufficient long-term effects of the current targeted therapies on the survival of NSCLC patients call for further research efforts to identify of novel actionable targets for NSCLC treatment.

1.3 The concept of synthetic lethality in the context of NSCLC therapy

1.3.1 Definition of synthetic lethality in cancer

Synthetic lethality refers to cell death induced by a concomitant perturbation of two genes or pathways, whereas disruption of either of the genes alone is non-lethal [113]. In the context of cancer, identification of synthetic lethal partners to tumor-specific mutations provides new opportunities for the discovery of novel therapeutic targets [113]. Traditionally, the concept of synthetic lethality implies an interaction between two genetic loss-of-function (LOF) perturbations, however, in the context of cancer, also gain-of-function (GOF) / loss-of-function (LOF) interactions, i.e., oncogenic activation of one partner gene coupled with inactivation of the other partner gene, as well as GOF/GOF associations are considered due to their potential usability for drug development [114]. As such, mutual exclusivity of *EGFR* and *KRAS* mutations in NSCLC is attributed to their GOF/GOF synthetic lethality [16]. Besides genetic associations, synthetic lethal interactions between gene-drug and drug-drug pairs also provide a valuable basis for drug development [115, 116].

1.3.2 Synthetic lethality and non-oncogene addiction in cancer

Synthetic lethal interactions between cancer-specific mutations and their non-oncogenic synthetic lethal partners are referred to as “non-oncogene addiction” (NOA) [117]. As such, NOA genes are defined as non-mutated genes upregulated in cancer cells that are essential to maintain the tumorigenic state [117]. Inhibition of a product of the NOA gene that forms a synthetic lethal relationship with a cancer-specific mutation would selectively affect tumor cells that harbor the mutation, making it a promising strategy for developing cancer therapeutics with minimal side effects. Additionally, it allows targeting of “undruggable” oncogenes that cannot be pharmacologically inhibited directly due to their molecular structure, as well as generally non-actionable LOF mutations in tumor suppressor genes [28, 118]. Moreover, indirect targeting of cancer mutations via their synthetic lethal

partners has the potential to overcome drug resistance arising in response to conventional targeted therapies [119].

1.3.3 Synthetic lethal interactions as basis for cancer drug development

Synthetic lethal interaction discovered between the poly(ADP-ribose) polymerase (PARP) and the breast-cancer susceptibility genes 1 and 2 (*BRCA1* and *BRCA2*) has led to the development of the first cancer treatment approach exploiting the concept of synthetic lethality [120]. *BRCA1/2* mutations disrupt the homologous recombination mechanism of DNA repair, and additional PARP inhibition further impair DNA damage repair resulting in *BRCA1/2*-mutant cell death [121]. To date, a total of four PARP inhibitors have been approved for clinical use in ovarian, breast, pancreatic and prostate cancers with deleterious *BRCA1/2* mutations [122]. Moreover, synthetic lethality-based combinatorial inhibition of EGFR and BRAF has recently been approved for treatment of *BRAF*-mutant metastatic colorectal cancers [123, 124].

Recent advances in high-throughput genetic screening methods, particularly RNA interference (RNAi) and CRISPR-based genetic perturbation techniques reviewed in Section 1.5, set a path for systematic discovery of novel putative drug targets based on synthetic lethality interactions in human cancer models [113, 118, 125]. In lung cancer, many initial genetic screening-based attempts to discover novel therapeutic targets focused on identifying LOF synthetic lethal partners to mutant *KRAS*, as no effective direct inhibitors of KRAS had been described until very recently [114, 126–131].

1.3.3.1 Mutant KRAS synthetic lethal partners as drug targets in NSCLC

One of the first synthetic lethal partners of mutant *KRAS* identified via an RNAi-based screen in human cancer models was the *STK33* gene encoding the Serine/Threonine Kinase 33 [132]. RNAi-mediated silencing of *STK33* resulted in a selective loss of cell viability in *KRAS*-mutant, but not -wildtype cancer cell lines. However, a selective kinase function inhibitor developed for STK33 failed to reproduce the synthetic lethality phenotype in KRAS-dependent cancer cell lines observed in the context of the *STK33* gene silencing [131]. Such result could potentially be explained by non-kinase functions of STK33 that could be essential for *KRAS*-driven cancer cell survival [115]. Thus, characterization of specific molecular functions of the

products of genes identified via LOF genetic screens is crucial for the efficient design of pharmacological inhibitors and poses an additional challenge for the development of cancer therapeutics [115].

1.3.3.2 Mutant EGFR synthetic lethal partners as drug targets in NSCLC

Since overcoming the resistance to EGFR TKIs remains a key challenge in lung cancer therapy, most *EGFR*-centric genetic screens described to date focused on the identification of genes which, when disrupted in the context of mutant *EGFR*, result in synthetic lethality in combination with EGFR-TKI treatments [133–138]. These studies have identified a number of genetic modifiers of EGFR-TKI sensitivity; however, no therapeutic strategies have yet been developed based on these findings [139]. Outside of the TKI resistance models, screening efforts aimed at the identification of factors mediating the dependence of *EGFR*-mutant NSCLC on oncogenic EGFR signaling have been limited and mostly focused on specific gene subsets, e.g., other oncogenes, tumor suppressors, or protein kinase genes [140, 141]. However, robust identification of synthetic lethal partners of mutant *EGFR* might uncover novel genetic determinants of EGFR-driven NSCLC survival and thus expand the therapeutic options for *EGFR*-mutant NSCLC.

1.4 lncRNAs as tumorigenic factors in lung cancer

The recent booming development of next-generation sequencing (NGS) technologies greatly prompted the discovery and characterization of novel functional non-coding transcripts, i.e., transcripts that lack protein-coding potential. Particularly, it demonstrated the genome-wide transcription of long non-coding RNAs (lncRNAs), which have previously largely remained under the radar due their generally low expression levels [142–144]. LncRNAs constitute the largest non-coding RNA family and are defined by the transcript length exceeding 200 nucleotides [145, 146].

1.4.1 lncRNA classification

Based on their position relative to protein-coding genes, lncRNAs can be categorized into intergenic RNAs, i.e., lncRNAs transcribed from intergenic genomic regions that are

commonly referred to as long intergenic non-coding RNAs (lincRNAs), and intragenic RNAs, i.e., lncRNAs transcribed from genomic loci intersecting other genes. Intragenic lncRNAs can be further classified by their direction of transcription relative to the genes they intersect into sense and antisense. Both intergenic and intragenic lncRNAs can be transcribed from independent promoters as well as promoters shared with their neighbor genes, including bidirectional promoters. Intragenic lncRNAs are often transcribed from internal promoters located within the bodies of genes they intersect [143, 146, 147].

1.4.2 lncRNA expression and function

Expression of lncRNAs is tightly controlled at the transcriptional and epigenetic level, although majority of lncRNAs are expressed at very low abundance, often below one copy per cell [143, 144, 148, 149]. Expression of many lncRNAs is dysregulated between normal and cancer cells, suggesting their involvement in carcinogenesis and tumor progression [150–154]. Particularly, a number of lncRNAs have been implicated in lung cancer pathogenesis, e.g., *MALAT1* and *AFAP1-AS1* [149, 155–158].

Individual lncRNAs may localize to different cellular compartments, interact with nucleic acids or proteins, and mediate a multitude of diverse cellular processes, with most lncRNA functions characterized to date attributed to regulation of gene expression at transcriptional or post-transcriptional levels. However, specific cellular and molecular functions of the majority of lncRNAs still remain uncharacterized [145, 146, 149, 159, 160].

1.4.3 Systematic identification of lncRNAs

The advent of high-throughput genetic screening approaches(described in Section 1.5) allowed systematic identification of functional lncRNAs in various biological processes. As such, a high-throughput RNAi-based screen previously conducted in the Diederichs lab has identified several lncRNAs essential for cancer cell viability, particularly *VELUCT, LINC000673* and *LINC00152* [148, 161, 162]. However, RNAi has a limited utility for lncRNA characterization, as discussed in Section 1.5.1.

Recent advancements of the CRISPR-based genome modification technologies, particularly the development of the CRISPR interference (CRISPRi) system of targeted transcriptional repression described in detail in Section 1.5.2, provided a novel tool for

probing lncRNA functionality in a high-throughput manner [163–165]. Several CRISPRi-based screens for functional lncRNAs have been conducted in various cancer models [165, 166]. However, CRISPRi screens focused on the lncRNA profiling in the context of NSCLC have not been reported to date.

1.5 Strategies for high-throughput genetic screening for cancer vulnerabilities

Genetic loss-of-function (LOF) screens are a powerful tool to identify molecular determinants of different biological processes. As such, negative selection (dropout) screens allow identifying factors essential for cell survival and proliferation, as their LOF leads to the depletion of the affected cells over time. Dropout screens are therefore widely utilized to pinpoint cancer vulnerabilities. On the contrary, positive selection (rescue) screens are used to identify genetic perturbations that confer enrichment of the affected cells in response to a selective pressure, e.g., drug treatment or apoptotic stimuli, and are used to identify modulators of the respective phenotypes, e.g., drug resistance or apoptosis [167]. RNA interference (RNAi) has been the gold standard technique for LOF screens for more than a decade at the beginning of the 21st century, until the recent advent of the CRISPR-based technologies [168–170].

1.5.1 RNA interference

1.5.1.1 Endogenous RNAi mechanism of gene expression control

RNA interference (RNAi) technology exploits the endogenous post-transcriptional response to the formation of double-stranded RNA (dsRNA) in the cytoplasm, which get degraded by the RNA-induced silencing complex (RISC) [170]. In mammalian cells the RNAi pathway is predominantly utilized to control endogenous gene expression via microRNA (miRNA) – mRNA pairing [171].

miRNAs are transcribed from the genome in a form of long precursor RNAs (pri-miRNAs). After initial processing of the pri-miRNAs in the nucleus, the resulting pre-miRNAs are transported into the cytoplasm, where they pair with their target mRNA. The resulting long dsRNAs are processed by the cellular RNase III enzyme Dicer into short miRNA duplexes

of 20-30 bp in length. Further, the strands of the miRNA duplexes that are complementary to the target mRNA, termed the “antisense” strands, are loaded to the RISC complex and act as guides for sequence-specific degradation and/or blocking of translation of the target mRNA [171, 172].

1.5.1.2 RNAi for targeted gene depletion

To induce exogenous RNAi, 20-30 bp synthetic double-stranded small interfering RNAs (siRNAs) against the mRNA of interest are introduced into the cells via transfection. Transfected siRNAs are loaded to the RISC complex, resulting in the transient target gene inactivation [173]. This method has been widely used to conduct arrayed LOF genetic screens, where cells are targeted with individual siRNAs grow in separate wells in a multiwell plate [168, 174]. However, it does not allow performing the screens in a pooled fashion, where the cells that receive different siRNAs can grow in one population, due to the lack of a molecular readout enabling to resolve the activity of individual siRNAs in a pool.

Alternatively, short hairpin RNA (shRNA), that act as precursors for siRNA, allow pooled screening as they can be expressed from plasmids which can be stably integrated into the genome via lentiviral transduction [168, 174]. The integrated shRNA sequences can be detected via NGS, and thus, under the condition of a single shRNA integration per cell, changes in representation of reads corresponding to individual shRNAs over the course of the experiment can be identified and linked to the observed phenotypes [168]. Compared to the arrayed setup, pooled screens allow considerably less laborious and more cost-effective scaling up as far as to genome-wide targeting [168]. On the downside, pooled screens are only suitable for phenotypes that can be assessed quantitatively, e.g., cell viability and proliferation [175].

Despite their wide use, the conventional RNAi methods of gene silencing are generally characterized by substantial off-target effects, which can result in false positive observations [176–178]. The strategy of using siRNA pools (siPOOLs) of up to 60 individual siRNAs, where low concentration of each siRNA results in dilution of sequence-specific off-target effects below detection limits, allows the reduction of the RNAi-associated off-target effects [177]. However, similarly to siRNAs, siPOOLs cannot be utilized for pooled genetic screens.

Additionally, RNAi has limited utility for lncRNA silencing. First, many lncRNAs are localized in the nucleus, whereas RNAi machinery is predominantly active in the cytoplasm

[179, 180]. Antisense oligonucleotides (ASOs) inducing RNase H-mediated degradation of target RNA in the nucleus can be utilized as an alternative to RNAi to target nuclear lncRNAs; however, they are complex to design and cannot be utilized for pooled screening [181]. Besides, the functionality of a lncRNA often depends on the act of transcription rather than on the transcript itself, making the RNAi gene silencing approach ineffective [182–184].

1.5.2 The CRISPR toolbox

The recent emergence of CRISPR-based technologies offered a new concept of LOF genetic screening [185–187]. The CRISPR (clustered regularly interspaced short palindromic repeats) system was originally discovered in bacteria as a mechanism of adaptive defense against bacteriophages, and later developed into a toolbox of targeted genome modification techniques [188, 189].

1.5.2.1 CRISPR/Cas9 genome editing

In its simplest form, the CRISPR genome editing tool consists of two components: the Cas9 (CRISPR-associated nuclease 9) enzyme and a single guide RNA (sgRNA). An sgRNA contains a universal Cas9-binding sequence and a short 20 bp protospacer region complementary to the target DNA sequence. Thereby, the sgRNA binds and directs the Cas9 to its target DNA site, where it introduces a double-strand break (DSB) at a precise location defined by the presence of an appropriate protospacer-adjacent motif (PAM) immediately downstream of the target DNA sequence. Cas9 nucleases derived from different species of bacteria recognize different PAM sites, with the most commonly used Cas9 from *S. pyogenes* recognizing the PAM sequence of NGG. DSB are predominantly repaired by an error-prone non-homologous end joining (NHEJ) DNA repair mechanism. NHEJ can generate frameshift mutations in the open reading frames (ORF) of targeted protein-coding genes that result in their aberrant translation, thereby mediating efficient gene knockouts [185, 190–192].

sgRNA-expressing constructs can be delivered into the cells via lentiviral transduction, allowing for the pooled screening setup [186, 191]. Given its design simplicity, robust on-target activity and high fidelity, the CRISPR/Cas9 technology has quickly become the preferred method for genetic screening [186, 193, 194]. Numerous modifications of the conventional

CRISPR/Cas9 system have been developed in the recent years to expand its utility beyond NHEJ-based gene knockouts [189].

1.5.2.2 CRISPR-based lncRNA targeting

The conventional CRISPR/Cas9 system based on targeted editing of the genomic sequence is not utilizable for lncRNA silencing as they are unlikely to be sensitive to sequence alterations due to their lack of ORF [145, 147, 181, 195, 196]. Modified CRISPR/Cas9-based high-throughput screening approaches, such as targeted disruption of the lncRNA splice sites or excision of lncRNA promoters or entire lncRNA loci using paired sgRNAs, exhibited a high risk of perturbation of neighboring and/or intersecting protein-coding genes [181, 195, 197–199] As such, my previously published analysis revealed that the majority of lncRNAs annotated in the human genome might not be amenable to CRISPR excision based on their relative genomic position to protein-coding genes [147]. Moreover, paired sgRNA libraries required for CRISPR excision are complex to design and laborious to scale up, limiting their utility for high-throughput genetic screening [195].

1.5.2.3 CRISPRi transcriptional repression system

Alternatively, a nuclease-deficient version of Cas9 (dCas9), which still possesses its RNA-dependent DNA-binding activity, can be used as a transcriptional roadblock to suppress target gene expression without altering the genomic sequence, making it applicable for lncRNA inactivation [147, 163, 164, 181, 200, 201]. Additionally, the dCas9 can be fused to effector domains, such as KRAB (Krüppel-associated box) domain of the Zinc-Finger Protein 10 (ZNF10) [201]. The KRAB domain catalyzes repressive chromatin modification H3K9me3, resulting in specific and potent suppression of transcription [164, 165, 201, 202].

The resulting targeted gene knockdown system is termed CRISPR interference (CRISPRi) and functions by impeding transcriptional initiation or elongation when targeted within 1000 bp downstream of a transcription start site (TSS), with maximum activity within 100 bp downstream of the target TSS [163, 164]. Thus, it potentially allows to independently target individual gene isoforms originating from different TSSs. However, the genomic position of target TSSs has to be carefully considered in order to avoid off-target disruption of neighbor TSS by the CRISPRi system, particularly relevant for lncRNA targeting [147, 164].

1.6 Lung cancer cell lines as a model to identify cancer vulnerabilities

Cancer cell lines are instrumental *in vitro* model systems for cancer research and drug discovery. Under appropriate culture conditions, authenticated cancer cell lines retain the genetic and molecular properties of the cancer of origin, providing an almost indefinite source of standardized biological material for experimental purposes [203]. Moreover, cell lines are cheaper and much less laborious to handle compared to other model systems, and provide an easy and standardized target for genomic manipulation. Genetic and molecular profiling of authenticated cancer cell lines allows designing robust and comprehensive cancer models based on their specific mutational and/or expression signatures. Conversely, it allows retrospectively correlating experimental outcomes with genetic and/or molecular landscapes of the tested cell lines. As such, the COSMIC Cell Lines Project database contains annotations of somatic mutations for over 1000 standardized cancer cell lines (https://cancer.sanger.ac.uk/cell_lines) [204]. Moreover, the Diederichs lab has recently profiled migration capacity as well as RNA and protein expression levels of over 50 lung cancer cell lines, providing a valuable tool for the experimental design in the field of lung cancer research [205].

2 Aim of the project

The primary aim of my PhD project was to perform a CRISPR interference (CRISPRi)-based negative selection screen to identify novel general and genotype-specific factors of LUAD cell viability. Additionally, I aimed at probing the functionality of lncRNAs as essential LUAD viability factors within the same screen setup. Subsequently, I aimed to validate and characterize the resulting screen hits, i.e., screen targets that significantly affected the viability of LUAD cell lines in a genotype-specific or general manner, in order to establish their potential utility for therapeutic targeting.

[This page intentionally left blank]

3 Materials

3.1 Reagents and chemicals

Table 1. *List of reagents and chemicals used in this study*

Reagent name	Provider	Reference No.
0.05% Trypsin-EDTA	Gibco	25200054
0.25% Trypsin-EDTA	Gibco	25200056
2-Propanol	Sigma-Aldrich	33539
5-Bromo-2'-deoxyuridine, BrdU	Sigma-Aldrich	B5002
Agar-agar	Carl Roth	2266.1
Agencourt AMPure XP magnetic beads	Beckman Coulter	A63881
Ampicillin sodium salt	Sigma-Aldrich	A9518
Blasticidine S hydrochloride	Sigma-Aldrich	15205
Bovine Serum Albumin (BSA) Fraction V	Roche	10735094001
BsrGI-HF	New England Biolabs	R3575
CI-1040	MedChemExpress	HY-50295
CutSmart Buffer	New England Biolabs	B7204S
Dimethyl sulfoxide, DMSO	AppliChem	A3672
DMEM high glucose medium	Sigma-Aldrich	D5671
RNase-free DNAse I set	QIAGEN	79254
dNTP set,100 mM each	ThermoFischer Scientific	R0181
ElectroMAX DH10B	hermoFischer Scientific	1829005
Esp3I	ThermoFisher Scientific	ER0452
Ethanol 99,8%	Sigma-Aldrich	32205
FastDigest Eco31I	ThermoFisher Scientific	FD0293
Fetal calf serum (FCS)	Gibco	10270
FITC-labeled anti-BrdU antibody	Biolegend	364104
GeneRuler 100bp Plus DNA Ladder	Fermentas	SM0322
GeneRuler 1kb DNA Ladder	Fermentas	SM0311

Gibson Assembly Master mix	New England Biolabs	E2611
GSK-923295	MedChemExpress	HY-10299
Hexadimethrine bromide / polybrene	Th. Geyer	SA/H9268/000005
Hydrochloric acid (HCl)	Sigma-Aldrich	30721
KpnI-HF	New England Biolabs	R3142
L-Glutamine	Sigma-Aldrich	G7513
LE Agarose	Biozym	840004
Lipofectamine RNAiMax	Invitrogen	13778150
Luria Broth base, LB	LIFE Technologies	12795027
MEM Eagle medium	PAN-Biotech	P04-08056
NEBNext Ultra II Q5 Master Mix	New England Biolabs	M0544
Opti-MEM I Reduced-Serum Medium	Gibco	31985047
Orange DNA loading dye 6x	LIFE Technologies	R0631
PacI	New England Biolabs	R0547
PEI MAX 4000	Polysciences	24765-1
PhiX control v3	Illumina	15017872
Phosphate-buffered saline, PBS	Sigma-Aldrich	D8537
PowerSYBR Green Master Mix	Applied Biosystems	4367659
Propidium Iodide 95-98%	Sigma-Aldrich	P4170
Puromycin	AppliChem	A2856
Q5 High-Fidelity DNA Polymerase	New England Biolabs	M0491
Random hexamer primers	ThermoFischer Scientific	S0142
RevertAid Reverse Transcriptase	ThermoFischer Scientific	EP0442
Ribolock	ThermoFischer Scientific	EO0382
RNase A	QIAGEN	19101
RPMI-1640 medium (ATCC-modification)	Gibco	A1049101
SH-4-54	MedChemExpress	HY-16975
SOC medium	Invitrogen	15544-034
Sodium dodecyl sulphate (SDS) pellets	Carl Roth	CN30.2
Sodium hydroxide (NaOH) pellets	Sigma-Aldrich	30620

Sodium hydroxinde (NaOH) 1N	Neolab	HI70457
Sodium tetraborate 99%	Sigma-Aldrich	221732
SYBR Safe DNA Gel Stain	LIFE Technologies	S33102
T-5224	MedChemExpress	HY-12270
T4 DNA Ligase	ThermoFisher Scientific	EL0011
T4 DNA Ligase Buffer	Thermo Scientific	B69
T4 PNK	New England Biolabs	M0201
TE buffer pH 8.0	Invitrogen	12090015
Trypan Blue solution	Sigma-Aldrich	T8154
UltraPure distilled water DNase/RNase free	LIFE Technologies	10977049
Wortmannin	MedChemExpress	HY-10197
z-VAD-FMK	MedChemExpress	HY-16658B

3.2 Kits

Table 2. *List of kits used in this study*

Kit name	Provider	Reference No.
Caspase-Glo 3/7	Promega	G8090
Caspase-Glo 8 assay	Promega	G8200
Caspase-Glo 9 assay	Promega	G8210
CellTiter-Glo Luminescent Cell Viability assay	Promega	G7572
GeneJET Gel Extraction kit	ThermoFisher Scientific	K0692
Illumina Ribo-Zero Plus rRNA Depletion kit	Illumina	20037135
NucleoSpin Plasmid Mini Prep kit	Macherey-Nagel	40588
PCR Mycoplasma Test kit	PromoKine	PK-CA91-1048
PureLink Maxiprep Plasmid DNA Isolation kit	Invitrogen	K210017
QIAGEN Blood&Tissue DNA Maxi kit	QIAGEN	51194
Qubit dsDNA HS assay kit	ThermoFisher Scientific	Q32851
Quick-RNA Microprep kit	Zymo Reseach	R1051
Quick-RNA Miniprep kit	Zymo Reseach	R1055
RNeasy Mini kit	QIAGEN	74106

3.3 Consumables

Table 3. *List of consumables used in this study*

Item	Provider	Reference No.
15-mL tubes	Greiner Bio-One	188271
50-mL tubes	Greiner Bio-One	227261
96-well transparent cell culture plates	Greiner Bio-One	655180
96-well white cell culture plates	Greiner Bio-One	655073
24-well cell culture plates	Greiner Bio-One	662160
6-well cell culture plates	Greiner Bio-One	657160
Cell culture plates 6 cm	TPP	TPP93060
Cell culture plates 10 cm	TPP	TPP93100
Cell culture plates 15 cm	TPP	TPP93150
Cell culture flasks T75	Corning	3276
Cell culture flasks T225	Corning	431082
Cell scrapers	Corning	3010
Cell strainers 100 µm	Corning	431752
CellSTACK 5-chamber	Corning	CLS3313-8EA
Combitips Advanced 0.2 mL	Eppendorf	0030089413
Combitips Advanced 1.0 mL	Eppendorf	0030089430
Cryopreservation vial 2 mL	Greiner Bio-One	122263
Cryopreservation vials 4 mL	Greiner Bio-One	127261
Electroporation cuvettes LE, 1 mm electrode gap	VWR Peqlab	71-2010-LE
Eppendorf Safe-Lock Tubes, 1.5 mL	Eppendorf	0030 121.023
Eppendorf Safe-Lock Tubes, 1.5 mL Amber	Eppendorf	0030 120.191
Eppendorf Safe-Lock Tubes, 2.0 mL tubes	Eppendorf	0030 120.094
Filtertips 1-10 µl long, sterile	Neptune	976-010XL-5
Filtertips 1-20 µl, sterile	Neptune	976-020-5
Filtertips 100-1000 µl, sterile	Neptune	976-1250-4
Filtertips 200 µl, sterile	Nerbe Plus	976-200-5

LightCycler 480 Multiwell Plate 384-well for RT-qPCR	Roche	04729749001
MicroAmp Fast Optical 96-Well Reaction Plate for RT-qPCR	Applied Biosystems	4346907
MicroAmp Optical Adhesive Film	Applied Biosystems	4311971
PCR tubes	Biozym	710980
Petri dishes for bacterial culture 15 cm	Sarstedt	821184500
Petri dishes for bacterial culutre 10 cm	Greiner Bio-One	632181
Pipette tips 10 µL (grey)	Sarstedt	701130100
Pipette tips 1000 µL (blue)	Sarstedt	70762100
Pipette tips 200 µL (yellow)	Sarstedt	70760502
Plating beads for bacterial culture	Zymo Research	S1001
Polysterene round-bottom FACS tubes 5 mL	Neolab	352058
PVDF filtration bottles (1L) .45 µm (SteriCup)	Merck-Millipore	SCHVU11RE
PVDF syringe filter units .22 µm (Millex)	Sigma-Aldrich	SLGP033R
PVDF syringe filter units .45 µm (Millex)	Sigma-Aldrich	Z355518
Qubit Assay tubes	LIFE Technologies	Q32856
Reagent Reservoir	Corning	4870
Round-bottom tubes 14 mL	Neolab	352059
Scalpel	Carl Roth	T998.1
Aspiration pipettes	Falcon	357558
Serological pipettes 5 mL	Corning	4051
Serological pipettes 10 mL	Corning	4101
Serological pipettes 25 mL	Corning	4251
Serological pipettes 50 mL	Corning	4501

3.4 Equipment

Table 4. *List of equipment used in this study*

Item	Provider	Location
Agarose gel electrophoresis system	Biometra	AG Diederichs
Automated cell counter T20	Bio-Rad	AG Diederichs
BD Fortessa Flow Cytometer	BD Biosciences	Lighthouse Core Facility Freiburg
CoolCell cell freezing box 432001	Corning	AG Diederichs
Dynamag-2 magnetic rack 12321D	Invitrogen	AG Diederichs
Eppendorf BioPhotometer spectrophotometer	Eppendorf	AG Diederichs
EVOS FL Fluorescence micropscope	LIFE Technologies	AG Diederichs
Gene Pulser Electroporation system	Bio-Rad	Dept. of Internal Medicine, University Clinic Freiburg
Gilson Classic P10 pipette	Gilson	AG Diederichs
Gilson Pipetman kit	Gilson	AG Diederichs
Heatblock neoBlock-HeizerDuo	Neolab	AG Diederichs
Hemacytometer	Neubauer	AG Diederichs
HERACell 240 I cell culture incubator	ThermoFisher Scientific	AG Diederichs
Heraeus Megafuge 16 centrifuge	ThermoFisher Scientific	AG Diederichs
Heraeus Megafure 16R cooling centrifuge	ThermoFisher Scientific	AG Diederichs
HiSeq 2000	Illumina	DKFZ Genomics and Proteomics Core Facility
HiSeq 4000	Illumina	DKFZ Genomics and Proteomics Core Facility
Infors HT Ecotron bacterial incubator / shaker	Infors HT	AG Diederichs
Intas Gel Visualization system	Intas	AG Diederichs
Lab Armor bead bath	Lab Armor	AG Diederichs
LighCycler 480	Roche	Lighthouse Core Facility Freiburg
Micro star 17 tabletop centrifuge	VWR	AG Diederichs
Micro star 17R cooling tabletop centrifuge	VWR	AG Diederichs
Microplate reader TriStar2 S LB 942	Berthold Technologies	AG Diederichs

Multichannel pipette 12-channel	Eppendorf	AG Diederichs
Multipette E3 multistep pipette	Eppendorf	AG Diederichs
NanoDrop One microvolume spectrophotometer	ThermoFisher Scientific	AG Diederichs
nanoPAC-300 Power supply	Cleaver Scientific	AG Diederichs
Nikon Eclipse TS2 microscope	Nikon	AG Diederichs
Nuaire CELLGARD ES Class II Bilogical Safety cabinet for cell culture	Nuaire	AG Diederichs
Orbital shaker DOS-10L	Neolab	AG Diederichs
pH-meter Accumet AE150, with electrode 13-620-299B	Fisher Scientific	AG Diederichs
Pipetboy Pro	Interga	AG Diederichs
Qubit 3 Fluorometer	Invitrogen	AG Diederichs
SpeedVac concentrator	Eppendorf	CCI/ITG Freiburg
StepOnePlus Real-Time PCR System	Applied Biosystems	AG Diederichs
Thermal cycler T100	Bio-Rad	AG Diederichs

3.5 Software

Table 5. *List of software used in this study*

Software	Provider
FlowJo v10	BD Biosciences
GraphPad Prism	GraphPad Software
ICE Microplate Reader software	Berthold Technologies
ImageJ Image Analysis software	National Institutes of Health (NIH)
LighCycler 480 software	Roche
Microsoft Office	MicroSoft
SnapGene Viewer	SnapGene
StepOne software	Applied Biosystems

3.6 Webtools

Table 6. *List of webtools used in this study*

Webtool	Reference	Weblink
cBio data portal for Cancer Genomics	[206, 207]	https://cbioportal.org
COSMIC Cell Lines Project	[204]	https://cancer.sanger.ac.uk/cell_lines
Ensembl hg19/GRCh37 Biomart	[208]	grch37.ensembl.org/biomart
European Galaxy server	[209]	https://usegalaxy.eu
The Atlas of Non-coding RNA in Cancer, TANRIC	[210]	https://www.tanric.org
The Cancer Genome Atlas, TCGA	[15, 18]	https://cancer.gov/tcga
The Toronto Knockout Library portal, TKO	[211]	http://tko.ccbr.utoronto.ca/

3.7 Bioinformatics algorithms

Table 7. *List of published bioinformatics algorithms used in this study*

Algorithm	Reference
Cutadapt	[212]
MAGeCK	[213]
FastQC	[214]
HISAT2	[215]
featureCounts	[216]
DESeq2	[217]

[This page intentionally left blank]

4 Methods

Parts of the text presented in this chapter are revised versions of the text submitted for a patent application to the European Patent Office under Application No./Patent No. 20200810.8-1111.

4.1 CRISPRi screen setup

4.1.1 Screen target selection

4.1.1.1 TCGA dataset analysis

The TCGA-LUAD dataset containing RPKM expression values of genes annotated in the hg19/GRCh37 human genome for 515 tumor and 59 normal lung patient samples was obtained from the cBio data portal (https://cancer.gov/tcga; https://cbioportal.org) and merged with the list HUGO-annotated gene identifiers obtained from Ensembl hg19/GRCh37 through Biomart (grch37.ensembl.org/biomart) [15, 18, 206–208]. Genes mapped to alternative genomic loci and patches in hg19/GRCh37 assembly were removed from the dataset. A total of n = 20501 genes were evaluated in the downstream analysis.

30 out of 515 TCGA-annotated tumor samples exhibited low RNA expression levels of LUAD molecular markers *NKX2-1* and *NAPSA* according to the analysis performed by Dr. Chul-Min Yang and were therefore excluded from the dataset. Four samples were additionally excluded due to sample identifier duplications.

The remaining 484 tumor samples were categorized by Prof. Dr. Sven Diederichs according to the annotations of genomic aberrations (mutations and copy number variations) in cancer driver genes provided by Campbell *et al*. in [15]. Out of 484 LUAD selected tumor samples, 464 were annotated in [15]. Eight cancer driver genes listed by Campbell *et al*. in [15] that were most frequently altered among the analyzed LUAD samples were included in the analysis: oncogenes *KRAS*, *EGFR, ERBB2*, *BRAF* and *MET,* and tumor suppressor genes *TP53*, *CDKN2A* and *NF1*. The following rules for sample stratification based on driver mutational signatures applied:

- in the analysis stratified by the *TP53* status, samples annotated as harboring mutations in *TP53* (n = 246) were compared to the rest of the samples (n = 218).

- In the analysis stratified by the *KRAS* status, samples annotated as harboring non-oncogenic mutations or copy number variations of *KRAS*, *HRAS* or *NRAS* as well as oncogenic mutations in *HRAS* or *NRAS* in the absence of oncogenic *KRAS* mutations (annotated as "oncogene positive alterations"), were excluded. In total, n = 156 tumor samples with oncogenic *KRAS* mutations were compared to n = 291 samples without mutant *KRAS*.
- In the analysis stratified by the *EGFR* status, samples annotated as harboring non-oncogenic mutations or copy number variations of *EGFR* in the absence of oncogenic *EGFR* mutations as well as samples harboring any mutations or high copy number gain of *ERBB2* were excluded. In total, n = 63 tumor samples with oncogenic *EGFR* mutations were compared to n = 361 samples without mutant *EGFR.*
- In the analysis stratified by the *ERBB2* status, samples annotated as harboring any mutation or high copy number gain of *ERBB2* (n = 20) were compared to samples that did not exhibit neither the said *ERBB2* nor any *EGFR* genomic alterations (n = 361).
- In the analysis stratified by the *CDKN2A* status, samples annotated as harboring mutations or high copy number loss of *CDKN2A* (n = 89) were compared to the rest of the samples (n = 375).
- In the analysis stratified by the *NF1* status, samples annotated as harboring additional mutations or high copy number gain of *NF1* in the absence of "oncogene positive alterations" of *NF1* were excluded. In total, n = 40 tumor samples with *NF1* mutations were compared to n = 411 samples without mutant *NF1.*
- In the analysis stratified by the *BRAF* status, samples annotated as harboring non-oncogenic mutations in *BRAF* and *RAF1* mutations were excluded from the comparison. In total, n = 30 tumor samples with oncogenic *BRAF* mutations were compared to n = 419 samples without mutant *BRAF*.
- In the analysis stratified by the *MET* status, samples annotated as harboring non-oncogenic mutations or high copy number gain of *MET* were excluded from the comparison. In total, n = 17 tumor samples with oncogenic *MET* mutations were compared to n = 432 samples without mutant *MET.*

The resulting dataset containing RPKM values for 20501 genes in a total of 464 LUAD samples stratified by eight driver mutational signatures was further analyzed by me using Microsoft Excel. Fold changes in median RPKM values of HUGO-annotated genes between LUAD samples with a particular oncogenic mutation and the rest of tumor samples were calculated and tested for significance using two-tailed unpaired Student's t-test with Welch's correction for unequal variances where appropriate. Genes that exhibited low median expression (RPKM < 1) in mutant LUAD samples were removed from every mutation-stratified dataset to minimize representation of transcriptional noise. Genes that were found significantly overexpressed in mutant tumor samples (FC ≥ 2, $p < 0.05$) or had detectable expression only in mutant tumor samples were selected as targets for the CRISPRi screen.

4.1.1.2 TANRIC dataset analysis

TCGA-LUAD dataset containing RPKM expression values of lncRNAs (annotation based on hg19/GRCh37 human genome assembly) for 488 tumor and 58 normal lung patient samples was downloaded from the TANRIC database (https://www.tanric.org) [210]. 28 out of 488 tumor samples exhibited low RNA expression levels of LUAD molecular markers *NKX2-1* and *NAPSA* according to the analysis performed by Dr. Chul-Min Yang and were therefore excluded from the dataset.

The dataset containing expression values corresponding to 12727 TANRIC-annotated lncRNAs for 460 LUAD samples and 58 normal lung tissue samples was further analyzed using Microsoft Excel. Fold change in median expression between tumor and normal samples across all annotated lncRNAs was calculated and tested for significance using two-tailed unpaired Student's t-test with Welch's correction for unequal variances where appropriate. LncRNAs that were significantly upregulated in tumor samples (FC ≥ 2, $p < 0.05$) or had detectable expression only in tumor samples were selected as targets for the subsequent CRISPRi dropout screen.

Additionally, TANRIC-annotated TCGA-LUAD samples were stratified by mutational signatures and analyzed using the workflow described in Section 4.1.1.1 with the following modifications. Genomic aberrations were annotated by Campbell *et al*. in [15] for 447 out of 460 *NKX2-1* and *NAPSA*-stratified LUAD samples present in the TANRIC dataset, which were therefore used for downstream analysis. All TANRIC-annotated lncRNAs (n = 12727) were analyzed irrespective of HUGO annotations and expression levels.

4.1.2 Selection of positive controls and manual targets

The following categories of target genes were additionally incorporated into the CRISPRi screen:

- Core essential positive controls: a list of 1580 core fitness genes identified by Hart *et al*. in [211] was downloaded from The Toronto Knockout Library portal (http://tko.ccbr.utoronto.ca/) and matched with the RNA expression dataset from the Diederichs lab created by Dr. Maïwen Caudron-Herger. The dataset contained RPKM values for HUGO-annotated genes corresponding to their expression in NCI-H460, A549, HLE and Huh-7 cancer cell lines, as well as the chromatin fraction of the NCI-460 cell line. 20 out of 1580 genes exhibiting the RPKM expression values above the set threshold for each sample were selected as positive controls for the CRISPRi screen (RPKM > 100 for NCI-460 or NCI-H460-chromatin and A549; RPKM > 10 or HLE and Huh-7).
- Driver genes: eight LUAD driver genes used for patient sample stratification by genetic background in the database analysis described in Section 4.1.1.1 and two additional *RAS* genes *HRAS* and *NRAS*.
- lncRNA controls from the siRNA screen: 15 lncRNAs that exhibited the strongest effects on cell viability of lung cancer cell lines in the previous siRNA-based screen conducted in the Diederichs lab [148, 161, 162].
- Manual positive controls: 10 protein-coding genes and 7 lncRNAs selected by me or Prof. Dr. Sven Diederichs that were previously annotated as lung cancer viability factors in the published literature.
- Additional candidates: 17 genes that have been studied in the Diederichs lab within other projects.

In total, 2098 genes were selected as the CRISPRi screen targets after removal of duplicates between different candidate subsets. Out of the 2098 genes, 74 were previously annotated as core essential by Hart *et al*. [211], including 53 genes that were selected from the TCGA dataset as differentially expressed between LUAD patient samples with different

driver mutational signatures. Together with the manually added targets, the screen candidate gene list contained a total of 90 protein-coding genes annotated as “positive controls”.

4.1.3 Custom sgRNA design algorithm

A canonical single guide RNA (sgRNA) consists of a universal trans-activating sequence recruiting the Cas9 into the RNP complex (tracrRNA) and a target-specific protospacer sequence recruiting the RNP complex to the genomic target site (crRNA) [190]. For simplicity, I refer to protospacer sequences as sgRNAs throughout the whole thesis, unless specified otherwise.

The custom pooled sgRNA design pipeline was based on the characteristics of the CRISPRi system defined by Gilbert *et al*. in [164] The algorithm was programmed by Dr. Andrew Walsh based on the hg19/GRCh37 human genome assembly [208]. Design of the algorithm and the final sgRNA library compilation were performed by me.

sgRNAs were designed against genomic regions adjacent to all TSSs annotated for the selected gene targets. TSSs that were located within 200 bp from each other were regarded as a single target site. 4633 target TSSs corresponded to the 2098 selected genes, out of which 378 TSSs corresponded to the 90 positive control genes. Successful sgRNA designs had to fulfill the following criteria:

1. 20 nt long;
2. GC content between 35% and 85%;
3. 5’-adjacent to the canonical NGG PAM sequence;
4. Target PAM sites had to fall within a [-100; +300] bp window around the target TSS (0); PAM sites within a [0; +100] bp window were preferred in the first iteration; PAM sites closest to the [0; +100 bp] window were preferred in the following iterations; PAM sites closest to the target TSS were preferred in all iterations;
5. Have not more than five TSS-specific off-target sites (see below); designs with least TSS-specific off-target sites were preferred in every iteration of the sgRNA selection. Additionally, designs with the least genome-wide off-target sites were preferred in every iteration.

TSS-specific off-target sites were identified by aligning the sgRNA designs against sequences corresponding to [-1000; +1000] bp windows around all TSSs annotated in the hg19/GRCh37 human genome assembly except TSSs corresponding to pseudogenes and TSSs mapping to alternate genomic loci and patches. Alignment matches that fell within 1000 bp around another TSS of the same target gene were not considered as off-targets. The reference sequences were downloaded from Ensembl hg19/GRCh37 via Biomart (grch37.ensembl.org/biomart) [208]. Overlapping sequences were concatenated to produce a TSS-specific off-target reference file. For genome-wide off-target mapping, sgRNA designs were aligned to hg19/GRCh37 human genome assembly.

Additionally, to be assigned as an off-target (both TSS-specific and genome-wide), the matching between the sgRNA and a genomic site had to fulfill the following criteria:

1. The predicted off-target site had to be 5'-adjacent to the canonical NGG PAM sequence;
2. No mismatches were allowed within seven PAM-proximal sgRNA nucleotides (positions 14 – 20);
3. No more than one mismatch was allowed within the following eight nucleotides (positions 6 – 13).

All sgRNAs that fit the design criteria were computed for all target sites, and an average of nine sgRNAs per target site were selected. Number or TSS-specific off-targets and proximity to target TSS were used as primary and secondary selection criteria, respectively.

"G" nucleotide was manually added in the first position to all sgRNA sequences that started with other nucleotides to ensure efficient transcription of all sgRNAs from the U6 promoter [218], resulting in a pooled library of 20 – 21 nt long sgRNAs. In total, 40357 sgRNA designs against 4620 out of 4633 TSSs corresponding to 2084 out of 2098 target genes out of were successfully selected for the sgRNA library.

4.1.4 Selection of negative control sgRNAs

20 – 21 nt long negative control sgRNA sequences were selected from the list provided in the previous publication by Gilbert *et al*. [164] and sorted by the assigned absolute growth

phenotype value from smallest to largest. Top 1643 negative control sgRNA sequences with the smallest absolute growth phenotype values were added to the custom target sgRNA list to produce a pooled single-stranded DNA oligonucleotide library corresponding to 42000 sgRNAs.

4.2 Experimental CRISPRi screening procedures

4.2.1 Cell line selection and general cell culture protocols

All cell lines used in this study were either purchased from ATCC or were already available in the Diederichs lab at the start of this project.

EGFR-mutant and -wildtype LUAD cell lines used for the CRISPRi screen were selected based on their mutational signatures annotated in the COSMIC Cell Lines Project database (https://cancer.sanger.ac.uk/cell_lines) [204]. The PC-9 cell line was not annotated in COSMIC, however, has been widely characterized as *EGFR*-mutant in literature [219–221]. Thus, PC-9, NCI-H1650 and HCC827 LUAD cell lines were annotated as harboring deletions in exon 19 of the *EGFR* gene, and NCI-H1975 cell lines expressed L858R and T790M *EGFR* mutations. The selected *EGFR*-wildtype cell lines Calu-6, NCI-H3122, NCI-H522 and NCI-H838 cell lines also did not harbor mutations in *ERBB2*, *ERBB3* and *ERRB4* genes.

4.2.1.1 Growth and propagation of human cell lines

All lung cancer cell lines used in this study were routinely cultured in RPMI-1640 medium (ATCC-modification, Gibco) supplemented with 10% FCS (Gibco) in a cell culture incubator at 37°C and 5% CO_2. Non-transformed lung cell lines IMR-90 and WI-38 were routinely cultured in MEM Eagle medium (PAN-Biotech) supplemented with 10% FCS. The HEK293T cell line was routinely cultured in DMEM high glucose medium (Sigma-Aldrich) supplemented with 4 mM L-Glutamine (Sigma-Aldrich) and 10% FCS.

For cell passaging, medium was removed from the cell culture dishes, and cells were washed with 1x PBS (Sigma-Aldrich) and incubated with trypsin-EDTA solution pre-warmed to 37°C until the cells were detached from the plastic (3 – 10 minutes, depending on the cell line). 0.05% trypsin-EDTA solution (Gibco) was used for HEK293T cells, 0.25% trypsin-EDTA solution (Gibco) was used for the rest of the cell lines. Afterwards, trypsin was inactivated by

the addition of at least one volume of the respective complete cell culture medium pre-warmed to 37°C, and the resulting cell suspensions were mixed by pipetting. Afterwards, cells were either seeded at the required densities for further downstream experimental procedures, split for further propagation, or frozen. The IMR-90 and WI-38 cells were used for experimental purposes until passage #29.

For cell counting, 10 µL of the cell suspensions were mixed by pipetting with 10 µL of Trypan Blue solution (Sigma-Aldrich) to allow discrimination between viable and non-viable cells. Unless otherwise specified, cells were counted using an automated cell counter.

All cell cultures were routinely tested for mycoplasma contamination using PCR Mycoplasma Test kit (PromoKine).

4.2.1.2 Cryopreservation of cell lines

For long-term storage, cells suspensions were pelleted after trypsin inactivation by centrifugation at 100x*g* for 5 minutes at room temperature. The cell pellets were resuspended in cold complete cell culture medium containing 5% DMSO (AppliChem), aliquoted into cryogenic vials, placed into freezing boxes (CoolCell, Corning) and incubated at -80°C for at least 24 hours. Subsequently, the vials were transferred to liquid nitrogen.

4.2.1.3 Thawing of the cryopreserved cells

To thaw the cryopreserved cells, vials were retrieved from liquid nitrogen and immediately placed into the bead bath at 37°C. After rapid thawing, the cells were resuspended in the pre-warmed complete cell culture medium and pelleted by centrifugation at 100x*g* for 5 minutes at room temperature. The supernatants were removed, the cell pellets were resuspended in fresh pre-warmed complete cell culture medium and plated onto appropriate cell culture dishes.

4.2.2 Plasmid sgRNA library cloning

Flanking sequences for PCR amplification and cloning (5'-flank 5'-TATCTTGTGGAAAGGACGAAACACC-3'; 3'-flank 5'-GTTTAAGAGCTATGCTGGAAACAGCATAGC-3') were added to the 20 – 21 nt sgRNA designs to a final length of 75 – 76 bp. The sgRNA

library was purchased from Twist Bioscience in a form of a lyophilized equimolar single-stranded DNA (ssDNA) oligo pool.

The lyophilized oligo pool was resuspended in TE Buffer pH 8.0 (Invitrogen) to a concentration of 10 ng/µL and stored at -20°C. 1 ng of the oligo pool was amplified by PCR in a 50 µL reaction using Q5 High-Fidelity DNA Polymerase (New England Biolabs), forward primer 5'-GTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC-3'; reverse primer 5'-GTTGATAACGGACTAGCCTTATTTAAACTTGCTATGCTGTTTCCAGCATAG CTCTTAAAC-3' under the following cycling conditions: initial denaturation at 98°C for 30 seconds, followed by 12 cycles of 98°C for 10 seconds, 60°C for 15 seconds and 72°C for 20 seconds, followed by final elongation at 72°C for 2 minutes. In total, 12 reactions were performed to amplify the oligo library. After the amplification, the PCR reactions were pooled, mixed with Orange DNA Loading Dye (LIFE Technologies), and resolved on a 1% agarose gel. The 140 bp PCR product was extracted from the gel using GeneJET Gel Extraction kit (ThermoFisher Scientific).

LentiCRISPRv2-dCas9-KRAB-puro(iv) vector [147] was linearized with Esp3I restriction enzyme (ThermoFisher Scientific). The linearized vector backbone (approx. 13.2 kb) was gel-extracted, and ten 50 µL Gibson assembly reactions were set up using 360 ng of the linearized vector and 50 ng of the gel-purified PCR-amplified oligo pool per reaction (molar ratio 1:13). The Gibson assembly reactions were incubated at 50°C for 1 hour, pooled, split back into ten 50 µL aliquots, and purified using 0.75x Agencourt AMPure XP magnetic beads (Beckman Coulter) according to the manufacturer's instructions. The DNA was eluted from the beads using 13 µL of water. In order to further concentrate the DNA, eluates were mixed into two pools of five, and the second round of purification with 0.75x magnetic beads was performed. The DNA from the second round of purification was eluted into of 13 µL of water; two concentrated eluates were pooled together.

Five negative control Gibson assembly reactions without the addition of oligo pools were performed, pooled, purified and eluted in the same manner.

Twelve electroporation reactions were performed with 1 µL of the eluate and 20 µM ElectroMAX DH10B electrocompetent *E. coli* cells per reaction (ThermoFisher Scientific) using Bio-Rad Gene Pulser electroporation system at 2.0 kV, 25 µF, 200 Ohm in 1 mm gap long electrode electroporation cuvettes (VWR Peqlab). 980 µL of SOC medium (Invitrogen) pre-warmed to 37°C was added to each reaction immediately after the electroporation, and the

bacteria were left to recover at 37°C, 220 rpm for 1 hour. Finally, each electroporation reaction was plated on ten 15 cm LB-agar plates with 100 µg/mL ampicillin (Sigma-Aldrich) and allowed to grow for 14 – 16 hours at 37°C. To collect bacterial colonies, each plate was scraped twice with a cell scraper, each time into 7 mL of LB medium.

Optical density of the resulting bacterial suspension was measured at a wavelength of 600 nm (OD600) to estimate the number of PureLink Maxiprep columns (Invitrogen) needed for DNA isolation according to the manufacturer's instructions. As a result, 16 Maxiprep columns were used to isolate the complete plasmid DNA library. The amount of DNA eluted from each column was quantified using Nanodrop. DNA samples of 500 ng from individual eluates were digested with the FastDigest Eco31I restriction enzyme (ThermoFisher Scientific) as a quality control. Afterwards, the eluates were pooled and stored at 2 µg/µL at -20°C.

4.2.3 Lentivirus production in HEK293T cells

Production of the lentiviral sgRNA library followed the general protocol for the lentivirus production in HEK293T cells described below. Each batch of the lentiviral sgRNA library was produced using eighteen T225 flask.

For lentivirus production, HEK293T cells were seeded 24 hours prior to transfection at a surface area-appropriate density resulting in approximately 60% confluence at the time of transfection indicated in **Table 8**. At the time of transfection, lentiviral packaging plasmid psPAX2 and VSV-G envelope protein-expressing plasmid pMD2.G (gifts from Didier Trono, Addgene plasmids #12260 and #12259, respectively) were mixed with the desired transgene plasmid or library of plasmids at a mass ratio of 2:1:4 in an appropriate volume of Opti-MEM I Reduced-Serum Medium (Gibco, **Table 8**). PEI MAX 4000 (Polysciences) resuspended in water to 1 g/L, filtered through .22 µM filter (Millex, Sigma-Aldrich), divided into 1 mL aliquots and stored long-term at -20°C, at 4°C for up to three month after thawing without repeated freezing) was added to the plasmid mix at a 3:1 w:w ratio, vortexed immediately for 15 seconds, and incubated for 15 minutes at room temperature, resulting in a transfection mix.

Subsequently, the transfection mix was added to an appropriate volume of Opti-MEM, added to the HEK293T cells in place of the cell culture medium, incubated with the cells for 12 hours, and then replaced by an appropriate volume of fresh Opti-MEM (**Table 8**). At 24 hours post-medium change, the first batch of lentivirus-containing supernatant was collected

and stored at 4°C. An appropriate volume of fresh Opti-MEM was added to the cells (**Table 8**). At 48 hours post-medium change, the second batch of lentiviral supernatant was collected, and cells were discarded. Two batches of viral supernatant were mixed and filtered through .45 µm PVDF filter membranes. Syringe filter units (Millex, Sigma-Aldrich) were used to filter up to 30 mL of viral supernatant; the lentiviral library supernatants were filtered using 1L filter bottles (SteriCup, Merck-Millipore). The resulting lentiviral supernatant was aliquoted into cryogenic vials (2 mL or 4 mL), snap-frozen in liquid nitrogen (2 mL vials) or dry ice-ethanol bath (4 mL vials; methods of choice consistent within replicates of the same experiment) and stored at -80°C.

Table 8. *Parameters of PEI-mediated transfection reactions for lentivirus production in HEK293T cells*

Format	6-well plate	T75 flask	T225 flask
HEK293T cell seeding density	$5*10^5$ cells	$5*10^6$ cells	$15*10^6$ cells
Mass of total DNA per reaction	2 µg	17.3 µg	52 µg
Volume of transfection mix	100 µL	1 mL	3 mL
Total transfection volume	700 µL	6 mL	18 mL
Volume of Opti-MEM per timepoint	1 mL	6 mL	18 mL

4.2.4 Lentiviral transduction and antibiotic selection

4.2.4.1 General protocol

Cell were seeded 24 hous prior to transduction at a density resulting in a 40 % (for fast-proliferating cells) – 70 % (for slow-proliferating cells) confluence at the time of transduction (**Table 9**).

At the time of transduction, the desired volume of virus for added to the complete cell culture medium to a final volume appropriate for the cell culture plate format: 400 µL/well on a 24-well plate, 2 mL/well on a 6-well plate, 15 mL per T75 flask, 45 mL per T225 flask, 650 mL per 5-stack CellSTACK chamber. Polybrene was added to the transduction mix to a final concentration of 8 µg/mL. Polybrene (hexadimethrine bromide, Th. Geyer) was resuspended in water to 8 mg/mL, filtered through .22 µM filter (Sigma-Aldrich), divided into 1 mL aliquots

and stored long-term at -20°C, at 4°C for up to three months after thawing without repeated freezing.

After 20 hours, depending on the cell line, either transduction mix was replaced by fresh complete cell culture medium, or the cells were split back to their original seeding densities. In case of the former, cells were split back to their original seeding densities after additional 24 hours. 24 hours after splitting, the cell culture medium was replaced by the medium containing an appropriate concentration of antibiotic. Antibiotic concentration that resulted in complete cell death within 48 – 72 hours was determined for every cell line and used throughout all experiments, unless specified otherwise. For puromycin selection, 2 µg/mL puromycin (AppliChem) was used to select the stably transduced HCC827 cells, 1.5 µg/mL puromycin to select NCI-H3122, PC-9 and NCI-H1650 cells, 1 µg/mL puromycin for NCI-H522 and NCI-H1975 cells, 0.5 µg/mL puromycin for Calu-6 and NCI-H838 cells. For blasticidin selection, 10 µg/mL Blasticidin S hydrochloride (Sigma-Aldrich) was used for Calu-6 cells. Puromycin and Blasticidin S hydrochloride were resuspended in water at 1 mg/mL and 2.5 mg/mL, respectively, filtered through .22 µM filter (Millex, Sigma-Aldrich), aliquoted and stored at -20°C long-term and 4°C up until three months after thawing without repeated freezing.

Table 9. *Cell seeding densities for lentiviral transduction per cell line*

Cell line \ Format	24-well plate; cells/well	6-well plate; cells/well	T75 flask; cells/flask	T225 flask; cells/flask	CellSTACK 5-stack chamber; cells/chamber
Calu-6	$7*10^4$	$3.5*10^5$	$3*10^6$	$9*10^6$	$12.6*10^7$
NCI-H3122	$7*10^4$	$3.5*10^5$	$3*10^6$	$9*10^6$	$12.6*10^7$
NCI-H522	$7*10^4$	$3.5*10^5$	$3*10^6$	$9*10^6$	$12.6*10^7$
NCI-H838	$1.8*10^4$	$9*10^4$	$1*10^6$	$3*10^6$	$4.2*10^7$
PC-9	$2.4*10^4$	$1.2*10^5$	$1.6*10^6$	$4.8*10^6$	$6*10^7$
HCC827	$2*10^4$	$1*10^5$	$1.4*10^6$	$4.2*10^6$	$6*10^7$
NCI-H1650	$2*10^4$	$1*10^5$	$1.4*10^6$	$4.2*10^6$	$6*10^7$
NCI-H1975	$3.6*10^4$	$1.8*10^5$	$2.5*10^6$	$7.5*10^6$	$10.8*10^7$

4.2.4.2 Transduction of the lentiviral sgRNA library

Transduction of the LUAD cell lines with the lentiviral sgRNA library followed the general lentiviral transduction protocol described above with the following modifications.

Volume titration of every batch of the lentiviral sgRNA library was performed to determine the appropriate volume of viral supernatant required to achieve the multiplicity of infection (MOI) of ≤ 0.3 for every screened cell line, which was consistent throughout all the batches produced. Tested cell lines were plated on 6-well cell culture plates at appropriate densities (**Table 9**). Four out of six wells were transduced with varying volumes of the lentiviral sgRNA library and subjected to the puromycin selection according to the general protocol described above. One of the untransduced wells was used as a control for antibiotic selection, the second untransduced well was used as a positive viability control and was not treated with puromycin. After 48 hours of antibiotic selection (upon complete death of the cells in the respective control well), cells in the test wells and the viability control well were manually counted in a hemacytometer (Neubauer) with the addition of Trypan Blue solution (Sigma-Aldrich) to discriminate between viable and non-viable cells. The volume of virus corresponding to the 20-30% of the cell viability registered in the control well was extrapolated according to the number of cells seeded for transduction of the CellSTACK cell culture format compared to the 6-well format (**Table 9**) and further used in the screening experiments.

4.2.5 Negative selection screens

The CRISPRi screens were performed in 5-stack CellSTACK cell culture chambers (Corning) to maintain the coverage of at least 500 cells per sgRNA for all tested cell lines. For each tested cell line, the screen was performed in two biological replicates. For the NCI-H838 cell lines, two CellSTACK chambers were used per biological replicate; cells from both chambers were pooled at every cell passaging round. Additionally, a T75 flask was seeded from the screened cells at every cell passaging round to enable monitoring of the cell density using light microscopy.

20 hours after lentiviral transduction, the lentivirus-containing medium was replaced by fresh complete cell culture medium. 24 hours later, the cells were seeded back to the densities indicated in **Table 9**. 24 hours later, the cell culture medium was replaced by fresh

medium containing the appropriate concentration of puromycin indicated in Section 4.2.4.1. Antibiotic selection was carried out for 72 – 96 hours, consistently within the replicates of the same cell line, to ensure complete removal of untransduced cells. In total, cells were passaged for 21 days after transduction, consistently within the replicates of the same cell line. At every passaging round, cells were washed with 250 mL of PBS, trypsinized using 100 mL of 0.25% trypsin-EDTA solution for 12 – 15 minutes to ensure compete cell detachment, followed by the neutralization of trypsin with 100 mL of complete cell culture medium. The resulting cell suspension was divided between an even number of 50 mL tubes and centrifuged for 5 minutes at 100x*g*. After centrifugation, the supernatant was removed, the pellets were resuspended and pooled together in 25 mL of cell culture medium. Cells in the resulting suspension were manually counted in a hemacytometer (Neubauer) with the addition of Trypan Blue solution (Sigma-Aldrich) to discriminate between viable and non-viable cells. The appropriate number of cells corresponding to the coverage of at least 500 cells per sgRNA was resuspended in the cell culture medium to a total volume of 650 mL, and seeded back to the CellSTACK chamber.

At 21 days, cells were washed with PBS, trypsinized, counted, divided into aliquots of at least $42*10^6$ cells per tube corresponding to the coverage of at least 1000 cells per sgRNA in the starting library and pelleted for 5 minutes at 100x*g*. Supernatant was removed, cell pellets were snap-frozen in liquid nitrogen and stored at -20°C prior to genomic DNA (gDNA) isolation.

4.2.6 gDNA isolation

Prior to the gDNA isolation, the frozen cell pellets were equilibrated to room temperature for 10 minutes. The gDNA isolation was performed using QIAGEN Blood&Tissue DNA Maxi kit (QIAGEN) with following modifications.

No more than $50*10^6$ cells were used per column. Cell pellets were resuspended in 10 mL PBS. 50 µL of RNase A (QIAGEN) was added to the cell suspension, briefly vortexed, and incubated for 10 minutes at room temperature. Next, 500 µL of Proteinase K supplied in the kit was added to the cell suspension, briefly vortexed, and incubated for 5 minutes at room temperature. 12 mL of buffer AL was added to the resulting suspension; the tubes were inverted 15 times followed by vigorous shaking and vortexing for 2 minutes and incubated at

70°C for 30 minutes. Next steps were carried out according to the manufacturer's centrifugation-based (spin) protocol. The DNA was eluted from the columns in two rounds. Every round, 1 mL of distilled room temperature water was incubated on the column for 5 minutes, followed by centrifugation at 4500x*g* for 2 minutes in the first round and 5 minutes in the second round. Eluates from both rounds were combined, and the DNA concentration was quantified using Nanodrop. The combined eluate was concentrated on SpeedVac at 60°C for 60 minutes under vacuum.

4.2.7 Next-generation sequencing library preparation

The NGS library preparation protocol was based on the general protocol described by Shalem *et al.* in [186] entailed two rounds of PCR amplification.

4.2.7.1 First-round PCR

All gDNA isolated from at least $42*10^6$ cells corresponding to the coverage of at least 1000 cells per sgRNA was used as a template for the first round of PCR, using 5 µg of gDNA per reaction. The first-round PCR was performed using NEBNext Ultra II Q5 Master Mix (New England Biolabs) in a total reaction volume of 50 µL with the following primers: forward primer 5'-AATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCG-3'; reverse primer 5'-TCTACTATTCTTTCCCCTGCACTGTTGTGGGCGATGTGCGCTCTG-3';
under the following cycling conditions: initial denaturation at 98°C for 3 minutes, followed by 18 cycles of 98°C for 30 seconds and 65°C for 75 seconds, followed by final elongation at 65°C for 5 minutes. Approximately 60 first-round PCR reactions were performed per sample and subsequently pooled together.

To prepare NGS libraries from the reference Plasmid sgRNA library (4.2.2), 20 ng of the plasmid library were used as a template per 25 µL first-round PCR reaction using Q5 High-Fidelity DNA Polymerase (New England Biolabs) performed under the following cycling conditions: initial denaturation at 98°C for 30 seconds, followed by 8 of 98°C for 10 seconds, 60°C for 15 seconds, 72°C for 30 seconds, followed by final elongation at 72°C for 2 minutes. Two sets of refence NGS libraries were prepared; for each, 8 first-round PCR reactions were performed and subsequently pooled together.

All PCR reactions performed per one sample were pooled together and subjected to electrophoresis on 1.2 % agarose gels. The DNA band of the expected size (approx. 300 bp) was gel-excised and gel-extracted using GeneJET Gel Extraction kit (ThermoFisher Scientific) according to the manufacturer's instructions. The resulting first-round PCR product was quantified on Qubit using dsDNA HS assay kit (ThermoFisher Scientific).

4.2.7.2 Second-round PCR

The second-round PCR was performed to attach Illumina sequencing adapters and indexes for sample multiplexing in a total reaction volume of 25 µL using 10 ng of the first-round PCR product, Q5 High-Fidelity DNA Polymerase (New England Biolabs, M0491) and the following primers: forward primer 5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCC TACACGACGCTCTTCCGATCT$(N)_{6-9}$GGAAAGGACGAAACACCG -3'; reverse primer 5'-CAAGCAGAAGACGGCATACGAGAT(Index)GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGC TGTTTCCAGCATAGCTC-3'. The forward primer contained degenerate sequences $(N)_{6-9}$ of 6-9 bp to improve cluster generation on the Illumina sequencing platforms. Reverse primers contained Illumina TruSeq LT indexes (Index) indicated in **Table 10** and were ordered from Sigma-Aldrich as PAGE-purified oligonucleotides. The second-round PCR was performed under the following conditions: initial denaturation at 98°C for 30 seconds, followed by 15 cycles of 98°C for 10 seconds, 67°C for 15 seconds, 72°C for 20 seconds, followed by final elongation at 72°C for 2 minutes.

5 uL of each completed second-round PCR reaction were simultaneously subjected to electrophoresis on a single 1 % agarose gel. The bands were visualized and photographed. The relative amounts of DNA in the samples were evaluated from the gel image using ImageJ software. Based on those evaluations, the completed second-round PCR reactions were pooled in equimolar proportions into two multiplexes of 9 samples. Both multiplexes contained the PCR products from the reference Plasmid sgRNA library.

The multiplexes were purified using 1.6x Agencourt AMPure XP magnetic beads (Beckman Coulter) according to the manufacturer's instructions. The resulting multiplexed NGS libraries were quantified on Qubit and used for NGS.

Table 10. *Illumina TruSeq LT indexes used for NGS sample multiplexing*

TruSeq LT Index	Sequence
Index 1	CGTGAT
Index 2	ACATCG
Index 3	GCCTAA
Index 4	TGGTCA
Index 5	CACTGT
Index 6	ATTGGC
Index 7	GATCTG
Index 8	TCAAGT
Index 9	CTGATC
Index 10	AAGCT
Index 11	GTAGC
Index 12	TACAAG

4.2.8 Next-generation sequencing

The NGS libraries were spiked with 15% PhiX Control v3 (Illumina) to increase the read diversity. All samples were sequenced at the coverage corresponding to 500 reads per sgRNA in the reference library. The single-end 50 bp sequencing was performed on Illumina HiSeq 2000 platform at one multiplexed library per flow cell lane in the DKFZ Genomics and Proteomics Core facility.

4.2.9 Next-generation sequencing analysis

First, the sequencing reads were trimmed down to the sgRNA sequences using the Cutadapt tool [212]. Afterwards, the NGS screen results were analyzed using the MAGeCK algorithm [213]. Negative control sgRNA (n = 1643) were randomly pooled into sets of 7-9 sgRNAs to computationally emulate the “negative control” targets (n = 205). The numbers of sequencing reads corresponding to each sgRNA were counted using the “MAGeCK count” command. The sgRNA ranking and the comparison of sgRNA read counts corresponding to

individual screen targets between the test samples and the reference sample was carried out using the “MAGeCK test” command.

Screen targets were regarded as depleted in a particular cell line if their average $\log_2$(FC) value in both screen replicates of the cell line calculated by the MAGeCK algorithm was ≤-1. Screen targets that were found significantly (FDR < 0.05) depleted in both screen replicates of one cell line were regarded as screen hits in the particular cell line.

4.3 Validation and characterization of the screen hits

4.3.1 Selection of sgRNAs for screen validation

To select individual sgRNAs for validation from the pooled sgRNA library, I calculated a cumulative ranking score of each sgRNA against every screen hit of interest by adding up individual rank values across all screen replicates in all tested cell lines calculated by the MAGeCK algorithm. For every screen hit, top three sgRNAs with the lowest cumulative ranking scores were selected for further validation. sgRNAs targeting EGFP published by Shalem *et al*. [186] were used as negative control sgRNA. sgRNA sequences used for the validation experiments are provided in **Table 11**.

Table 11. *sgRNA sequences used for the validation of the screen results*

Target name	sgRNA name	Forward DNA sequence 5'-3'
Negative control	sgNC_1	GGGCGAGGAGCTGTTCACCG
Negative control	sgNC_2	GAGCTGGACGGCGACGTAAA
Negative control	sgNC_3	GGAGCGCACCATCTTCTTCA
CENPE	sgCENPE_1	TGGCCGTCTGCGTGCGAGTG
CENPE	sgCENPE_2	GCCTGTGAGCCCTGAAGTGC
CENPE	sgCENPE_3	GCCGGCACTTCAGGGCTCAC
CEP55	sgCEP55_1	GAAGCGGCATCCACACCTGA
CEP55	sgCEP55_2	GGGGCACGTCACTGCCGGCA
CEP55	sgCEP55_3	TGGCCCAAGGGAGGCGACCG
EGFR	sgEGFR_1	GCCCGCGCGAGCTAGACGTC
EGFR	sgEGFR_2	GCCCGGACGTCTAGCTCGCG
EGFR	sgEGFR_3	CGGGCGCTCACACCGTGCGG
CASP8AP2	sgCASP8AP2_1 (sgC8AP2_1)	GATGCCAGGGAGACCTCGGT
CASP8AP2	sgCASP8AP2_2 (sgC8AP2_2)	CTGCCCGGCCCAAGACAACC
CASP8AP2	sgCASP8AP2_3 (sgC8AP2_3)	CGGTTCCTTTCTGCCCACCG

4.3.2 Cloning of individual sgRNAs

The DNA sequences corresponding to the forward strands of the selected sgRNAs (**Table 11**) were flanked at the 5'-end with a 5'-CACC(G)-3' overhang compatible with sticky-end cloning using Esp3I restriction digestion of the destination vector LentiCRISPRv2-dCas9-KRAB-puro(iv) [147]. "G" nucleotide was added to the sgRNAs starting with other nucleotides to ensure efficient transcription from the U6 promoter [218]. Reverse strand sgRNA sequences were flanked at the 5'-end with a 5'-AAAC-3' overhang. The resulting sequences were synthesized in the form of desalted single-stranded DNA oligonucleotides by Sigma-Aldrich, resuspended to 100 µM with deionized water and stored at -20°C.

The single-stranded oligonucleotides corresponding to individual sgRNAs were annealed and phosphorylated in a one-step 10 µL reaction using 1 µL of the respective 100 µM forward and reverse oligonucleotide dilutions and 1 µL (10 U) of T4 polynucleotide kinase (T4 PNK, New England Biolabs) in T4 DNA Ligase Buffer (Thermo Scientific) in a thermocycler under the following conditions: 37°C for 30 minutes, 95°C for 5 minutes, followed by cooling to 25°C at a rate of 1°C/second. The completed reaction was diluted with deionized water at 1:500 (v:v) ratio.

LentiCRISPRv2-dCas9-KRAB-puro(iv) vector [147] was linearized with Esp3I restriction enzyme (ThermoFisher Scientific). The linearized vector backbone was gel-extracted using GeneJET Gel Extraction kit (ThermoFisher Scientific). 100 ng of the linearized vector and 1 µL of the 1:500 dilution of the sgRNA oligonucleotide duplexes were used per 20 µL ligation reaction with 1 µL (5 U) of T4 DNA Ligase (ThermoFisher Scientific). The ligation reactions were incubated overnight at 4°C, followed by heat-shock transformation into chemically competent Stabl3 *E. coli* cells.

4.3.3 Bacterial transformation

Chemically competent Stabl3 *E. coli* cells were stored at -80°C degrees and thawed on ice for 30 minutes prior to use. Per one transformation, 50 µL of the cell suspension was added to 5 µL of the ligation reaction, mixed by gently flicking the tube, and incubated for 20 minutes on ice, followed by heat-shock at 42°C for 1 minute. After the heat-shock, the samples were immediately transferred to ice. After 5 minutes on ice, the samples were diluted by 1 mL of pre-warmed LB medium, transferred to round-bottom tubes and left to recover at 37°C, 220

rpm for 1 hour. Next, the bacterial suspensions were pelleted down by centrifugation at 3000 rpm for 4 minutes, and 750 µL of supernatant was removed from each sample. Bacterial pellets were resuspended in the remaining 250 µL of the supernatant, and the resulting suspensions were plated on LB-agar plates containing 100 µg/mL ampicillin (Sigma-Aldrich) pre-warmed to 37°C. The plates were incubated at 37°C for 14 – 16 hours. Afterwards, individual bacterial colonies were picked and inoculated with 2 mL of LB media at 37°C, 220 rpm for 14 – 16 hours.

4.3.4 Plasmid DNA isolation from bacterial cultures

Plasmid DNA was isolated from 2 mL overnight liquid bacterial cultures using NucleoSpin Plasmid Mini Prep kit (Macherey-Nagel) according to the manufacturer's instructions. Isolated constructs were quality-controlled by restriction digestion using FastDigest Eco31I (ThermoFisher Scientific). The correct ligation of the sgRNA-encoding inserts into the CRISPRi constructs was confirmed by Sanger sequencing carried out by GATC Biotech/Eurofins Genomics using the following primer annealing to the U6 promoter: 5'-GAGGGGCCTATTTCCCATGATTCC-3'.

4.3.5 Lentiviral transduction of individual sgRNAs and sgRNA pools

Lentiviral stocks for the CRISPRi constructs expressing individual sgRNAs were produced according to the general protocol described in Section 4.2.3. For transductions with sgRNAs pools, lentiviral stocks corresponding to three individual sgRNAs against the gene of interest were mixed in equal parts, and the resulting stock was used in the subsequent transductions.

The lentiviral transductions were performed according to the general protocol described in the Section 4.2.4 at high MOI on 24-well plates using 200 µL of individual sgRNA lentiviral stocks and 300 µL of pooled lentiviral stocks in a total volume of 400 µL and 500 µL, respectively, with 8 µg/mL polybrene (Th. Geyer).

4.3.6 Proliferation assays and RNA isolation upon lentiviral transduction

At the timepoint of the first cell passaging following the lentiviral transduction, before the start of puromycin selection, the cells were divided between 24-well plates and 6-well

plates for subsequent analysis of cell viability and RNA isolation, respectively. Calu-6, NCI-H3122 and NCI-H1975 cells were split 20 hours post- transduction at a ratio of 1:3; NCI-H838 cells were split 20 hours after the transduction at a ratio of 1:4; NCI-H1650 and PC-9 cells were split 44 hours post-transduction at a ratio of 1:3 and 1:4, respectively. 24 hours after splitting, cells were subjected to antibiotic selection with the appropriate concentrations of puromycin indicated in Section 4.2.4, and further kept in puromycin-containing medium until the end of the experiment. In between cell passaging, puromycin-containing medium was refreshed every 48 hours.

On the 24-well plates, all wells transduced with individual sgRNAs from the same experimental panel were passaged at consistent timepoints between replicates of the same cell line, when cells transduced with negative control sgRNAs were reaching confluence. After one cell passaging round following the addition of puromycin, Calu-6, NCI-H3122 and PC-9 cells, as well as NCI-H1975 cells (during the validation of sgRNAs against *CASP8AP2*) and NCI-H838 cells (during the validation of sgRNAs against *CENPE* and *CEP55*) were grown to confluence of the wells transduced with negative control sgRNAs. Afterwards, 1/3rd (NCI-H3122, NCI-H838) or 1/4th (Calu-6, PC-9, NCI-H1975) of cells from each well were divided between three wells on a transparent 96-well plate for the subsequent cell viability assay.

NCI-H1650 cells, as well as NCI-H1975 cells during the validation of sgRNAs against *CENPE* and *CEP55* and NCI-H838 cells during the validation of sgRNAs against *CASP8AP2,* were grown to confluence of the wells transduced with negative control sgRNAs following the addition of puromycin. Afterwards, 1/3rd (NCI-H1650, NCI-H838) or 1/4th (NCI-H1975) of cells from each well were divided between three wells on a transparent 96-well plate for the subsequent cell viability assay.

Viability of the cells growing on the 96-well plates was assessed using CellTiter-Glo Luminescent Cell Viability assay (CTG, Promega) described in Section 4.3.7 at the following timepoints corresponding to the confluence of wells transduced with negative control sgRNAs: 8 days post-transduction for Calu-6 and PC-9 cells; 11 days post-transduction for NCI-H3122 and NCI-H1650 cells; 8 and 13 days post-transduction for NCI-H1975 cells and 10 and 7 days post-transduction for NCI-H838 cells transduced with sgRNAs against *CENPE*/*CEP55* and *CASP8AP2*, respectively.

On the 6-well plates, cells were subjected to puromycin selection and lysed for subsequent RNA isolation described in Section 4.3.8 at the following timepoints

corresponding to the confluence of wells transduced with negative control sgRNAs: 4 days post-transduction for Calu-6 and NCI-H838 cell lines; 5 days post-transduction for PC-9 and NCI-H1975 cell lines; 7 days post-transduction for NCI-H1650 and NCI-H3122 cell lines.

4.3.7 Cell viability assay

CellTiter-Glo Luminescent Cell Viability assay (CTG, Promega) for the IMR-90, WI-38, Calu-6, NCI-H1975, NCI-H838 and PC-9 samples corresponding to Figure 14 was performed using the CellTiter-Glo reagent diluted in PBS at a ratio of 1:4 (v:v), otherwise according to the manufacturer‘s protocol, with cells initially seeded on white 96-well plates. CellTiter-Glo Cell Viability assay for the rest of the experiments in this study was performed according to the manufacturer‘s protocol with the following modifications. Cells were originally seeded on transparent 96-well plates. CellTiter-Glo reagent was diluted in PBS at a ratio of 1:4 (v:v). At the time of readout, medium was removed from the cells, and 60 µL of the diluted reagent equilibrated to room temperature was added per well. Plates were then placed on an orbital shaker for 2 minutes at 400 rpm protected from light to facilitate cell lysis, and incubated for another 10 minutes to stabilize luminescent signal. 40 µL of resulting cell lysates were transferred into white 96-well plates, and the sample luminescence was recorded using the microplate reader.

4.3.8 RNA isolation, reverse transcription, and RT-qPCR

RNA isolation for all samples except the samples for RNA sequencing was performed using Quick-RNA Microprep kit or Quick-RNA Miniprep kit (Zymo Reseach) consistently within replicates of the same experiment according to the manufacturer’s instructions, including the optional DNase I treatment step, with the following modifications: DNase I treatment was prolonged to 30 minutes, columns were centrifuged after the DNase I treatment at 13000 rpm.

RNA was reverse-transcribed in a two-step RT-PCR using the RevertAid Reverse Transcriptase (ThermoFisher Scientific). 1 µg, 500 ng, 250 ng or 200 ng of of RNA was used per reaction consistently within every sample set. First, RNA was mixed with 2 µL of 10 mM dNTPs, 1 µL of random hexamer primers (ThermoFischer Scientific) and RNAse-free water (UltraPure, Invitrogen) to a total volume of 15 µL, incubated at 65°C for 5 minutes and

immediately transferred to ice. Next, RT buffer, RevertAid Reverse Transcriptase and Ribolock (ThermoFischer Scientific) were added per reaction to a total volume of 20 µL, and the samples were incubated in the thermal cycler under the following conditions: 10 minutes at 25°C, 1 hour at 42°C, 10 minutes at 70°C. The completed RT reactions were diluted with RNase-free water at a ratio of 1:40 (v:v) when starting from 1 µg RNA; 1:20, 1:10 or 1:8 when starting from 500 ng, 250 ng or 200 ng of RNA, respectively, and directly used as templated for RT-qPCR.

RT-qPCR was performed using 4 µL of the diluted RT reaction and Power SYBR Green Master Mix (Applied Biosystems), with 200 nM of forward and reverse primers each in a total volume of 10 µL in in a 96-well format on Applied Biosystems StepOne Lightcycler, except for RT-qPCR reactions corresponding to Supplementary Figures 2-3 that were performed in a 384-well format on Roche Lightcycler 480. *GAPDH* was used as a housekeeping gene for the $2^{-\Delta\Delta Ct}$ normalization. Each reaction was performed in three technical replicates.

Sequences of primers used for RT-qPCR in this study are listed in **Table 13**.

RNA isolation for RNA sequencing is described in Section 4.3.13.

4.3.9 Inhibitor treatments

GSK-923295, CI-1040, Wortmannin, SH-4-54, z-VAD-FMK and T-5224 were purchased from MedChemExpress, diluted in DMSO (AppliChem) and stored according to the manufacturer's instructions.

For the treatments with GSK-923295, CI-1040, Wortmannin and SH-4-54, cells were seeded on transparent 96-well plates 24 hours before at the following densities: $6*10^3$ cells/well for the Calu-6 cell line, $3*10^3$ cells/well for NCI-H838, $4*10^3$ cells/well for PC-9, $4*10^3$ cells/well for NCI-H1975. At the time of treatment, the cell culture medium was replaced by 100 µL of complete cell culture medium containing 0.2% of the required concentration of the compound diluted in DMSO or 0.2% of DMSO as vehicle control for normalization. 72 hours later, viability of the cells was assessed using the CellTiter-Glo Luminescent Cell Viability assay (CTG, Promega) according to the protocol described in Section 4.3.7.

Treatments with z-VAD-FMK and T-5224 are described in Section 4.3.11.

4.3.10 Ectopic *EGFR* overexpression

pHAGE-EGFR (gift from Gordon Mills and Kenneth Scott, [222], Addgene plasmid #116731), pHAGE-EGFR-L858R and pHAGE-EGFR-T790M were a kind gift from Prof. Dr. Daniel A. Haber (Massachusetts General Hospital, Boston, USA). Puromycin resistance gene was replaced by Blasticidin resistance gene using 4-fragment Gibson assembly reactions.

Each of the three pHAGE-EGFR constructs was individually double digested with KpnI-HF (New England Biolabs) and PacI (New England Biolabs) restriction enzymes in a single 50 µL reaction using 1 µg of the respective plasmid in the CutSmart Buffer (New England Biolabs) at 37°C for 15 minutes. The competed reactions were resolved on 1% agarose gel, and the DNA fragments corresponding to the linearized plasmid (approx. 9300 bp) in each sample were extracted from the gel using GeneJET Gel Extraction kit (ThermoFisher Scientific, K0692). The Blasticidin-S deaminase (BSD) gene was amplified from the Lenti-dCas9-KRAB-Blast plasmid (gift from Gary Hon, Addgene plasmid #89567) in a 25 µL PCR reaction using Q5 High-Fidelity DNA Polymerase (New England Biolabs) with the following primers: forward primer 5'-TGAAAAACACGATGATAATATGGCCACACATATGGCCAAGCCTTTGTCTCA-3'; reverse primer 5'-TTTGTAATCCAGAGGTTGATTAGGATCTATTTAGCCCTCCCACACATAACCAG-3'; under the following conditions: initial denaturation at 98°C for 30 seconds, followed by 25 cycles of 98°C for 10 seconds, 68°C for 20 seconds, 72°C for 30 seconds, followed by final elongation at 72°C for 2 minutes.

The DNA fragment of 165 bp flanking the BSD insertion site from the KpnI-end was amplified from the undigested pHAGE-EGFR plasmid with the following primers: forward primer 5'-GGGCTGAAGGATGCCCAGAAG-3', reverse primer 5'-TGTGTGGCCATATTATCATCGTGTTTTTCA-3'; under the following conditions: initial denaturation at 98°C for 30 seconds, followed by 25 cycles of 98°C for 10 seconds, 69°C for 20 seconds, 72°C for 30 seconds, followed by final elongation at 72°C for 2 minutes.

The DNA fragment of approx. 1000 bp flanking the BSD insertion site from the PacI-end was amplified from the undigested pHAGE-EGFR plasmid with the following primers: forward primer 5'-ATAGATCCTAATCAACCTCTGGATTACAAA-3'; reverse primer 5'-TCAAGAATGATCTAGCCCTTTCCTTAAT-3'; under the following conditions: initial denaturation at 98°C for 30 seconds, followed by 25 cycles of 98°C for 10 seconds, 65°C for 20 seconds, 72°C for 30 seconds, followed by final elongation at 72°C for 2 minutes.

The resulting DNA fragments were mixed into three independent 50 µL Gibson assembly reactions with each of the three linearized pHAGE-EGFR backbones, using 50 ng of the respective backbone, 2-fold molar excess of each of the 460 bp and 1000 bp fragments, and 5-fold molar excess of the 165 bp fragment.

The Gibson assembly reactions were incubated at 50°C for 1 hour, followed by heat-shock transformation into chemically competent Stabl3 *E. coli* cells as described in Section 4.3.3 and DNA isolation as described in Section 4.3.4.

To create an empty vector, EGFR was excised from pHAGE-EGFR-Blast by restriction digestion with BsrGI-HF (New England Biolabs). The completed reaction was resolved on 1% agarose gel, and the DNA fragment corresponding to the linearized plasmid (approx. 8200 bp) was extracted from the gel using GeneJET Gel Extraction kit (ThermoFisher Scientific). The extracted fragment was re-circularized using T4 DNA Ligase (ThermoFisher Scientific) according to the manufacturer's instructions.

Lentiviral stocks for the resulting pHAGE-EGFR-Blast, pHAGE-EGFR-L858R-Blast, pHAGE-EGFR-T790M-Blast and pHAGE-empty-Blast were produced according to the protocol described in Section 4.2.3. Calu-6 cells were transduced on 24-well plates at high MOI using 200 µL of the respective lentiviral stocks with 8 µg/mL polybrene (Th. Geyer) according to the protocol described in Section 4.2.4.1. Stably transduced Calu-6 were selected using 10 µg/mL Blasticidin S hydrochloride (Sigma-Aldrich) for 7 days. Afterwards, the cells were used for secondary transductions with the respective CRISPRi constructs according to the protocol described in Section 4.3.5. To discriminate between endogenous and ectopic *EGFR* mRNA expression in the transduced cells, the following primer pair, with the forward primer annealing to the 5'-UTR of the ectopic *EGFR* mRNA, was used for RT-qPCR: forward primer 5'-ACAAGTTTGTACAAAAAAGCAGGCA-3'; reverse primer 5'-TTTTCCTCCAGAGCCCGACT-3'.

4.3.11 Transfection with siPOOLs

Cells were reverse-transfected with respective siPOOLs (designed and manufactured by siTOOLs Biotech) using Lipofectamine RNAiMax (Invitrogen) as a transfection reagent at 10 nM final concentration for single siPOOL transfections and 10 nM of each siPOOLs for double siPOOL transfections to 20 nM final concentration. Sequences of all siPOOLs used in this study are provided in **Table 14**.

For transfections on a 96-well plate format, 10 µL of of Opti-MEM I Reduced-Serum Medium (Gibco) containing siPOOLs was pipetted to the wells, followed by 10 µL of Opti-MEM containing 0.2 µL Lipofectamine RNAiMax, mixed by pipetting, and incubated for 20 minutes at room temperature. Afterwards, 80 µL of cell suspension containing an appropriate number of cells indicated in **Table 12** in the complete cell culture medium was added to the experimental wells.

For transfections on a 6-well plate format, 100 µL of Opti-MEM with siPOOLs was mixed with 100 µL of Opti-MEM containing 4 µL Lipofectamine RNAiMax and 1.8 mL of the cell suspension per transfection. For transfections on 10 cm cell culture plates, 1 mL of Opti-MEM with siPOOLs was mixed with 1 mL of Opti-MEM containing 20 µL Lipofectamine RNAiMax and 8 mL of the cell suspension per transfection.

For cell viability assays, cells were transfected on 96-well plates in technical triplicates. At 72 hours post-transfection, CellTiter-Glo Luminescent Cell Viability assay was performed according to the protocol described in Section 4.3.7. For caspase activity assays, cells were transfected on white 96-well plates in technical duplicates, and further analyzed as described in Section 4.3.12. For conventional RNA isolation, cells were transfected on 6-well plates and lysed for RNA isolation at 48 hours post-transfection with RNA Lysis buffer (Zymo Research) according to the manufacturers' instructions. The lysates were stored at -20°C until further use. The samples were further processed according to the protocol descried in 4.3.8. For RNA isolation for RNA-Seq and FACS analysis, cells were transfected on 10 cm cell culture plates, and processed as described in Sections 4.3.13 and 4.3.16, respectively.

For treatments with z-VAD-FMK and T-5224 (MedChemExpress), appropriate amounts of inhibitors diluted in DMSO to a volume corresponding to 1% and 0.5% of the final transfection volume, respectively, were added to the cell suspensions and incubated for 90 minutes prior to adding the cells to the wells containing the siPOOLs-RNAiMax complexes. Cells treated with 1% and 0.5% of DMSO, respectively, and transfected with negative control siPOOLs were used as normalization controls.

Table 12. *Cell line seeding densities for reverse-transfection with siPOOLs*

Cell line \ Format	96-well plate; cells/well	6-well plate; cells/well	10 cm plate; cells/plate
IMR-90	$1*10^3$	$3*10^4$	$1.8*10^5$
WI-38	$0.85*10^3$	$2.55*10^4$	$1.53*10^5$
Calu-6	$6*10^3$	$1.8*10^5$	$1.08*10^6$
PC-9	$4*10^3$	$1.2*10^5$	$7.2*10^5$
NCI-H1975	$4*10^3$	$1.2*10^5$	$7.2*10^5$
NCI-H838	$3*10^3$	$9*10^4$	$5.4*10^5$
NCI-H460	$4*10^3$	$1.2*10^5$	$7.2*10^5$
NCI-H1299	$3*10^3$	$9*10^4$	$5.4*10^5$
NCI-H1563	$2*10^3$	$6*10^4$	$3.6*10^5$

4.3.12 Caspase activity assays

Cells were reverse-transfected with the respective siPOOLs at 10 nM final concentration on white 96-well plates in technical duplicates. Two wells were filled with the respective medium mixed with same volume of Opti-MEM used in experimental wells for adding the transfection reagents for obtaining background reads.

At 48 hours post-transfection, Caspase-Glo assays (Caspase-Glo 3/7, Caspase-Glo 8 and Caspase-Glo 9, Promega) were performed according to the manufacturer‘s instructions. Briefly, Caspase-Glo reagents were mixed from the provided substrates and buffers (supplied with MG-132 for Caspase-Glo 8 and Caspase-Glo 9 as recommended), aliquoted into opaque tubes and stored at -80°C. The reagents were equilibrated to room temperature for 1.5 hours before the assay. Each aliquot has only been used once to avoid signal loss due to repeated freezing-thawing.

Plates were equilibrated to room temperature for 30 minutes before addition of an equal volume (100 µL/well) of Caspase-Glo reagents. Plates were then placed on an orbital shaker for two minutes at 400 rpm protected from light to facilitate cell lysis, and incubated

for one hour at room temperature to stabilize the luminescent signal before readout. Respective background reads were subtracted from experimental reads.

4.3.13 RNA sequencing

Cells were transfected with siPOOLs on 10 cm plates 48 hours prior to RNA isolation according to the protocol described in Section 4.3.11 in three biological replicates. RNA isolation was performed using RNeasy Mini kit (QIAGEN) according to the manufacturer's instructions, including the optional DNase I (QIAGEN) treatment step. The isolated RNA samples were subsequently treated with Illumina Ribo-Zero Plus rRNA Depletion kit (Illumina) by Jeanette Seiler according to the manufacturer's instructions. Stranded RNA-Seq libraries were prepared and sequenced by the DKFZ Genomics and Proteomics Core facility on Illumina HiSeq 4000 platform to obtain paired-end 100 bp reads.

4.3.14 Differential gene expression analysis

RNA-Seq data analysis was performed on the European UseGalaxy server (https://usegalaxy.eu) [209]. The FastQC tool was used to assess the quality of the reads [214]. Strandness of the RNA-Seq reads was assessed using the Infer Experiment tool. RNA-Seq reads were aligned to the reference human hg38/GRCh38 genome using the HISAT2 tool [215]. The aligned reads were counted using the featureCounts tool [216]. Based on the featureCounts outputs, genes differentially expressed between cells transfected with siPOOLs against *CASP8AP2* and the negative control siPOOLs across three technical replicates were identified using the DESeq2 tool [217]. Further analysis was carries out according to the selection criteria outlined in Section 5.3.5.

4.3.15 Lung cell line transcriptome and proteome analysis

The lung cell line RNA and protein expression datasets from the Diederichs lab were previously published by Pal *et al*. [205]. The dataset analysis criteria used in this study are outlined in Section 5.3.5.

4.3.16 Flow cytometry

4.3.16.1 BrdU incorporation assay and cell fixation

At 46 hours post-transfection, BrdU was added to the cell culture medium to the final concentration of 10 µM. After two hours (48 hours post-transfection), medium was removed, cells were washed with PBS, trypsinized, collected in 15 mL tubes and pelleted by centrifugation at 600x*g* for 5 minutes at room temperature. Cell pellets were washed in PBS twice, resuspended in 1 mL PBS and fixed by addition of 2.5 mL of ice-cold 100% ethanol under vortexing followed by an overnight incubation at 4°C. For further storage, cells were transferred to -20°C for up to 2 weeks.

4.3.16.2 anti-BrdU/PI-staining and flow cytometry analysis

Fixed cells were pelleted as described above. The supernatant was removed, and the pellets were washed in PBS supplemented with 0.5% BSA (w/v), followed by another round of centrifugation. The supernatant was removed, and the cell pellets were resuspended and incubated in 2M HCl for 20 minutes to facilitate DNA hydrolysis. Afterwards, cells were pelleted, washed in PBS supplemented with 0.5% BSA (w:v), pelleted again, and the supernatant was removed. The pellets were resuspended in 0.1 M sodium tetraborate, pH 8.5. and incubated for 2 minutes incubation Then, cells were washed again, pelleted, and the supernatant was removed. The cell pellets were incubated with FITC-labeled anti-BrdU antibody (Biolegend) for 1 hour at room temperature protected from light. Afterwards, samples were washed and pelleted and as described above. The supernatants were removed, and the cells were resuspended in 50 ng/mL propidium iodide solution supplemented with 100 µg/mL RNase A and incubated with for 10 minutes at 37°C protected from light. Samples were transferred to ice, protected from light, and analyzed by flow cytometry. Cells were filtered through a 100 µm cell strainer immediately before analysis. For the general cell cycle phase analysis and BrdU incorporation analysis, acquired signals were gated for single cells using FSC-A/SSC-A and FSC-W/FSC-H-based exclusion criteria. For the subG1-analysis, the ungated total cell population was used. The analysis was carried out and the data were visualized using the FlowJo v10 software.

4.4 Statistical analysis and data visualization

Significance testing was performed using two-tailed Student's t-test in Microsoft Excel (TTEST function). For fold change data, testing was performed after $\log_2$-transformation. For datasets normalized to a single negative control measurement per biological replicate, two-tailed one-sample (paired) t-tests were used. For datasets normalized to mean of several negative control measurements per biological replicate, two-tailed unpaired t-tests were used. In the latter case the need for Welch's correction for unequal variances was determined using an F-test (FTEST function). Alpha cutoff for significance was set to 0.05. At p-value < 0.05, the null hypothesis was rejected, and the difference between tested datasets was considered significant.

The combinatorial effects of the simultaneous silencing of two genes as well as the combinatorial effects of target gene silencing and another gene overexpression or inhibitor treatment on the cell viability were quantified using the Bliss independence model C = (A + B) – (A x B), where A and B stood for individual effects on the cell viability measured independently, and C stood for the predicted combinatorial effect. If the difference between the predicted and the observed combinatorial effects $\Delta C^{\text{observed-predicted}} = 0$, the combinatorial effect was classified as additive; if $\Delta C^{\text{observed-predicted}} > 0$, the combinatorial effect was classified as synergistic; if $\Delta C^{\text{observed-predicted}} < 0$, the combinatorial effect was classified as antagonistic.

Unless stated otherwise, data were visualized using the GraphPad Prism software. The final figures were prepared using Microsoft PowerPoint.

Table 13. *RT-qPCR primers used in this study (in alphabetical order)*

Gene name	Forward primer 5'-3'	Reverse primer 5'-3'
CASP3	TGGAATTGATGCGTGATGTT	TCCAAAAATTATTCCTTCTTCACC
CASP7	GCACTTGGCAAAAAGAATCC	CCACATTCTGCAATAAAGAGCA
CASP8	TCCAAATGCAAACTGGATGA	TCTCCCAGGATGACCCTCTT
CASP8AP2	ATCATGGCAGCAGATGATGA	CCCAGCGTATATGTCCAGTG
CASP9	GTTTGAGGACCTTCGACCAGCT	CAACGTACCAGGAGCCACTCTT
CDKN1A	AGGTGGACCTGGAGACTCTCAG	TCCTCTTGGAGAAGATCAGCCG
CENPE	GAGTTATACCCAGGGCAATTCA	CTTCTGTGAGATCAGCAACATACA
CEP55	CCAGCATGCTAGTGAATCATGT	GGGAGGTATCACTGCCAAGA
EGFR	GGCTCTGGAGGAAAAGAAAGTT	ACCACCTCACAGTTATTGAACATC
GAPDH	GTGAAGGTCGGAGTCAACG	TGAGGTCAATGAAGGGGTC
JUNB	CGATCTGCACAAGATGAACCACG	CTGCTGAGGTTGGTGTAAACGG
TCF3	GTTTCCAGGCCTGAGGTG	CTTCCCGTTGGTGACAGG
TP53	GTGACACGCTTCCCTGGATT	TGTTTCCTGACTCAGAGGGG

Table 14. *Sequences of siPOOLs used in this study (in order of use)*

Target name	siPOOLs name	siRNA #	Sense sequence 5'-3'	Antisense sequence 5'-3'
Negative control	siNC	1	TGTACGCGTCTCGCGATTT	AAATCGCGAGACGCGTACA
		2	TATACGCGGTACGATCGTT	AACGATCGTACCGCGTATA
		3	TTCGCGTAATAGCGATCGT	ACGATCGCTATTACGCGAA
		4	TCGGCGTAGTTTCGACGAT	ATCGTCGAAACTACGCCGA
		5	TCGCGTAAGGTTCGCGTAT	ATACGCGAACCTTACGCGA
		6	TCGCGATTTTAGCGCGTAT	ATACGCGCTAAAATCGCGA
		7	TCGCGTATATACGCTACGT	ACGTAGCGTATATACGCGA
		8	TTTCGCGAACGCGCGTAAT	ATTACGCGCGTTCGCGAAA
		9	TCGTATCGTATCGTACCGT	ACGGTACGATACGATACGA
		10	TTATCGCGCGTTATCGCGT	ACGCGATAACGCGCGATAA
		11	TCTCGTAGGTACGCGATCT	AGATCGCGTACCTACGAGA
		12	TCGTACTCGATAGCGCAAT	ATTGCGCTATCGAGTACGA
		13	TTTGCGATACCGTAACGCT	AGCGTTACGGTATCGCAAA
		14	TGCGTAAGGCATGTCGTAT	ATACGACATGCCTTACGCA
		15	TTATCGGCAGTTCGCCGTT	AACGGCGAACTGCCGATAA
		16	TAGCGCGACATCTATCGCT	AGCGATAGATGTCGCGCTA
		17	TCGTCGTATCAGCGCGTTT	AAACGCGCTGATACGACGA
		18	TACGCGAAACTGCGTTCGT	ACGAACGCAGTTTCGCGTA
		19	TCGACGATAGCTATCGCGT	ACGCGATAGCTATCGTCGA
		20	TCGCGTAATACGCGATCGT	ACGATCGCGTATTACGCGA
		21	TCGCGATAATGTTACGCGT	ACGCGTAACATTATCGCGA
		22	TTAACGCGCTACGCGTATT	AATACGCGTAGCGCGTTAA
		23	TCGCGTATAGGTAACGCGT	ACGCGTTACCTATACGCGA
		24	TTACGCGATCACGTAACGT	ACGTTACGTGATCGCGTAA
		25	TTATCGCGCGTCGCGTAAT	ATTACGCGACGCGCGATAA
		26	TTACGTACTAGTGCGTACT	AGTACGCACTAGTACGTAA
		27	TATACGCCGGTTGCGTAGT	ACTACGCAACCGGCGTATA
		28	TTCGCGTGCATAGCGTAAT	ATTACGCTATGCACGCGAA
		29	TACGCGACCTAATCGCGAT	ATCGCGATTAGGTCGCGTA
		30	TCGTACGCTGAACGCGTAT	ATACGCGTTCAGCGTACGA
CASP8AP2	siC8AP2	1	GAAACAGACTTAACTAATA	TATTAGTTAAGTCTGTTTC
		2	GCAAGGTTCTGATCTTTTA	TAAAAGATCAGAACCTTGC
		3	GTGACTACCTTACAGAAGA	TCTTCTGTAAGGTAGTCAC
		4	GTGATGTATTGAAATCTTA	TAAGATTTCAATACATCAC
		5	GTGATGAAGTGGCTGATGA	TCATCAGCCACTTCATCAC
		6	GTTCGAAGTTTAAAAGTAA	TTACTTTTAAACTTCGAAC
		7	GTGATATTATAGAGTCTAA	TTAGACTCTATAATATCAC
		8	GTTTTGGTATATTGTAAAA	TTTTACAATATACCAAAAC
		9	GAGAATTGCTGAAAGAAAA	TTTTCTTTCAGCAATTCTC
		10	GACATTTGATACAGTTAAA	TTTAACTGTATCAAATGTC
		11	GCTTGGAATTGGATACCAA	TTGGTATCCAATTCCAAGC
		12	GTGGGAAAATACACCTTTA	TAAAGGTGTATTTTCCCAC
		13	GAAAATAGTTTGTTAGTTA	TAACTAACAAACTATTTTC
		14	GGCTTTCAGTACCACCCTA	TAGGGTGGTACTGAAAGCC
		15	GTCCTAAAATGATTAGTGA	TCACTAATCATTTTAGGAC
		16	GCTGATATTTTGCCAATAA	TTATTGGCAAAATATCAGC

CASP8AP2	siC8AP2	17	GAAATTGTTTGGACTTATA	TATAAGTCCAAACAATTTC
		18	GCAGAAGACTAATAAGGAA	TTCCTTATTAGTCTTCTGC
		19	GACCTAAAATTGAGTTTTA	TAAAACTCAATTTTAGGTC
		20	GGCATATGTTCCCTCCATA	TATGGAGGGAACATATGCC
		21	GGTCATTATTATCATCTGA	TCAGATGATAATAATGACC
		22	GCCCAGACATGGAAAGGAA	TTCCTTTCCATGTCTGGGC
		23	GGACCAAGATCACTGGAGA	TCTCCAGTGATCTTGGTCC
		24	GAAATGACTTTCTAGATTA	TAATCTAGAAAGTCATTTC
		25	GGAAAATTATCAAGACCAA	TTGGTCTTGATAATTTTCC
		26	GGCTTTGCCTAGCCAAGGA	TCCTTGGCTAGGCAAAGCC
		27	GGAAAGAAGTTAAATCACA	TGTGATTTAACTTCTTTCC
		28	GAGTGGAAATAAACCGCAA	TTGCGGTTTATTTCCACTC
		29	GCCCTGCGTTTCTTCCATA	TATGGAAGAAACGCAGGGC
		30	GCTTCTCCATGTCCTAAAA	TTTTAGGACATGGAGAAGC
CASP3 + CASP7	siCASP3/7	1	GGATCGTTGTAGAAGTCTA	TAGACTTCTACAACGATCC
		2	GACGCTACTTTTCATGCAA	TTGCATGAAAAGTAGCGTC
		3	GAGATTTCTTGTTGCTCAA	TTGAGCAACAAGAAATCTC
		4	GAGGGTACTTTAAGACATA	TATGTCTTAAAGTACCCTC
		5	GGTGGAGTTTTAACTGTAA	TTACAGTTAAAACTCCACC
		6	GCTTCTGAGCCATGGTGAA	TTCACCATGGCTCAGAAGC
		7	GGAAAACCCAAACTTTTCA	TGAAAAGTTTGGGTTTTCC
		8	GAGATGGGTTTATGTATAA	TTATACATAAACCCATCTC
		9	GTGCCATGCTGAAACAGTA	TACTGTTTCAGCATGGCAC
		10	GCAGCAAACCTCAGGGAAA	TTTCCCTGAGGTTTGCTGC
		11	GAGAGGCAATGATTGTTAA	TTAACAATCATTGCCTCTC
		12	CTGAAACAGTATGCCGACA	TGTCGGCATACTGTTTCAG
		13	GGCTAAAACTTAACATTCA	TGAATGTTAAGTTTTAGCC
		14	GTAGAAGTCTAACTGGAAA	TTTCCAGTTAGACTTCTAC
		15	GGACCTGTTGACCTGAAAA	TTTTCAGGTCAACAGGTCC
		16	GGCAGATTTTCATGCAAAA	TTTTGCATGAAAATCTGCC
		17	GATGCAAGATCTGCTTAAA	TTTAAGCAGATCTTGCATC
		18	GAGGAAGAGTTTATGGCAA	TTGCCATAAACTCTTCCTC
		19	CTGACTTCCTCTTCGCCTA	TAGGCGAAGAGGAAGTCAG
		20	GTGGTTGCTTAATTCCTAA	TTAGGAATTAAGCAACCAC
		21	GACCCACACTTCCATGAGA	TCTCATGGAAGTGTGGGTC
		22	GGGCGTTCGAAACGGAACA	TGTTCCGTTTCGAACGCCC
		23	GCATCATAATAAACAACAA	TTGTTGTTTATTATGATGC
		24	GGCTATTACTCGTGGAGGA	TCCTCCACGAGTAATAGCC
		25	GCACGGAAAAGACCTGGAA	TTCCAGGTCTTTTCCGTGC
		26	GCCGAGGGACCGAGCTTGA	TCAAGCTCGGTCCCTCGGC
		27	GAACATAAGTAATCACTAA	TTAGTGATTACTTATGTTC
		28	GTGCCTACATATCAGTACA	TGTACTGATATGTAGGCAC
		29	GGGCAAATGCATCATAATA	TATTATGATGCATTTGCCC
		30	GGTACATTCTAGCTGAGAA	TTCTCAGCTAGAATGTACC
CASP8	siCASP8	1	GCCTCCCTCAAGTTCCTGA	TCAGGAACTTGAGGGAGGC
		2	CACGGGAGAAAGTGCCCAA	TTGGGCACTTTCTCCCGTG
		3	GGTCTTTTAAGTTTCTTTT	AAAAGAAACTTAAAAGACC
		4	GGACTTCAGCAGAAATCTT	AAGATTTCTGCTGAAGTCC
		5	GGGCTCAAATTTCTGCCTA	TAGGCAGAAATTTGAGCCC
		6	CCATCAAGGATGCCTTGAT	ATCAAGGCATCCTTGATGG

CASP8	siCASP8	7	CAGAATTGAGGTCTTTTAA	TTAAAAGACCTCAATTCTG
		8	GAAAGTTGGACATCCTGAA	TTCAGGATGTCCAACTTTC
		9	GAGTCTGTGCCCAAATCAA	TTGATTTGGGCACAGACTC
		10	CCCAAATCAACAAGAGCCT	AGGCTCTTGTTGATTTGGG
		11	GCCTTGATGTTATTCCAGA	TCTGGAATAACATCAAGGC
		12	CCGAATTAATAGACTGGAT	ATCCAGTCTATTAATTCGG
		13	CCTACCTAAACACTAGAAA	TTTCTAGTGTTTAGGTAGG
		14	CTATCAGATTTCAGAAGAA	TTCTTCTGAAATCTGATAG
		15	GTCATGCTCTATCAGATTT	AAATCTGATAGAGCATGAC
		16	GACATGAACCTGCTGGATA	TATCCAGCAGGTTCATGTC
		17	GATAATCAACGACTATGAA	TTCATAGTCGTTGATTATC
		18	CTCCAGGAAAAGAGAATGT	ACATTCTCTTTTCCTGGAG
		19	GGAAAGGGAACTTCAGACA	TGTCTGAAGTTCCCTTTCC
		20	GCAAGAACCCATCAAGGAT	ATCCTTGATGGGTTCTTGC
		21	GTCTGATCATCAACAATCA	TGATTGTTGATGATCAGAC
		22	GAGCTGCTCTTCCGAATTA	TAATTCGGAAGAGCAGCTC
		23	GACATCCTGAAAAGAGTCT	AGACTCTTTTCAGGATGTC
		24	CTGGATATTTTCATAGAGA	TCTCTATGAAAATATCCAG
		25	GGAGAAGGAAAGTTGGACA	TGTCCAACTTTCCTTCTCC
		26	CAGCATTAGGGACAGGAAT	ATTCCTGTCCCTAATGCTG
		27	GCAGAAATCTTTATGATAT	ATATCATAAAGATTTCTGC
		28	GGAACAACTGGACAGTGAA	TTCACTGTCCAGTTGTTCC
		29	GCCTGGACTACATTCCGCA	TGCGGAATGTAGTCCAGGC
		30	CTGATTACCTACCTAAACA	TGTTTAGGTAGGTAATCAG
CASP9	siCASP9	1	CCAAGTAGCTCTTACTACA	TGTAGTAAGAGCTACTTGG
		2	CTAACAGGCAAGCAGCAAA	TTTGCTGCTTGCCTGTTAG
		3	GGGTTTGAGGTGGCCTCCA	TGGAGGCCACCTCAAACCC
		4	CAGAGATTCGCAAACCAGA	TCTGGTTTGCGAATCTCTG
		5	GGTTAGGTCTCTTGTCCAA	TTGGACAAGAGACCTAACC
		6	CTTTCTTCTTCCTTACAAA	TTTGTAAGGAAGAAGAAAG
		7	CCAACCCTAGAAAACCTTA	TAAGGTTTTCTAGGGTTGG
		8	CTAATGCTGTTTCGGTGAA	TTCACCGAAACAGCATTAG
		9	CAGCTGGACGCCATATCTA	TAGATATGGCGTCCAGCTG
		10	CCTGAGTGGTGCCAAACAA	TTGTTTGGCACCACTCAGG
		11	GCTTTAATTTCCTCCGGAA	TTCCGGAGGAAATTAAAGC
		12	CACCCAGACCAGTGGACAT	ATGTCCACTGGTCTGGGTG
		13	GGACATTGGTTCTGGAGGA	TCCTCCAGAACCAATGTCC
		14	CAGTGGTGCTCAGACCAGA	TCTGGTCTGAGCACCACTG
		15	GTCGAAGCCAACCCTAGAA	TTCTAGGGTTGGCTTCGAC
		16	GAGGTTCTCAGACCGGAAA	TTTCCGGTCTGAGAACCTC
		17	CCACAAGGCTCCACCCTGA	TCAGGGTGGAGCCTTGTGG
		18	GCAGTGGGCTCACTCTGAA	TTCAGAGTGAGCCCACTGC
		19	CTGGCTTCGTTTCTGCGAA	TTCGCAGAAACGAAGCCAG
		20	GTGTGAACGTGCTGTGCAA	TTGCACAGCACGTTCACAC
		21	CCTGCAGTCCCTCCTGCTT	AAGCAGGAGGGACTGCAGG
		22	GCTGAACAGTGGAGGAAGA	TCTTCCTCCACTGTTCAGC
		23	CGCAAACCAGAGGTTCTCA	TGAGAACCTCTGGTTTGCG
		24	CCGTTCCAGGAAGGTTTGA	TCAAACCTTCCTGGAACGG
		25	CCCATTTATTCATGTCTTA	TAAGACATGAATAAATGGG
		26	GGACCTTCGACCAGCTGGA	TCCAGCTGGTCGAAGGTCC

CASP9	siCASP9	27	GCAGAAAGACCATGGGTTT	AAACCCATGGTCTTTCTGC
		28	CCCACACCCAGTGACATCT	AGATGTCACTGGGTGTGGG
		29	GTGGCTCCTGGTACGTTGA	TCAACGTACCAGGAGCCAC
		30	CCCTGGACGACATCTTTGA	TCAAAGATGTCGTCCAGGG
CDKN1A	siCDKN1A	1	GGACACTCAGACCTGAATT	AATTCAGGTCTGAGTGTCC
		2	GCAGGGACCACACCCTGTA	TACAGGGTGTGGTCCCTGC
		3	GCACCCTAGTTCTACCTCA	TGAGGTAGAACTAGGGTGC
		4	CCAGCTCAATGGACTGGAA	TTCCAGTCCATTGAGCTGG
		5	GAGACTCTCAGGGTCGAAA	TTTCGACCCTGAGAGTCTC
		6	GCTTAGTGTACTTGGAGTA	TACTCCAAGTACACTAAGC
		7	CTCAGTTTGTGTGTCTTAA	TTAAGACACACAAACTGAG
		8	GCAGTAGAGGCTATGGACA	TGTCCATAGCCTCTACTGC
		9	CGCTAATGGCGGGCTGCAT	ATGCAGCCCGCCATTAGCG
		10	CCTAAGAGTGCTGGGCATT	AATGCCCAGCACTCTTAGG
		11	CTCCCACAATGCTGAATAT	ATATTCAGCATTGTGGGAG
		12	GCCGGCTGATCTTCTCCAA	TTGGAGAAGATCAGCCGGC
		13	GGAGGCACTGAAGTGCTTA	TAAGCACTTCAGTGCCTCC
		14	CCTCTGGCATTAGAATTAT	ATAATTCTAATGCCAGAGG
		15	CGCGACTGTGATGCGCTAA	TTAGCGCATCACAGTCGCG
		16	GTTCATTGCACTTTGATTA	TAATCAAAGTGCAATGAAC
		17	GTCACAGGCGGTTATGAAA	TTTCATAACCGCCTGTGAC
		18	CGGAACAAGGAGTCAGACA	TGTCTGACTCCTTGTTCCG
		19	GCCTCAAAGGCCCGCTCTA	TAGAGCGGGCCTTTGAGGC
		20	GGCGGTTGAATGAGAGGTT	AACCTCTCATTCAACCGCC
		21	CCTAATCCGCCCACAGGAA	TTCCTGTGGGCGGATTAGG
		22	CTCCAAGAGGAAGCCCTAA	TTAGGGCTTCCTCTTGGAG
		23	GCGATGGAACTTCGACTTT	AAAGTCGAAGTTCCATCGC
		24	GGGTGTGGCTCCTTCCCAT	ATGGGAAGGAGCCACACCC
		25	GGGAGCCCGTCTCAGTGTT	AACACTGAGACGGGCTCCC
		26	GGTCCCATGTGGTGGCACA	TGTGCCACCACATGGGACC
		27	GGCTTCATGCCAGCTACTT	AAGTAGCTGGCATGAAGCC
		28	GCTGGGAGTAGTTGTCTTT	AAAGACAACTACTCCCAGC
		29	GCGGCAGACCAGCATGACA	TGTCATGCTGGTCTGCCGC
		30	GGGAAGGGACACACAAGAA	TTCTTGTGTGTCCCTTCCC
TP53	siTP53	1	GCAGTTAAGGGTTAGTTTA	TAAACTAACCCTTAACTGC
		2	GGCATTTGCACCTACCTCA	TGAGGTAGGTGCAAATGCC
		3	GTGAACCTTAGTACCTAAA	TTTAGGTACTAAGGTTCAC
		4	CCCTGTCTGACAACCTCTT	AAGAGGTTGTCAGACAGGG
		5	CCCAACAACACCAGCTCCT	AGGAGCTGGTGTTGTTGGG
		6	CTAACTTCAAGGCCCATAT	ATATGGGCCTTGAAGTTAG
		7	GGCCATCTACAAGCAGTCA	TGACTGCTTGTAGATGGCC
		8	CCTCACTGTTGAATTTTCT	AGAAAATTCAACAGTGAGG
		9	CCCAGGGAGCACTAAGCGA	TCGCTTAGTGCTCCCTGGG
		10	GGATTTCATCTCTTGTATA	TATACAAGAGATGAAATCC
		11	CCCACACCCTGGAGGATTT	AAATCCTCCAGGGTGTGGG
		12	GGAAGACTCCAGTGGTAAT	ATTACCACTGGAGTCTTCC
		13	GCAATAGGTGTGCGTCAGA	TCTGACGCACACCTATTGC
		14	CTGGCCTTGAAACCACCTT	AAGGTGGTTTCAAGGCCAG
		15	GGGCAGCTGGTTAGGTAGA	TCTACCTAACCAGCTGCCC
		16	CAGTCTACCTCCCGCCATA	TATGGCGGGAGGTAGACTG

TP53	siTP53	17	GCACCCAGGACTTCCATTT	AAATGGAAGTCCTGGGTGC
		18	GCGCTGCTCAGATAGCGAT	ATCGCTATCTGAGCAGCGC
		19	CTGTGAAATGCTGGCATTT	AAATGCCAGCATTTCACAG
		20	GTCGGTGGGTTGGTAGTTT	AAACTACCAACCCACCGAC
		21	GGCCTGACTCAGACTGACA	TGTCAGTCTGAGTCAGGCC
		22	GGCTCTGACTGTACCACCA	TGGTGGTACAGTCAGAGCC
		23	GCATCTTATCCGAGTGGAA	TTCCACTCGGATAAGATGC
		24	GCTGAATGAGGCCTTGGAA	TTCCAAGGCCTCATTCAGC
		25	CACATGACGGAGGTTGTGA	TCACAACCTCCGTCATGTG
		26	GGGCTCCACTGAACAAGTT	AACTTGTTCAGTGGAGCCC
		27	GGCGCACAGAGGAAGAGAA	TTCTCTTCCTCTGTGCGCC
		28	GGATGTTTGGGAGATGTAA	TTACATCTCCCAAACATCC
		29	CCACTGGATGGAGAATATT	AATATTCTCCATCCAGTGG
		30	GGACATACCAGCTTAGATT	AATCTAAGCTGGTATGTCC
TCF3	siTCF3	1	GCCTATGCCTCCTTCGGGA	TCCCGAAGGAGGCATAGGC
		2	GGACGAGGAGAACACGTCA	TGACGTGTTCTCCTCGTCC
		3	CGGCCGCCAGCGAGATCAA	TTGATCTCGCTGGCGGCCG
		4	GCCTGGCAGGAACGTCACA	TGTGACGTTCCTGCCAGGC
		5	GCGCTGGCCTCAGGTTTCA	TGAAACCTGAGGCCAGCGC
		6	GCCTAGACACGCAGCCCAA	TTGGGCTGCGTGTCTAGGC
		7	CGGCTGACCACTCGGAGGA	TCCTCCGAGTGGTCAGCCG
		8	CAATAAACGTGACATTTTA	TAAAATGTCACGTTTATTG
		9	CCGGATCACTCAAGCAATA	TATTGCTTGAGTGATCCGG
		10	GCCTCAAGATGGTTTTCAA	TTGAAAACCATCTTGAGGC
		11	GGCGAGCTGGCCCTCAACA	TGTTGAGGGCCAGCTCGCC
		12	GCGCCTGTGGGCACAGACA	TGTCTGTGCCCACAGGCGC
		13	GGTGAACGGTGGGCTCCCA	TGGGAGCCCACCGTTCACC
		14	CCGGTGGGCAGCAGTGGAA	TTCCACTGCTGCCCACCGG
		15	GGTGTACCCACCCAGCTCA	TGAGCTGGGTGGGTACACC
		16	GCACAGCCGGCGACATGCA	TGCATGTCGCCGGCTGTGC
		17	GTGTGATCTGAGTGCCTCA	TGAGGCACTCAGATCACAC
		18	CTGGATGATTGGGACTTTA	TAAAGTCCCAATCATCCAG
		19	GAACTGTGCGTTCTGCATA	TATGCAGAACGCACAGTTC
		20	CTCCTGGACTTCAGCATGA	TCATGCTGAAGTCCAGGAG
		21	CCAGCTCAGGTGAGGACTA	TAGTCCTCACCTGAGCTGG
		22	GGCAAAGCACTGGCCTCGA	TCGAGGCCAGTGCTTTGCC
		23	GGCGCAGTTCGGAGGTTCA	TGAACCTCCGAACTGCGCC
		24	GCATAGAATTCAAACGAGA	TCTCGTTTGAATTCTATGC
		25	GCAGGGATGCCACCGCCTA	TAGGCGGTGGCATCCCTGC
		26	CGGCCTCCCGACTCCTACA	TGTAGGAGTCGGGAGGCCG
		27	GCGCGAGGAGGAAGAAACA	TGTTTCTTCCTCCTCGCGC
		28	GCAGTTCCTGGGAACCACA	TGTGGTTCCCAGGAACTGC
		29	CCCGGTGCCTTATCGCCCA	TGGGCGATAAGGCACCGGG
		30	GCACCAGCCTCATGCACAA	TTGTGCATGAGGCTGGTGC

[This page intentionally left blank]

5 Results

Parts of the text and figures presented in this chapter are revised versions of the text and figures submitted for a patent application to the European Patent Office under Application No. / Patent No. 20200810.8-1111.

5.1 Dropout CRISPRi-based cell viability screen for novel oncogenic targets in lung adenocarcinoma cell lines

The primary goal of my PhD thesis project was to design and perform a CRISPRi-based negative selection cell viability screen to identify novel oncogenic factors in lung adenocarcinoma (LUAD). Therefore, I first set out to select the targets for prospective screening.

5.1.1 Oncogenic target selection in LUAD patient sample datasets

5.1.1.1 TCGA database analysis for genotype-specific expression patterns in LUAD

First, I performed a systematic analysis of gene expression profiles of lung adenocarcinoma (LUAD) patients from The Cancer Genome Atlas database (TCGA, https://cancer.gov/tcga) to identify genes putatively involved in LUAD oncogene addiction pathways. The schematic representation of the analysis approach is depicted in **Figure 1A, *left***.

The original LUAD-TCGA dataset contained data for 515 cancer patient samples. 30 samples were excluded from further analysis based on the lack of expression of LUAD molecular markers *NKX2-1* and *NAPSA* identified by Dr. Chul-Min Yang. Four samples were additionally excluded due to sample identifier duplications.

The remaining TCGA-LUAD patient samples were categorized by presence of genomic aberrations (mutations or copy number variations) in LUAD oncogenes *KRAS*, *EGFR, ERBB2*, *BRAF* and *MET* and tumor suppressors *TP53*, *CDKN2A* and *NF1*, based on the information provided by Campbell *et al*. in [15]. The eight considered genes were found most frequently altered in the LUAD-TCGA samples in the analysis of the data provided in [15] performed by Prof. Dr. Sven Diederichs.

To minimize representation of potentially misannotated genes, I merged the resulting eight TCGA-based datasets with the list of gene identifiers annotated by the Human Genome Organization (HUGO) retrieved from the Ensembl hg19/GRCh37 database [208]. To minimize the representation of transcriptional noise, I removed genes that exhibited low median expression (RPKM < 1) in mutant LUAD samples from every mutation-stratified dataset.

Finally, I compared median expression values between LUAD samples harboring one of the mutations and the rest of LUAD samples in every dataset across all HUGO-annotated genes. Per each comparison, I selected genes that were significantly ($p < 0.05$) overexpressed in mutant LUAD samples with a fold change equal or greater than two, as well as genes expressed in mutant samples LUAD whose expression was undetectable in non-mutant LUAD samples (median RPKM = 0).

As a result, I found significantly overexpressed genes between mutant and wildtype LUAD samples with 260 genes for *TP53*-, 55 genes for *KRAS*, 347 genes for *EGFR*, 22 genes for *ERBB2*, 69 genes for *CDKN2A*, 12 genes for *NF1*, 7 genes for *BRAF,* and 80 genes for *MET* mutant versus wildtype LUAD samples. After accounting for overlaps between genes selected across all comparisons, I ended up with 783 genes putatively involved in LUAD oncogene addiction pathways to use as targets in the planned CRISPRi screen.

5.1.1.2 TANRIC database analysis for genotype-specific lncRNAs in LUAD

To expand the target gene library, I further set out to identify long non-coding RNAs (lncRNAs) differentially regulated between LUAD samples exhibiting different mutational signatures as described above. To achieve that, I applied a similar analysis workflow to the TCGA-derived LUAD lncRNA expression dataset obtained from The Atlas of Non-Coding RNA in Cancer database (TANRIC, [210]) with the following modifications (**Figure 1A, *right***). Out of 488 LUAD samples present in the original dataset, 447 were annotated in [15] and met the sample selection criteria, and hence were used in the downstream analysis. Since the annotation of lncRNAs has proven to be a challenging area, I did not restrict the present analysis to HUGO-annotated genes [223]. Additionally, I did not apply the expression level filter as lncRNAs are generally characterized by low expression [143, 223]. As a result, I found significantly overexpressed lncRNAs between mutant and wildtype LUAD samples with 134 lncRNAs for *TP53*, 29 lncRNAs for *KRAS*, 275 lncRNAs for *EGFR*, 19 lncRNAs for *ERBB2*, 28 lncRNAs for *CDKN2A*, 29 lncRNAs for *NF1*, 3 lncRNAs for *BRAF*, and 32 lncRNAs for *MET* mutant

and wildtype LUAD samples. After accounting for overlaps, I ended up with a total of 524 lncRNA candidates to be included among the CRISPRi screen targets.

5.1.1.3 TANRIC database analysis for putatively oncogenic lncRNAs

In order to enrich the CRISPRi target library with novel putatively tumorigenic factors, I identified lncRNAs overexpressed between LUAD and normal lung samples from the LUAD TANRIC dataset. I compared the median expression of all TANRIC-annotated lncRNAs across 460 LUAD samples to their median expression values across 58 normal lung samples included in the dataset, and selected lncRNAs significantly ($p < 0.05$) overexpressed in LUAD samples with a fold change equal or greater than two as well as lncRNAs that were not detectably expressed in normal samples (median RPKM = 0) but detectable in LUAD samples (median RPKM > 0) (**Figure 1B**). In total, I found 813 lncRNAs overexpressed in LUAD tumor samples compared to normal lung tissue and included them among the CRISPRi screen targets.

All in all, I ended up with a list of 2036 gene targets for the prospective CRISPRi screening for novel oncogenic factors in LUAD derived from systematic analysis of publicly available cancer patient expression datasets (**Figure 1C**). After manual addition of 40 positive control genes [116, 211], as well 36 manually curated target genes of interest and removing overlaps, I obtained a list of 2098 genes to target.

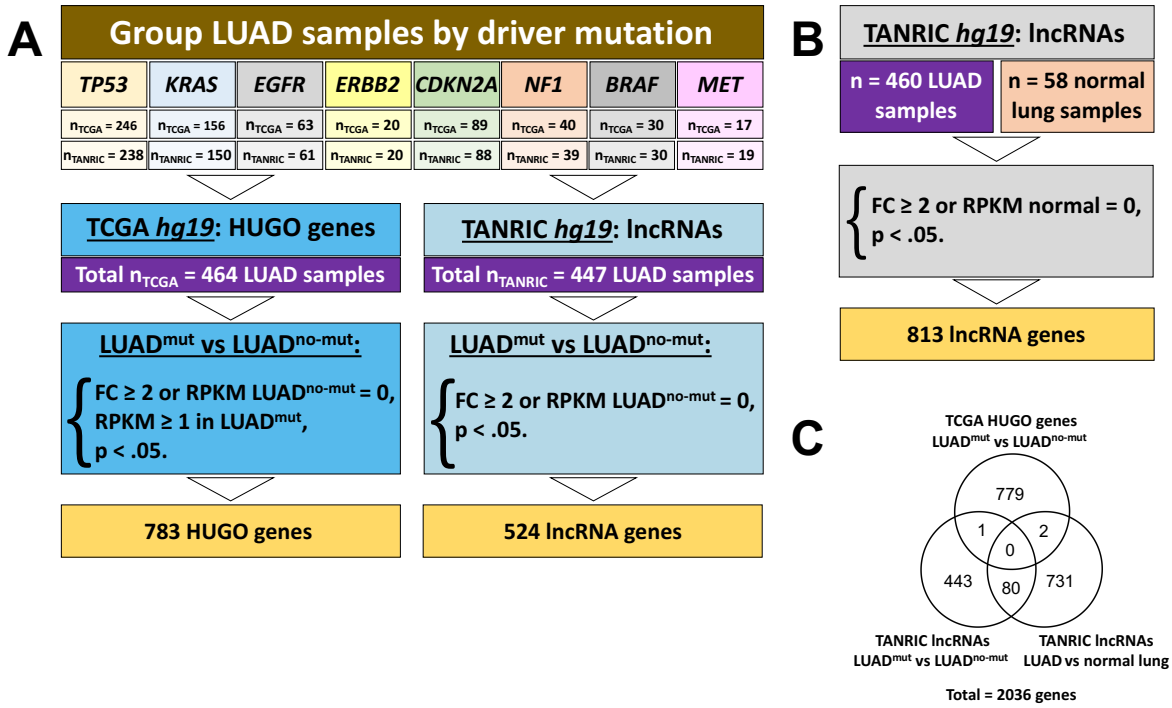


Figure 1. Screen target selection workflow.
(A) Identification and selection of HUGO-annotated genes and TANRIC-annotated lncRNAs differentially expressed in LUAD tumors with different mutational signatures. $LUAD^{mut}$ stands for LUAD samples expressing a certain driver mutation, $LUAD^{no\text{-}mut}$ stands for tumor samples without mutations in the particular driver gene; FC = fold change.
(B) Identification and selection of lncRNAs upregulated in LUAD vs. normal lung samples annotated in the TANRIC database.
(C) Venn diagram representing overlaps between screen targets selected via different approaches.

5.1.2 Custom sgRNA design algorithm for pooled CRISPRi screening

Since the CRISPRi system functions most effectively when targeted immediately downstream of transcription start sites (TSSs) [164], prediction of functional TSSs is essential for a successful screening setup. The 2098 genes selected for the CRISPRi screening corresponded to 7038 TSSs, with an average of 3.4 TSSs per gene.

Since I intended to produce a universal library to screen a panel of LUAD cell lines, and TSS usage is specific to cell type [224], I set out to differentially target every TSS within each gene of interest. To achieve an appropriate targeting resolution, I binned all TSSs corresponding to the selected target genes located within 200 bp from each other into a single

target based on their genomic coordinates retrieved from Ensembl hg19/GRCh37. As a result, I obtained a list of 4633 TSS target sites.

Next, I designed sgRNAs against the selected TSSs. Together with Dr. Andrew Walsh, we created a custom algorithm tailored to CRISPRi-specific pooled sgRNA library design outlined in **Figure 2**. Since CRISPRi was found to efficiently repress gene expression only when targeted within a 1000 bp window around the TSS [164], we implemented a CRISPRi-specific TSS-wide off-target prediction as the main feature of our algorithm. Thus, only off-target sites that fell within 1000 bp around any TSS other than the target TSS (or another TSS corresponding to the same target gene) were considered, and corresponding sgRNA designs were removed from the output. Additionally, we used genome-wide off-target mapping as a secondary sgRNA quality filter (**Figure 2A**).

In the end, we were able to generate a library of 40357 sgRNAs against 4620 out of 4633 target TSSs corresponding to 2084 out of 2098 genes of interest (**Figure 2B**). 78% of the TSSs were targeted with 9 sgRNAs/TSS, and only 1.1% of TSSs was targeted with less than 5 sgRNAs (**Figure 2C**). The final library included negative control sgRNAs selected from [164] and contained a total of 42000 sgRNAs.

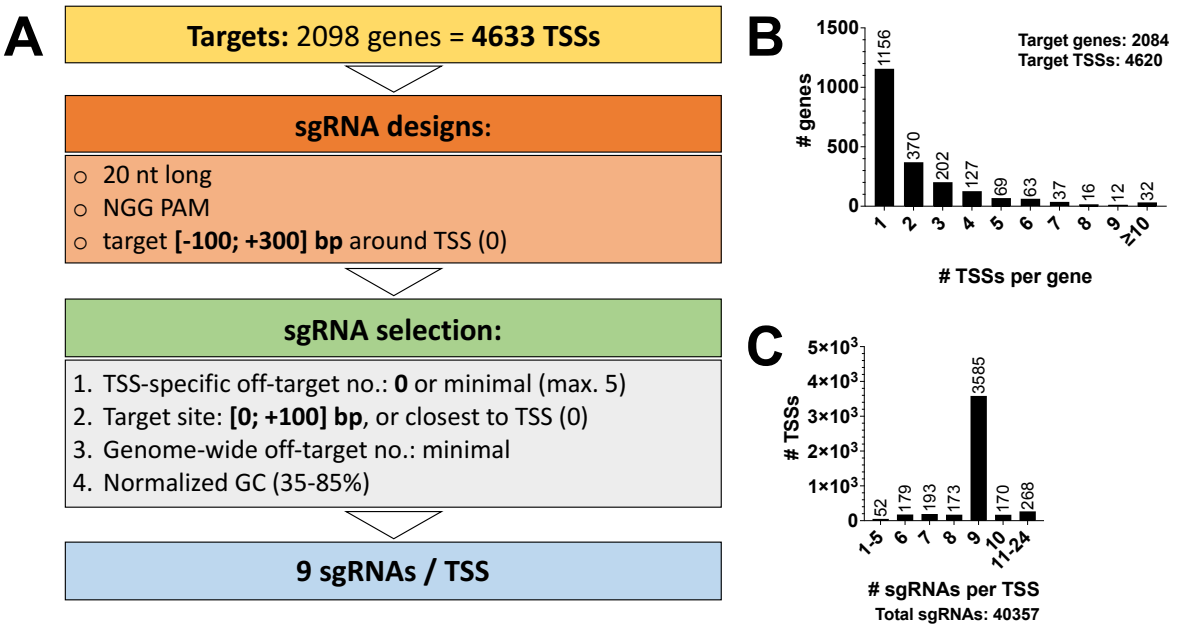


Figure 2. Pooled CRISPRi sgRNA library design.
(A) Custom sgRNA design algorithm for pooled CRISPRi screening library. TSS = transcriptional start site; nt = nucleotide; PAM = protospacer adjacent motif.
(B) Distribution of number of CRISPRi-targeted TSSs per gene in genes with successfully designed sgRNAs.
(C) Distribution of number of successfully designed sgRNAs per targeted TSS.

5.1.3 Pooled negative CRISPRi screening in *EGFR*-mutant vs. -wildtype LUAD cell lines yielded reproducible results

Since most of the screen targets originated from the comparison between *EGFR*-mutant and -wildtype LUAD samples (**Figure 1**), I decided to perform the CRISPRi screen in four LUAD cell lines harboring oncogenic *EGFR* mutations (PC-9, NCI-H1650, NCI-H1975, HCC827) and four *EGFR*-wildtype LUAD cell lines (Calu-6, NCI-H3122, NCI-H522, NCI-H838). Mutational signatures of LUAD cell lines used in this study were derived from the COSMIC Cell Lines Project database (https://cancer.sanger.ac.uk/cell_lines) [204]. The PC-9 cell line was not annotated in COSMIC, however, has been widely characterized as *EGFR*-mutant in literature [219–221]. PC-9, NCI-H1650 and HCC827 LUAD cell lines were annotated as harboring deletions in exon 19 of the *EGFR* gene, and NCI-H1975 cell lines expressed L858R and T790M *EGFR* mutations. The selected *EGFR*-wildtype cell lines also did not bear any

mutations in *ERBB2* and other genes from the epidermal growth factor receptor (ERBB) family.

The schematic screening process is outlined in **Figure 3**. Briefly, the 42000 sgRNA library was synthesized in a form of an oligo pool and cloned into the LentiCRISPRv2-dCas9-KRAB(iv) lentiviral backbone [147] via Gibson assembly, resulting in a plasmid library. The plasmid library was then transfected into HEK293T cells along with lentiviral packaging plasmids to produce the lentiviral library. Then, the target cell lines were transduced with the lentiviral library and cultured for 21 days, allowing for negative selection (dropout) of cells expressing sgRNAs targeting essential genes in the respective cell line.

At the endpoint, I harvested the screened cells, PCR-amplified the sgRNA-encoding lentivirally integrated sequences from the genomic DNA and subjected the resulting amplicons to next-generation sequencing (NGS). Then, I compared the representation of the sgRNA reads at the endpoint to the representation of sgRNAs in the reference plasmid library using the MAGeCK algorithm [213] to identify the depleted sgRNAs and their corresponding targets. Throughout all steps of the screening process, I maintained a minimal 500-fold coverage per sgRNA.

I performed two biological replicates of the screen per cell line. Targets were regarded as depleted in a particular cell line if their average $\log_2$(FC) value in both screen replicates of the particular cell line was equal to or smaller than -1. Screen targets that were found significantly (FDR < 0.05) depleted in both replicates of one cell line were regarded as hits in the particular cell line.

As a result, I identified 71 hits in Calu-6, 34 hits in NCI-H3122, 6 hits in NCI-H522 and 25 hits in NCI-H838 *EGFR*-wildtype cell lines, and 34 hits in PC-9, 48 hits in HCC-827, 13 hits in NCI-H1650 and 6 hits in NCI-H1975 *EGFR*-mutant cell lines. For five out of eight tested cell lines, the Pearson correlation coefficient for $\log_2$(FC) values of all targets (R_{all}) between the two replicates was greater than 0.8, for NCI-H1975 the coefficient was 0.79, for NCI-H1650 the coefficient was 0.67 and for NCI-H3122 the coefficient was 0.41. The correlation coefficient between $\log_2$(FC) values of the screen hits (R_{hits}) was high in seven out eight cell lines, with values greater than 0.9 for PC-9 and NCI-H522, between 0.8 and 0.9 for Calu-6, NCI-H838 and HCC827, 0.73 for NCI-H1975 and 0.65 for NCI-H3122, indicating good overall reproducibility of the screen results (**Figure 4**). For only one out eight cell lines, namely NCI-H1650, the correlation coefficient between $\log_2$(FC) values of the screen hits was 0.16.

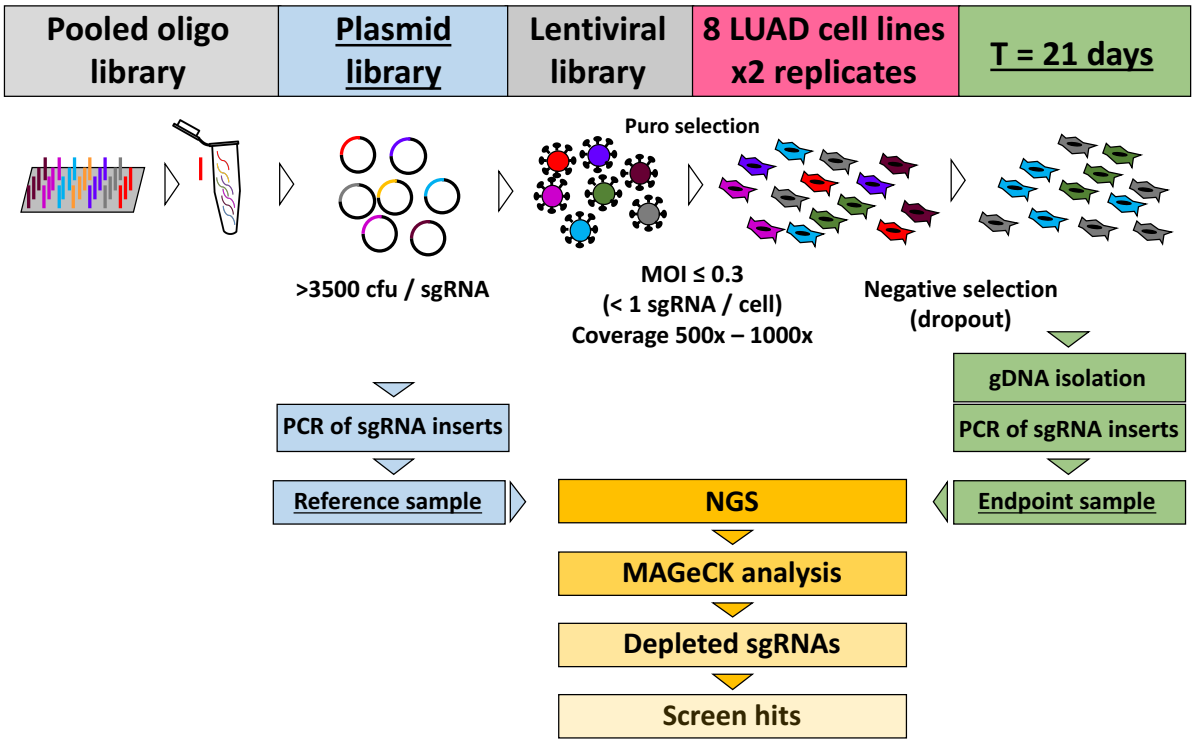


Figure 3. Pooled CRISPRi screening workflow.
Cfu = colony forming units; MOI = multiplicity of infection; LUAD = lung adenocarcinoma; NGS = next-generation sequencing; puro = puromycin.

Oligos encoding the 42000 sgRNA library were cloned at high coverage to produce the plasmid library. The plasmid library was transfected into HEK293T cells along with lentiviral packaging plasmids to produce the lentiviral library.
Eight LUAD cell lines were transduced with the lentiviral library at MOI 0.3 and 500x-1000x minimal coverage. Stably transduced cells were selected with puromycin and passaged for 21 days after transduction at 500x-1000x minimal coverage. At the endpoint, cells were harvested at a minimal coverage of 1000x, followed by genomic DNA isolation and a two-step PCR aimed at amplifying integrated sgRNA-encoding sequences.
The resulting amplicons were subjected to NGS on Illumina HiSeq2000 platform at the minimal coverage of 500 reads per sgRNA sequence. The sequencing data were analyzed using MAGeCK pipeline [213] to identify depleted sgRNAs and the corresponding screen hits.

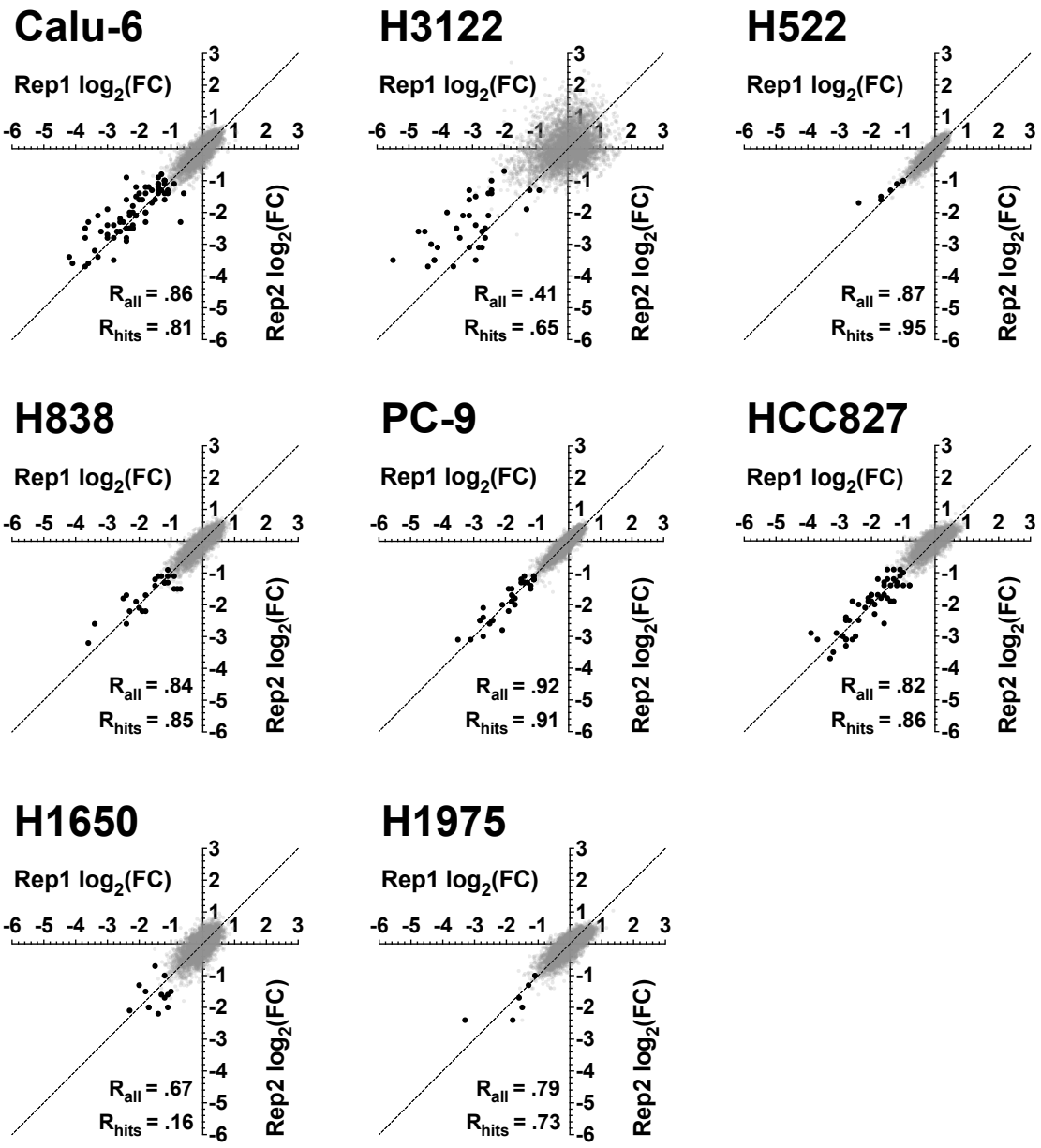


Figure 4. Correlations of fold change (FC) values representing the performance of the screen targets between the CRISPRi screen replicates per tested LUAD cell line.

Each black datapoint represents a TSS selected in the CRISPRi screen in both replicates of the particular cell line based on the log_2(FC) and FDR thresholds calculated using MaGECK pipeline. Grey datapoints represent TSSs that were not selected in the particular cell line. FC values are represented as log_2(FC).

R_{all} stands for Pearson correlation coefficient calculated based on the log_2(FC) values across all targeted TSSs between two replicates of the particular cell line; R_{hits} stands for Pearson correlation coefficient calculated for the selected hits.

5.1.4 Screen hits were detectably expressed in the respective LUAD cell lines

Analysis of the lung cell line RNA expression dataset from the Diederichs lab [205] revealed that the target library was enriched in transcripts exhibiting low expression (FPKM < 1) values in the panel of screened cell lines (**Supplementary figure 1**). Notably, 75 – 100 % of screen hits selected in each cell line were expressed at FPKM values equal or greater than 1 (**Figure 5**). In the NCI-H522 cell line, all six selected hits were expressed at FPKM ≥ 1. Across the remaining seven cell lines, screen targets with expression values of FPKM ≥ 1 were on average 25 times more likely to be selected as screen hits compared to targets with FPKM < 1 ($p = 2.8*10^{-6}$).

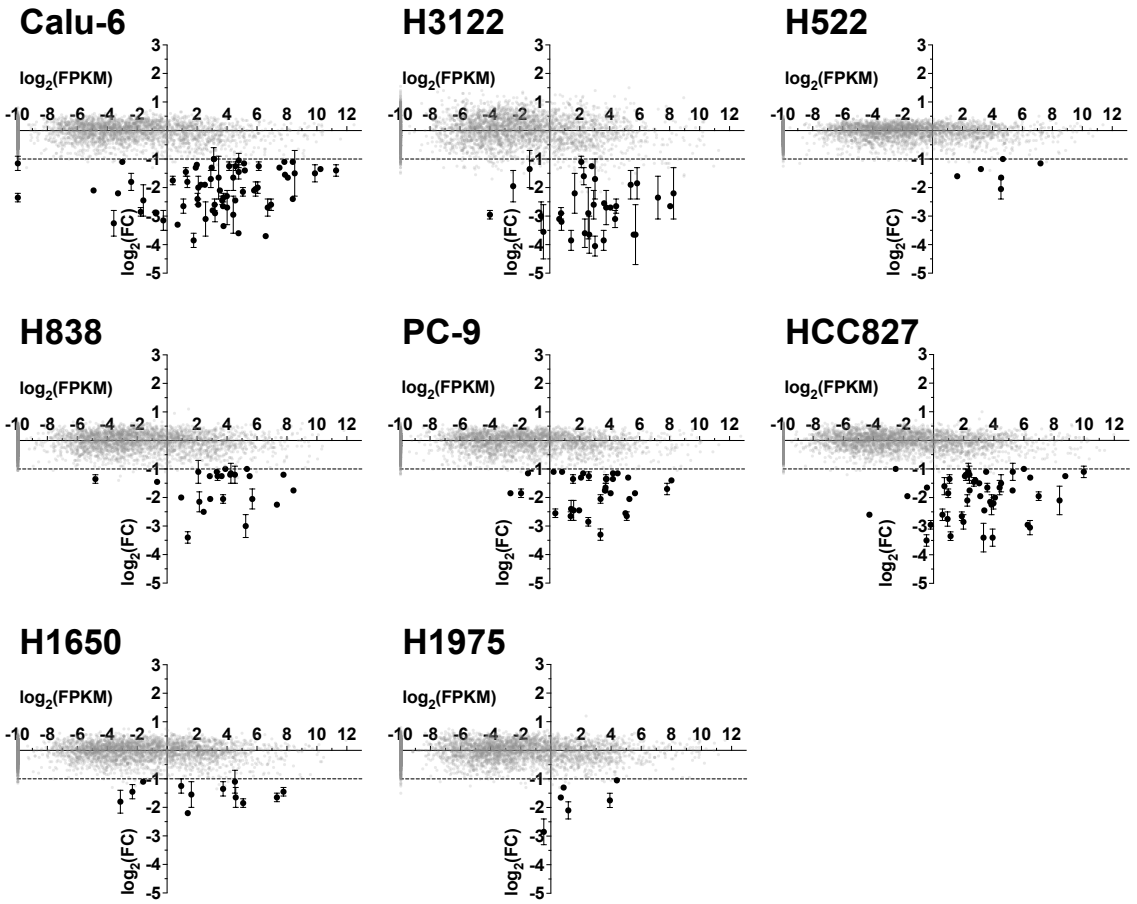


Figure 5. Hits identified in the CRISPRi dropout screen were detectably expressed in the tested LUAD cell lines.

Correlations between the fold change (FC) and the RNA expression FPKM values of screen targets per tested LUAD cell line. Black datapoints represent TSSs selected in the CRISPRi screen based on the $\log_2$(FC) and FDR thresholds in each cell line. Grey datapoints represent target TSSs that were not selected in the particular cell line. FC values are represented as mean $\log_2$(FC) between two replicates of each cell line, bars represent range. FPKM values are represented as $\log_2$(FPKM+0.001).

5.1.5 The TSS-level screen resolution allowed identification of hits that could not be predicted by position or RNA expression profiling

Next, I examined the relationship between the expression level, genomic position, and screen performance of TSSs selected as screen hits for genes expressed from multiple TSSs. This analysis was performed to determine whether targeting of every TSS within each gene of interest provided an additional advantage to the screen design or whether a selection of the first TSS in the target gene or the TSS giving rise to the highest expression, judged by the expression of the immediate downstream exon, could have yielded the same level of information.

The average frequency of the TSS with the highest expression value of the immediate downstream exon being selected as a screen hit within a gene with multiple TSS annotations (a multi-TSS gene), calculated across all cell lines, was not significantly different from the frequency occurring by chance, indicating that there was no bias towards selection of most highly expressed TSSs in the CRISPRi screen. Meanwhile, TSSs originating from gene starts were selected as screen hits in multi-TSS genes 2.1-fold more frequently than expected by chance ($p = 2.7*10^{-5}$), suggesting that the first TSSs were selected relatively often.

Additionally, I performed a detailed analysis of the screen results from the Calu-6 cell line to examine the specific expression and position characteristics of the TSSs selected in the screen. I chose to focus on the Calu-6 cell line as it yielded the most screen hits compared to the other cell lines.

71 TSS hits selected in Calu-6 corresponded to 58 genes. Out of the 58 genes, 44 had multiple TSSs targeted in the screen. Out of those 44 genes, two TSSs were selected as screen hits in 13 genes, and one TSS was selected in 31 genes.

The expression analysis based on the lung cell line RNA expression dataset from the Diederichs lab [205] revealed that in 14 genes out of the 31, the expression value corresponding to the first exon downstream of the selected TSS was at least two-fold lower compared to the TSS with the highest expression within the respective gene, suggesting that not the TSS giving rise to the highest expression had been selected in the screen in these cases. For the remaining 17 genes, the selected TSS was either upstream of the most strongly expressed exon, or the fold change of expression values between the most strongly expressed exon within the gene and the exon following the selected TSS was lower than two-fold.

The combined expression and position analysis based on the genomic coordinates provided in the GRCh37/hg19 human genome assembly revealed that for 21 out of 31 genes, the selected TSS corresponded to the gene start. In four genes out of 31, the selected TSS was not the first by genomic position, but the first TSS with detectable expression of the subsequent exon within the gene with the detectability cutoff set to FPKM > 1. In six genes out of the 31, the selected TSS was neither the first by genomic position nor the first detectably expressed TSS within the gene, indicating that neither expression nor genomic position would have allowed the identification of the experimentally selected TSS in these cases. In three genes out of these six (*RPL11, PLK4, CHEK1*), the expression value corresponding to the selected TSS was at least two-fold lower compared to the TSS with the highest expression within each gene (**Figure 6A**).

TSSs selected within the *PLK4* and *CHEK1* genes gave rise to transcripts annotated as protein coding in the hg19/GRCh37 human genome assembly, while the selected TSS of *RPL11* gave rise to a non-coding transcript (**Table 15**). Each TSS located both upstream and downstream of the selected *PLK4, CHEK1* and *RPL11* TSSs gave rise to a number of transcripts, including transcripts annotated as protein-coding, suggesting that the screen setup designed to target all TSSs allowed distinguishing between different functional isoforms.

Among the 13 genes with two TSS hits, the first and second consecutive TSSs were selected for 10 genes. In all 10 genes the genomic distance between the selected TSSs was smaller than the CRISPRi efficiency window of 1000 bp [164].

For the remaining three genes, *CDCA3*, *PLK1* and *MCM2*, other than the first two TSSs were selected in the screen. For *CDCA3*, two non-consecutive TSSs with a genomic distance greater than 1000 bp were selected, which could indicate that two independent isoforms were phenotypically selected (**Figure 6B**). The downstream exon of the first TSS selected within *CDCA3* was expressed at an approximately 50-fold lower level compared to the maximum expression level of the other TSSs. The other selected TSS, third by genomic position, corresponded to the maximum expression value within the gene. The first TSS selected within the *CDCA3* gene gave rise to a protein-coding transcript, while the other one gave rise to a non-coding transcript (**Table 15**), suggesting putative functional relevance or a possible misannotation of the non-coding transcript. The non-selected TSS located between the two selected *CDCA3* TSSs gave rise to several protein-coding as well as non-coding

transcripts. *CDCA3* TSSs located downstream of the selected TSSs gave rise to non-coding transcripts.

In the genes *PLK1* and *MCM2*, two consecutive TSSs were selected that were neither first by genomic position nor first detectably expressed within the respective gene. Moreover, both TSSs in both genes were expressed at least two-fold lower than the respective most highly expressed TSSs (**Figure 6B**). Within the *PLK1* gene, the two selected TSSs corresponded to protein-coding transcripts, while the remaining TSSs gave rise exclusively to non-coding transcripts, suggesting that the screen setup allowed to differentially detect functionally relevant *PLK1* protein-coding isoforms. For *MCM2,* both selected TSSs gave rise to protein-coding transcripts, and upstream and downstream TSSs gave rise to protein- as well as non-coding isoforms, indicating that in this case the screen setup allowed to differentiate between functionally relevant protein-coding isoforms.

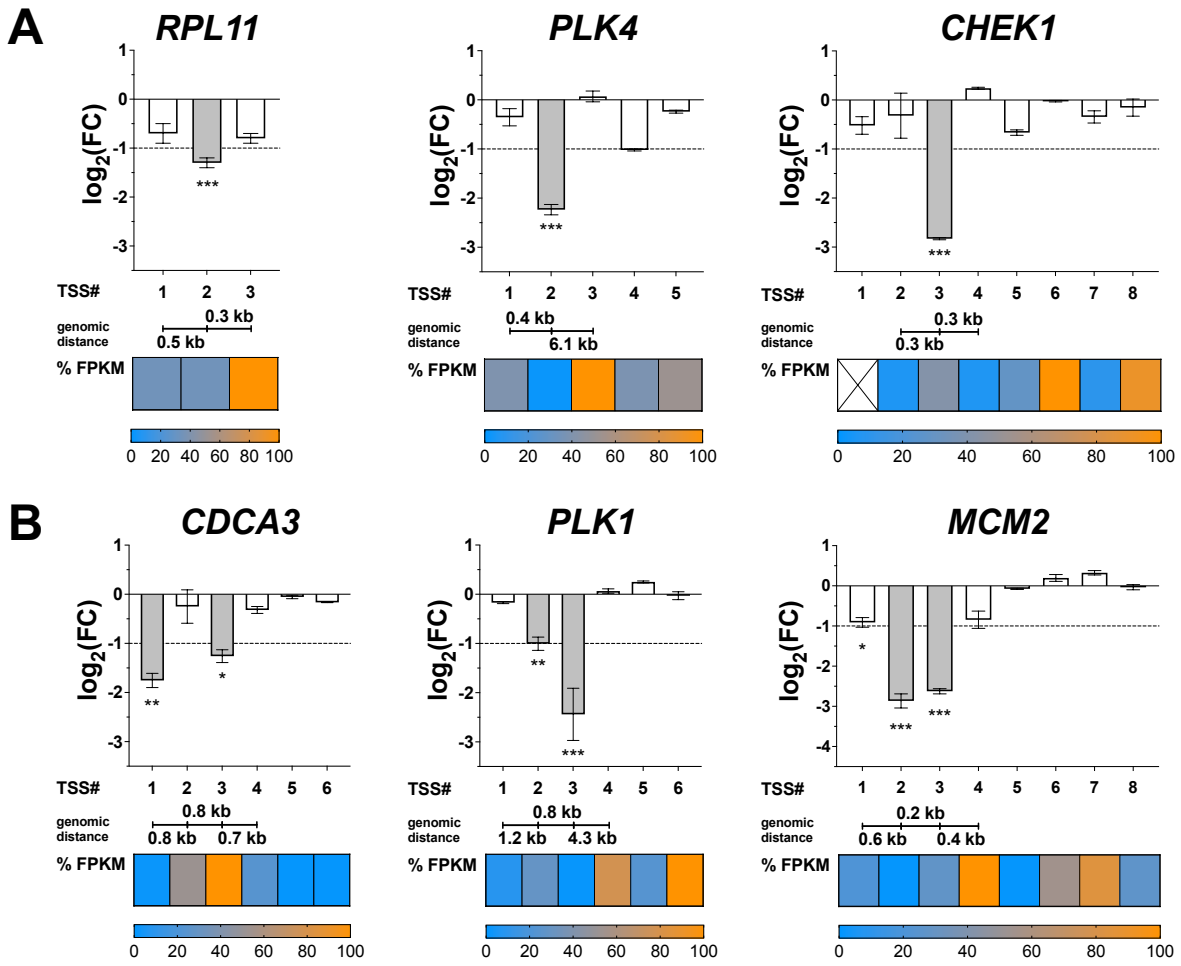


Figure 6. A fraction of TSSs selected within a gene in the dropout CRISPRi screen could not be predicted by genomic position or RNA expression.

Genes containing single **(A)** and multiple **(B)** TSSs selected by the CRISPRi screen in Calu-6 cells that were not first and most expressed within a gene.

Log_2(FC) values per TSS were calculated using MaGECK algorithm.

Grey bars correspond to TSSs selected in the CRISPRi screen using log_2(FC) and FDR thresholds. Bar heights represent mean of two screen replicates, error bars represent range. FDR values were calculated by MaGECK algorithm per each replicate; asterisks represent mean FDR per TSS of two screen replicates; ***, FDR < 0.001; **, FDR < 0.01; *, FDR < 0.05.

Genomic distance was calculated using coordinates provided in GRCh37/hg19 human genome assembly.

Heatmaps represent row-normalized FPKM values corresponding to the first exon downstream of each TSS (100% assigned to the maximal value per comparison). White crossed out cells represent conditions where data were not available.

Table 15. *hg19/GRCh37 annotations of all transcripts corresponding to the genes where the TSSs selected in the CRISPRi screen in the Calu-6 cell line was neither first nor most expressed within the gene*

Gene name; Gene ID	Transcript ID	Transcript type hg19/GRCh37	Target TSS ID	Screen hit
***RPL11*; ENSG00000142676**	ENST00000374550	protein_coding	RPL11-1	NO
	ENST00000467075	processed_transcript		
	ENST00000482370	**processed_transcript**	**RPL11-2**	**YES**
	ENST00000443624	protein_coding	RPL11-3	NO
	ENST00000458455	protein_coding		
***PLK4*; ENSG00000142731**	ENST00000270861	protein_coding	PLK4-1	NO
	ENST00000511942	processed_transcript		
	ENST00000503914	retained_intron		
	ENST00000515069	protein_coding		
	ENST00000513090	protein_coding		
	ENST00000507249	protein_coding		
	ENST00000514379	**protein_coding**	**PLK4-2**	**YES**
	ENST00000510605	retained_intron	PLK4-3	NO
	ENST00000507454	retained_intron	PLK4-4	NO
	ENST00000508113	protein_coding		
	ENST00000510192	retained_intron	PLK4-5	NO
***CHEK1*; ENSG00000149554**	ENST00000438015	protein_coding	CHEK1-1	NO
	ENST00000525396	protein_coding	CHEK1-2	NO
	ENST00000427383	protein_coding		
	ENST00000428830	**protein_coding**	**CHEK1-3**	**YES**
	ENST00000532449	**processed_transcript**		
	ENST00000528761	**processed_transcript**		
	ENST00000544373	**protein_coding**		
	ENST00000527013	**protein_coding**		
	ENST00000526937	**protein_coding**		
	ENST00000534685	**protein_coding**		
	ENST00000533778	protein_coding	CHEK1-4	NO
	ENST00000534070	protein_coding		
	ENST00000524737	protein_coding	CHEK1-5	NO
	ENST00000531607	processed_transcript		
	ENST00000532669	protein_coding		
	ENST00000278916	protein_coding		
	ENST00000531062	retained_intron	CHEK1-6	NO
	ENST00000528276	retained_intron	CHEK1-7	NO
	ENST00000498122	processed_transcript	CHEK1-8	NO

Gene name; Gene ID	Transcript ID	Transcript type hg19/GRCh37	Target TSS ID	Screen hit
***CDCA3*; ENSG00000111665**	**ENST00000538862**	**protein_coding**	**CDCA3-1**	**YES**
	ENST00000545368	processed_transcript	CDCA3-2	NO
	ENST00000446553	retained_intron		
	ENST00000422785	protein_coding		
	ENST00000536241	retained_intron		
	ENST00000535406	protein_coding		
	ENST00000229265	protein_coding		
	ENST00000540683	protein_coding		
	ENST00000535871	**processed_transcript**	**CDCA3-3**	**YES**
	ENST00000544610	processed_transcript	CDCA3-4	NO
	ENST00000604599	processed_transcript	CDCA3-5	NO
	ENST00000603043	processed_transcript	CDCA3-6	NO
***PLK1*; ENSG00000166851**	ENST00000562272	retained_intron	PLK1-1	NO
	ENST00000300093	**protein_coding**	**PLK1-2**	**YES**
	ENST00000570220	**nonsense_mediated_decay**		
	ENST00000564202	**processed_transcript**		
	ENST00000567897	**protein_coding**	**PLK1-3**	**YES**
	ENST00000568568	**protein_coding**		
	ENST00000562407	retained_intron	PLK1-4	NO
	ENST00000564947	retained_intron	PLK1-5	NO
	ENST00000564794	retained_intron	PLK1-6	NO
***MCM2*; ENSG00000073111**	ENST00000265056	protein_coding	MCM2-1	NO
	ENST00000474964	nonsense_mediated_decay		
	ENST00000480910	**protein_coding**	**MCM2-2**	**YES**
	ENST00000472731	**protein_coding**	**MCM2-3**	**YES**
	ENST00000477668	nonsense_mediated_decay	MCM2-4	NO
	ENST00000468659	processed_transcript	MCM2-5	NO
	ENST00000491422	protein_coding	MCM2-6	NO
	ENST00000473785	retained_intron	MCM2-7	NO
	ENST00000468414	processed_transcript	MCM2-8	NO

5.1.6 Neither common nor *EGFR* status-specific lncRNAs essential for LUAD viability could be confidently identified by the CRISPRi screen

In all eight tested LUAD cell lines, only one common TANRIC-annotated lncRNA was selected as a screen hit, which, upon closer examination, turned out to be a protein-coding gene, *CASP8AP2*, which was misannotated in the hg19/GRCh37 genome assembly (**Table 16**). In four out of eight cell lines, no additional TANRIC-annotated targets were selected as screen hits. In total, one lncRNA, *DDX11-AS1,* was selected in two cell lines and additional six lncRNAs were selected in one of three cell lines. Thus, due to the absence of both common and *EGFR* status-specific lncRNA screen hits, I did not further pursue the lncRNA line of research within this project.

Besides, all seven selected lncRNAs were transcribed in an antisense orientation to protein-coding genes, which posed a risk of the concurrent transcriptional repression of the neighboring TSS. As such, six out of seven lncRNAs were deemed “non-CRISPRable” according to my previously published analysis based on the presence of a TSS from another gene within the 1000 bp CRISPRi efficiency window around the target TSS (**Table 16**) [147]. Therefore, it could not be excluded that their screen performance was attributable to the silencing of the neighboring gene.

Table 16. *Characteristics of TANRIC-annotated screen hits*

Transcript ID; Gene name GRCh37	**Screen hit in cell lines**	**Transcript type GRCh37**	**Transcript type GRCh38**	**CRISPRable lncRNA** [147]	**Distance to neighbor TSS GRCh37**
ENST00000552401; *CASP8AP2*	all	processed_ transcript	protein_ coding	NA	>10 kb to *MDN1*
ENST00000500527; *DDX11-AS1*	HCC827; H1650	antisense	lncRNA	No	2 bp to *DDX11*
ENST00000446492; *AC093391.2*	HCC827	antisense	lncRNA	No	18 bp to *DARS*
ENST00000456519*; AC010884.1*	Calu-6	antisense	lncRNA	No	513 bp to *MRPS9*
ENST00000582938*; RP11-1148O4.2*	Calu-6	antisense	lncRNA	No	325 bp to *NSRP1*
ENST00000428791; *RP11-269F19.2*	Calu-6	antisense	lncRNA	No	174 bp to *RPS8*
ENST00000578831; *RP11-739L10.1*	Calu-6	antisense	lncRNA	No	62 bp to *RBBP8*
ENST00000608442; *AFAP1-AS1*	H838	antisense	lncRNA	Yes	>25 kb to *AFAP1*-

5.1.7 CRISPRi screens identified *CASP8AP2* as an essential viability factor for LUAD cell lines

CASP8AP2 and *TPX2* TSSs were selected as screen hits across all cell lines (**Figure 7**). For both corresponding genes, only the selected TSS was targeted in the screen. *TPX2* was listed among the core essential positive controls included to the screen from [211], while *CASP8AP2* was included in the library based on its deregulation in LUAD patient samples. Thus, I set out to further investigate the function of *CASP8AP2* in LUAD (see Section 5.3).

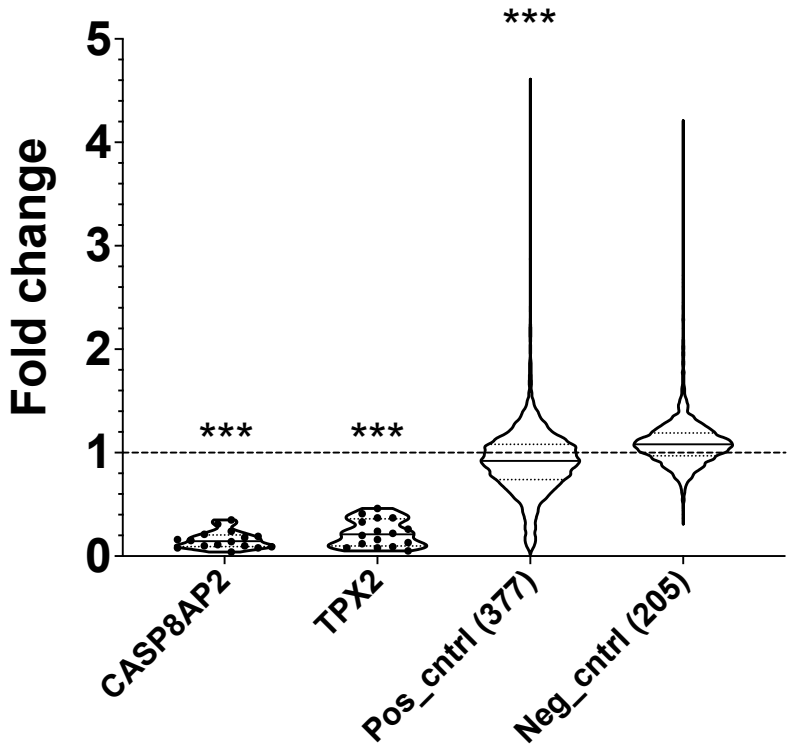


Figure 7. CRISPRi dropout screen identified *CASP8AP2* as an essential viability factor in LUAD cell lines.
Fold change (FC) of the screen hits *CASP8AP2* and *TPX2* that were selected based on $\log_2$(FC) and FDR thresholds in all tested LUAD cell lines are plotted alongside cumulative FC of positive control (Pos_cntrl) and negative control (Neg_cntrl) screen targets. Each datapoint represents a fold change value in an individual screen replicate. Fold change values were calculated using the MaGECK pipeline. Significance testing against negative control (Neg_cntrl) was performed on $\log_2$-transformed fold change data using two-tailed unpaired Student's t-test with Welch's correction, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

5.1.8 CRISPRi screens suggested differential response of LUAD cell lines to depletion of mitotic factors depending on *EGFR* status

Comparison between screen hits selected in *EGFR*-wildtype and -mutant cell lines revealed five differentially selected genes encoding various mitotic factors (**Figure 8**).

CDC20 and *PLK1* were selected in three *EGFR*-wildtype cell lines Calu-6, NCI-H3122 and NCI-H838 and in none of the *EGFR*-mutant cell lines. *SPAG5* and *CENPE* were selected in three *EGFR*-wildtype cell lines (Calu-6, NCI-H3122 and either NCI-H522 or NCI-H838, respectively) and in one *EGFR*-mutant cell line HCC827. *CEP55* was selected in Calu-6 and NCI-H838 *EGFR*-wildtype cell lines and in none of the *EGFR*-mutant cells.

Comparison of the cumulative results demonstrated that the difference in the negative selection fold change between *EGFR*-wildtype and -mutant cell lines was significant for four out of the five hits (*CDC20*, *PLK1*, *CENPE* and *CEP55*) (**Figure 8A**). Analysis of the lung cell line RNA expression dataset from the Diederichs lab [205] showed a significant difference in the expression level between *EGFR*-wildtype and -mutant cells for only one (*SPAG5*) out of the five TSSs (**Figure 8B**).

Therefore, I proceeded to the validation of the suggested differential phenotypes upon the depletion of the selected mitotic factors. I chose to knockdown *CENPE* and *CEP55* for the initial follow-up studies as the screen results suggested the biggest difference in phenotypic response to their repression between *EGFR*-wildtype and -mutant cell lines (**Figure 8A**).

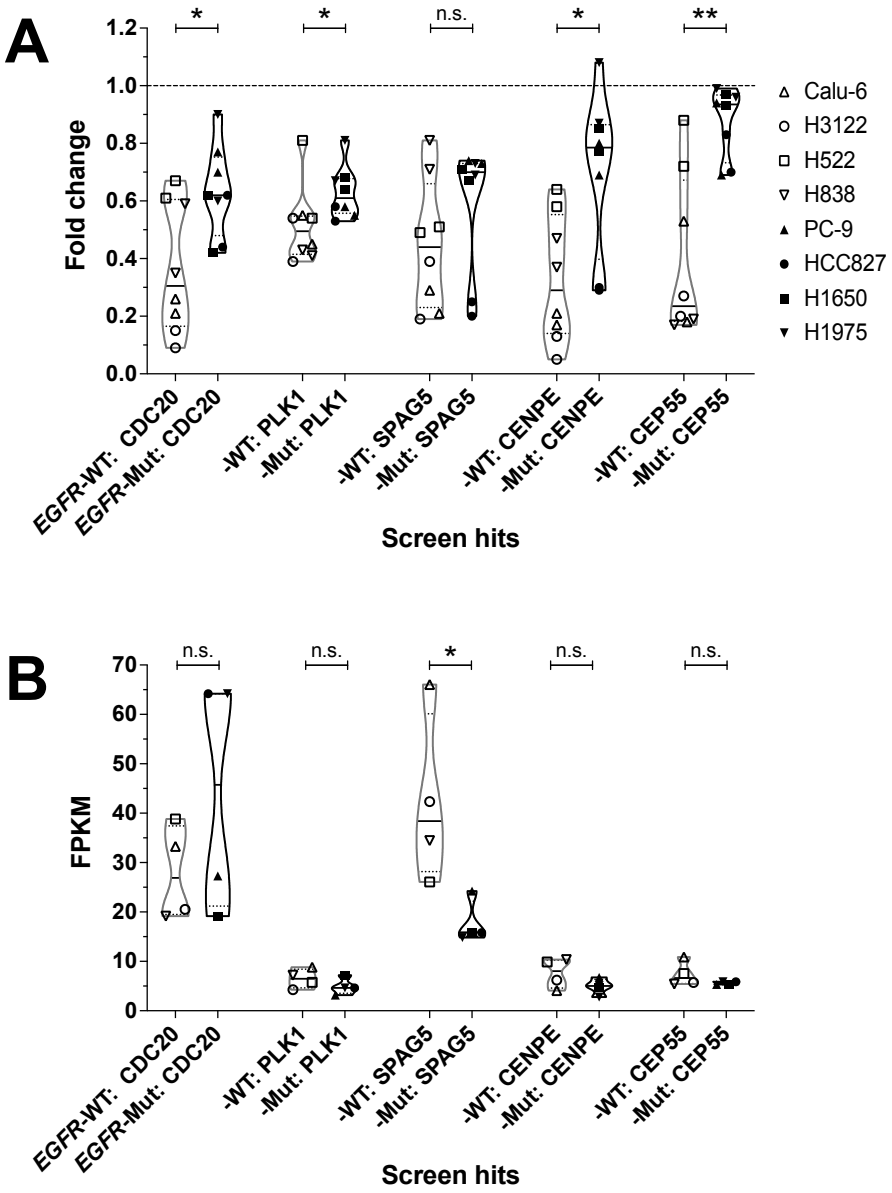


Figure 8. Screen hits differentially selected between *EGFR*-wildtype and -mutant LUAD cell lines.

EGFR-wildtype (WT) group includes Calu-6, NCI-H3122 (H3122), NCI-H522 (H522) and NCI-H838 (H838) cell lines. *EGFR*-mutant (Mut) group includes PC-9, HCC827, NCI-H1650 (H1650) and NCI-H1975 (H1975) cell lines.

(A) Fold change (FC) and **(B)** RNA expression FPKM values of screen hits that were selected based on $\log_2$(FC) and FDR thresholds in min. three out of four WT cell lines and max. one out of four Mut cell lines, or in min. two out of four WT cell lines and in no Mut cell lines. Each datapoint represents a fold change value per individual screen replicate. Fold change values were calculated using MaGECK pipeline. Significance was calculated on $\log_2$-transformed fold change data using two-tailed unpaired Student's t-test with Welch's correction where appropriate, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; "n.s.", not significant.

5.2 Characterization of the differential responses of *EGFR*-wildtype and -mutant LUAD cell lines to the loss of *CENPE* expression

5.2.1 *EGFR*-wildtype LUAD cell lines were more sensitive to the silencing of the mitotic factors *CENPE* and *CEP55* compared to *EGFR*-mutant cells

To validate the phenotypes observed in the CRISPRi screen, I silenced *CENPE* and *CEP55* with three individual sgRNAs per gene in three *EGFR*-wildtype LUAD cell lines exhibiting the strongest negative selection of *CENPE* and *CEP55* in the CRISPRi screen (Calu-6, NCI-H522 and NCI-H838, **Figure 8**) and the three least-responding *EGFR*-mutant cell lines (PC-9, NCI-H1650 and NCI-H1975). I selected the top ranking sgRNAs per gene from the CRISPRi screening library based on their negative selection scores across all cell lines.

All sgRNAs induced an efficient *CENPE* knockdown in all tested cell lines (60 – 100% efficiency) (**Supplementary figure 2**), which resulted in a complete loss of cell viability in *EGFR*-wildtype cell lines (decrease by 80 - 97%) and moderate to no negative effect on cell viability in *EGFR*-mutant cells (**Figure 9**). Silencing of *CEP55* expression was also highly efficient in all cell lines (76 – 99% knockdown) (**Supplementary figure 3**), however, it did not affect the cell viability of *EGFR*-wildtype and -mutant cell lines in the predicted differential pattern (**Figure 9**). Therefore, I set out to further investigate the link between *EGFR* status and the sensitivity of LUAD cell lines to the depletion of *CENPE* expression.

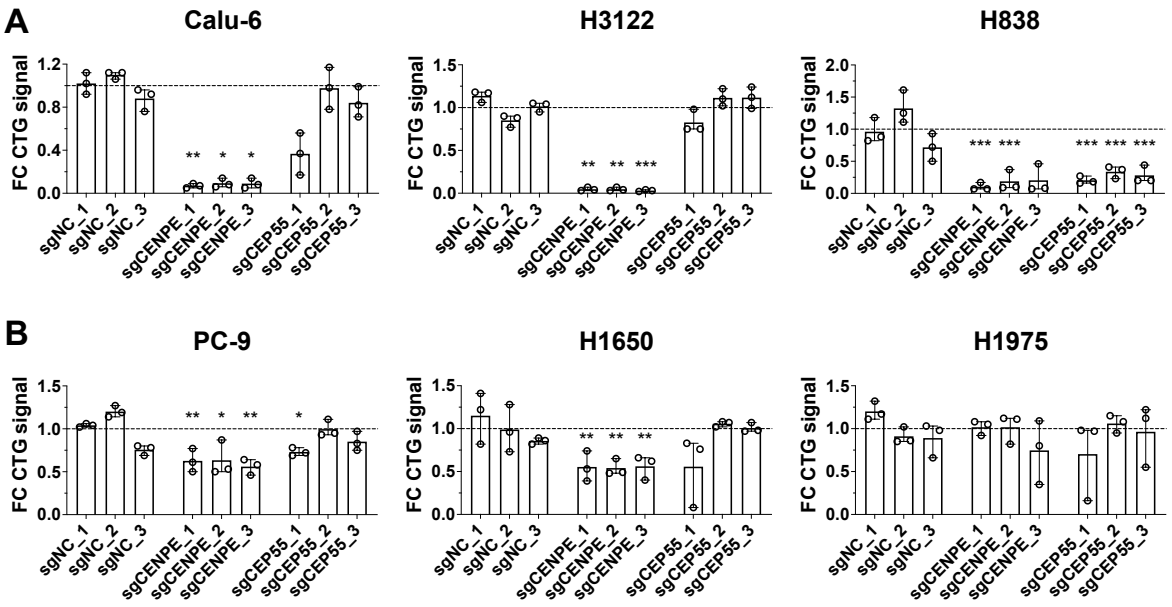


Figure 9. Validation of the differential effects of *CENPE* and *CEP55* knockdowns (KDs) on cell viability in *EGFR*-wildtype and -mutant LUAD cell lines observed in the CRISPRi screen.

CellTiter-Glo cell viability assay (CTG) was performed to assess the phenotype of LUAD cells transduced independently with three CRISPRi constructs expressing sgRNAs against *CENPE* (sgCENPE) and sgRNAs against *CEP55* (sgCEP55) compared to negative control sgRNAs (sgNC) (n = 3). *EGFR*-wildtype (WT) group includes Calu-6, NCI-H3122 (H3122) and NCI-H838 (H838) cell lines **(A)**. *EGFR*-mutant (Mut) group includes PC-9, NCI-H1650 (H1650) and NCI-H1975 (H1975) cell lines **(B)**.

The readouts were performed at 8 days post-transduction for Calu-6, H838, PC-9 and H1975 and 11 days for H3122 and H1650 cells.

For each biological replicate, data were normalized to the average of negative control signals and plotted as individual values. Bar heights represent mean, error bars represent range. Significance was calculated on $\log_2$-transformed fold change data using two-tailed unpaired Student's t-test with Welch's correction, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

5.2.2 The *EGFR*-mutant LUAD cell line PC-9 was less sensitive to CENPE inhibition by an allosteric inhibitor compared to the *EGFR*-wildtype LUAD cells

Next, I used the specific allosteric inhibitor of CENPE kinesin motor function GSK-923295 to assess whether it exerted differential effects on the cell viability of *EGFR*-wildtype and -mutant LUAD cell lines comparable to the effects observed upon the *CENPE* gene knockdown. The viability of the *EGFR*-wildtype Calu-6 and H838 cell lines was strongly impacted in response to the increasing concentrations of the inhibitor (**Figure 10**), recapitulating the effects observed upon CRISPRi-mediated *CENPE* knockdown. Contrary to the expectations, the *EGFR*-mutant LUAD cell line NCI-H1975 was highly sensitive to the compound starting from the lowest tested concentration of 0.1 µM, while the other tested *EGFR*-mutant cell line PC-9 was the least sensitive to the treatment throughout the range of tested concentrations up until 5 µM, corroborating the gene silencing-based observations (**Figure 10**).

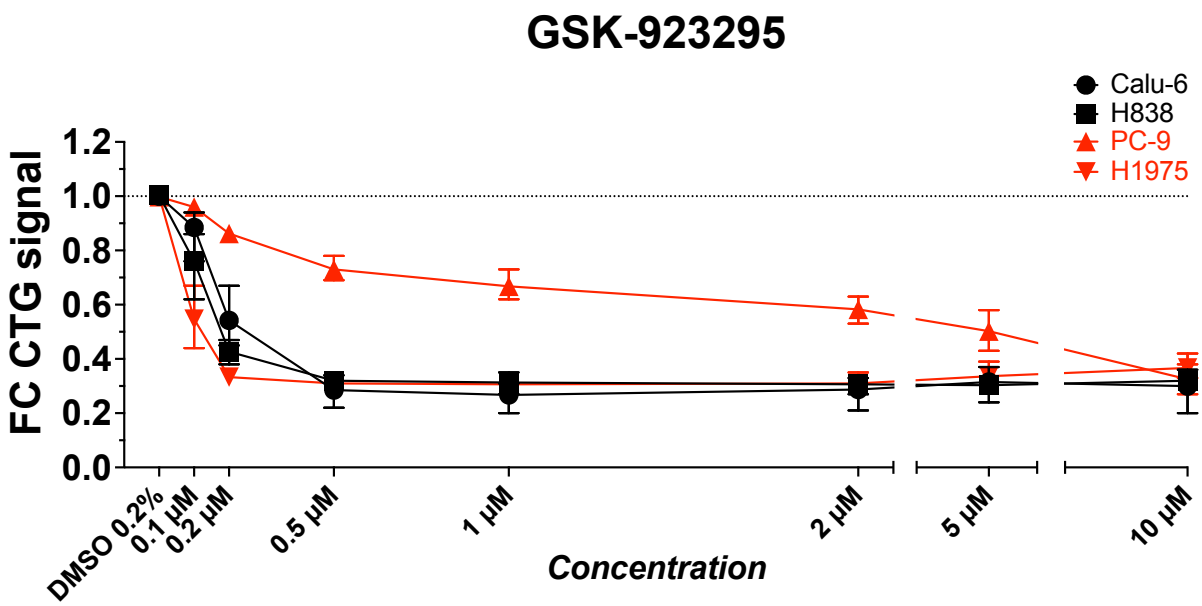


Figure 10. Viability of *EGFR*-wildtype and -mutant LUAD cell lines upon the treatment with CENPE inhibitor.
CellTiter-Glo cell viability assay (CTG) was performed on *EGFR*-wildtype (WT) cell lines Calu-6 and NCI-H838 (H838) and *EGFR*-mutant (Mut) cell lines PC-9 and NCI-H1975 (H1975) treated with 0.1-10 µM GSK-923295 or vehicle solvent (DMSO 0.2%, normalization control) 72 hours post-treatment (n = 3-4).
For each biological replicate, data were normalized to the respective negative control signal. Data are represented as mean per condition; error bars represent range.

5.2.3 Simultaneous depletion of *EGFR* and *CENPE* expression in *EGFR*-mutant cells exerted a synergistic effect on cell viability

To test whether the expression of mutant *EGFR* mediated the resistance of LUAD cell lines to *CENPE* silencing, I performed a simultaneous CRISPRi-based knockdown of *EGFR* and *CENPE* expression in PC-9 cells harboring an oncogenic deletion in exon 19 of the *EGFR* gene, and assayed the resulting cell viability. I used a pool of three individually pre-validated sgRNAs per target to maximize the knockdown efficiency and reduce potential off-targeting. I repeated the experiments in the *EGFR*-wildtype Calu-6 cell line to control for the specificity of the double-knockdown phenotype in the *EGFR*-mutant context.

Gene expression of both *EGFR* and *CENPE* was efficiently suppressed by the CRISPRi system in both cell lines (**Supplementary figure 5**).

In accordance with previous results, knockdown of *CENPE* with the CRISPRi pool only mildly reduced the cell viability of PC-9 cells by 24% and dramatically decreased the viability of Calu-6 cells by 83%. Conversely, *EGFR* knockdown had a strong negative impact on the viability of PC-9 cells and no negative effect on Calu-6 cells, illustrating the dependence of the PC-9 cells on the oncogenic EGFR signaling. Finally, the combination of anti-*CENPE* and anti-*EGFR* sgRNAs showed a weak, albeit significant synergistic negative effect on the viability of PC-9 cells, while the strength of the impact of the double-knockdown on Calu-6 cells was comparable to the effect of *CENPE*-only knockdown (**Figure 11**).

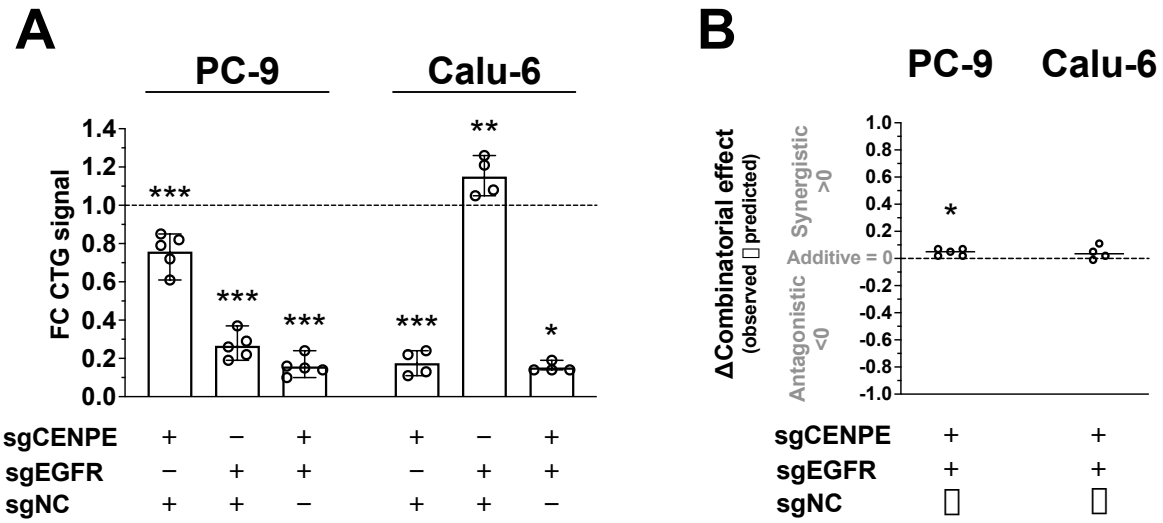


Figure 11. Simultaneous knockdown (KD) of *EGFR* and *CENPE* expression exerted a weak synergistic effect on the cell viability of *EGFR*-mutant PC-9 cells.

EGFR-mutant PC-9 cells and *EGFR*-WT Calu-6 cells were transduced with indicated combinations of pools of three CRISPRi constructs expressing sgRNAs against *CENPE* (sgCENPE), *EGFR* (sgEGFR) and non-targeting negative contol sgRNAs (sgNC). Cells transduced with sgNC were used as a normalization control. Total amount of lentiviral stock used for each transduction was equal across all conditions.

(A) CellTiter-Glo cell viability assay (CTG) was performed 8 days post- transduction (n = 4-5). For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance was calculated on $\log_2$-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

(B) Theoretical additive effect of simultaneous *CENPE* and *EGFR* KD was calculated using the Bliss independence model and subtracted from the observed effect per condition (n = 4-5). Data are plotted as individual values per replicate, lines represent mean, bars represent range. Significance was calculated using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

5.2.4 Silencing of *CENPE* in the context of ectopic overexpression of mutant *EGFR* did not render *EGFR*-wildtype cells resistant to *CENPE* knockdown

To further investigate the link between the expression of mutant *EGFR* and the resistance of LUAD cells to *CENPE* knockdown, I ectopically overexpressed *EGFR* harboring L858R or T790M mutations as well as wildtype *EGFR* in Calu-6 cells, and subsequently silenced *CENPE* expression with CRISPRi. I verified the *CENPE* knockdown and ectopic *EGFR* overexpression by RT-qPCR (**Supplementary figure 6**).

I observed a similar *CENPE* silencing-induced decrease of cell viability in Calu-6 cells in all samples compared to empty vector control (reduction by 81 - 89%) (**Figure 12A**). Neither mutant nor wildtype *EGFR* overexpression had an impact on *CENPE* knockdown-induced loss of Calu-6 cell viability (**Figure 12B**).

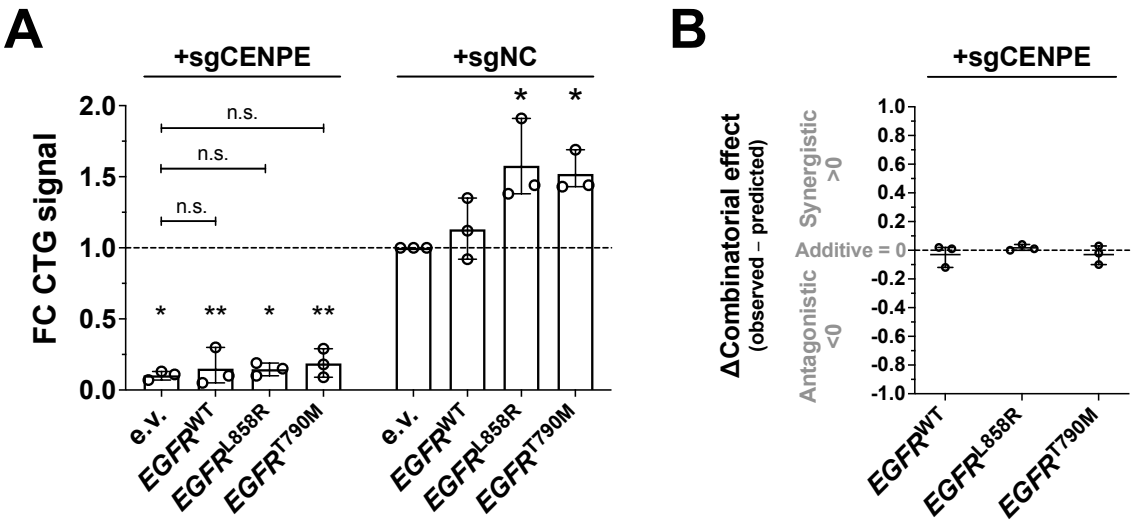


Figure 12. Overexpression (OE) of mutant *EGFR* in *EGFR*-wildtype cells did not render the cells resistant to *CENPE* knockdown (KD).

EGFR-wildtype (WT) Calu-6 cells stably overexpressing empty vector (e.v.), $EGFR^{WT}$, $EGFR^{L858R}$ and $EGFR^{T790M}$ were transduced with a pool of three CRISPRi constructs expressing sgRNAs against *CENPE* (sgCENPE) vs. a pool of three negative control sgRNAs (sgNC). Cells stably transduced with e.v. and sgNC were used as normalization control.

(A) CellTiter-Glo cell viability assay (CTG) was performed 8 days after transduction with the CRIPSRi constructs (n = 3). For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance was calculated on log_2-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

(B) Theoretical additive effects of $EGFR^{WT/L858R/T790M}$ OE and *CENPE* KD were calculated using the Bliss independence model and subtracted from the observed effects per condition (n = 3). Data are plotted as individual values per replicate, lines represent mean, bars represent range. Significance was calculated using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

[This page intentionally left blank]

5.3 *CASP8AP2* as a novel essential viability factor in NSCLC

5.3.1 Validation of the *CASP8AP2* knockdown-mediated phenotype in LUAD cell lines with CRISPRi

To validate the negative effects of *CASP8AP2* depletion on LUAD cell viability, I transduced three of the screened LUAD cell lines with three top ranking sgRNAs against *CASP8AP2* selected from the CRISPRi screening library based on their negative selection scores across all cell lines. All three sgRNAs yielded a high *CASP8AP2* knockdown efficiency, which resulted in a complete loss of cell viability in all three tested LUAD cell lines Calu-6 (decrease in viability over 95%), NCI-H1975 (over 87%) and NCI-H838 (over 95%) (**Figure 13**).

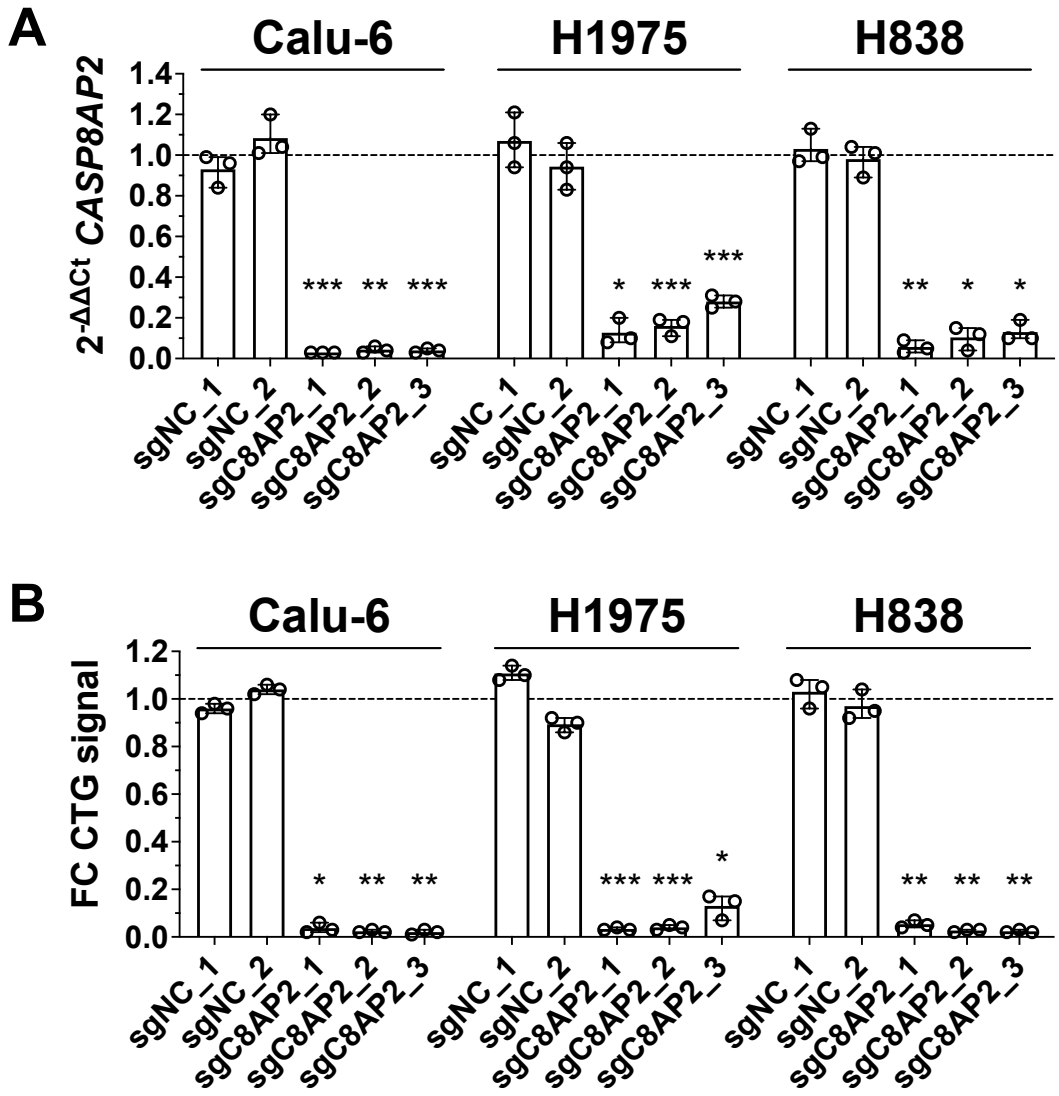


Figure 13. CRISPRi-mediated knockdown (KD) of *CASP8AP2* resulted in a significant decrease in cell viability of LUAD cell lines.

(A) RT-qPCR to confirm the *CASP8AP2* knockdown (KD) in all tested cell lines transduced independently with three CRISPRi constructs expressing sgRNAs against *CASP8AP2* (sgC8AP2) compared to negative control sgRNAs (sgNC). Calu-6 and H838 cells were harvested for RNA isolation 4 days post-transduction, H1975 cells at 5 days post-transduction (n = 3).

(B) CellTiter-Glo cell viability assay (CTG) to assess the phenotype of LUAD cells transduced independently with three sgRNAs against *CASP8AP2* (sgC8AP2) compared to two negative control sgRNAs (sgNC). CTG assays were performed at 8 days post-transduction for Calu-6, 13 days post-transduction for H1975 and 7 days post-transduction for H838 cell line (n = 3).

For each biological replicate, data were normalized to the average of negative control signals and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed on $\log_2$-transformed fold change data using two-tailed unpaired Student's t-test with Welch's correction, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

5.3.2 RNAi-mediated knockdown revealed differential responses to *CASP8AP2* silencing between non-transformed and malignant lung cells

Due to the laboriousness of the CRISPRi-based knockdown protocol, I switched to siPOOLs for RNAi-based validation of the phenotypic effects of the *CASP8AP2* depletion in an expanded cell line panel (**Figure 14**).

In all tested cell lines, the efficiency of *CASP8AP2* knockdown was comparable (between 62% and 82%; **Figure 14A**). siPOOLs-mediated *CASP8AP2* silencing recapitulated the complete loss of cell viability in NCI-H1975 and NCI-H838 cells observed after the CRISPRi-based knockdown (decrease in cell viability by 94% and 88%, respectively) despite lower knockdown efficiencies. Depletion of *CASP8AP2* in PC-9 also resulted in a complete loss of cell viability (by 88%), and in Calu-6 cells, the viability was strongly decreased (by 69%) (**Figure 14B**).

Next, I tested three additional NSCLC cell lines NCI-H460 (LCLC), NCI-H1299 and NCI-H1563 (both LUAD) for a *CASP8AP2* silencing-mediated phenotype. Notably, NCI-H1563 cells exhibited only a slight decrease in cell viability by 13% compared to control. NCI-H460 cells were highly sensitive to the loss of *CASP8AP2* expression, exhibiting a 96% decrease in cell viability, and NCI-H1299 lost 64% of cell viability.

Further, I assayed two non-transformed lung fibroblast cell lines, IMR-90 and WI-38, and found a moderate decrease by 35% and 40% in cell viability upon *CASP8AP2* knockdown, respectively.

The effect of *CASP8AP2* silencing on the cell viability in the assayed NSCLC cells (except NCI-H1563) was significantly stronger than in the non-transformed lung cell lines. Thus, I concluded that non-transformed lung cell lines IMR-90 and WI-38 as well as the LUAD cell line NCI-H1563 were tolerant to *CASP8AP2* depletion compared to the remaining NSCLC cell lines. For simplicity, I termed the two cell line groups “tolerant” and “sensitive” based on their response to *CASP8AP2* knockdown. I further set out to investigate molecular mechanisms underlying the observed differential responses.

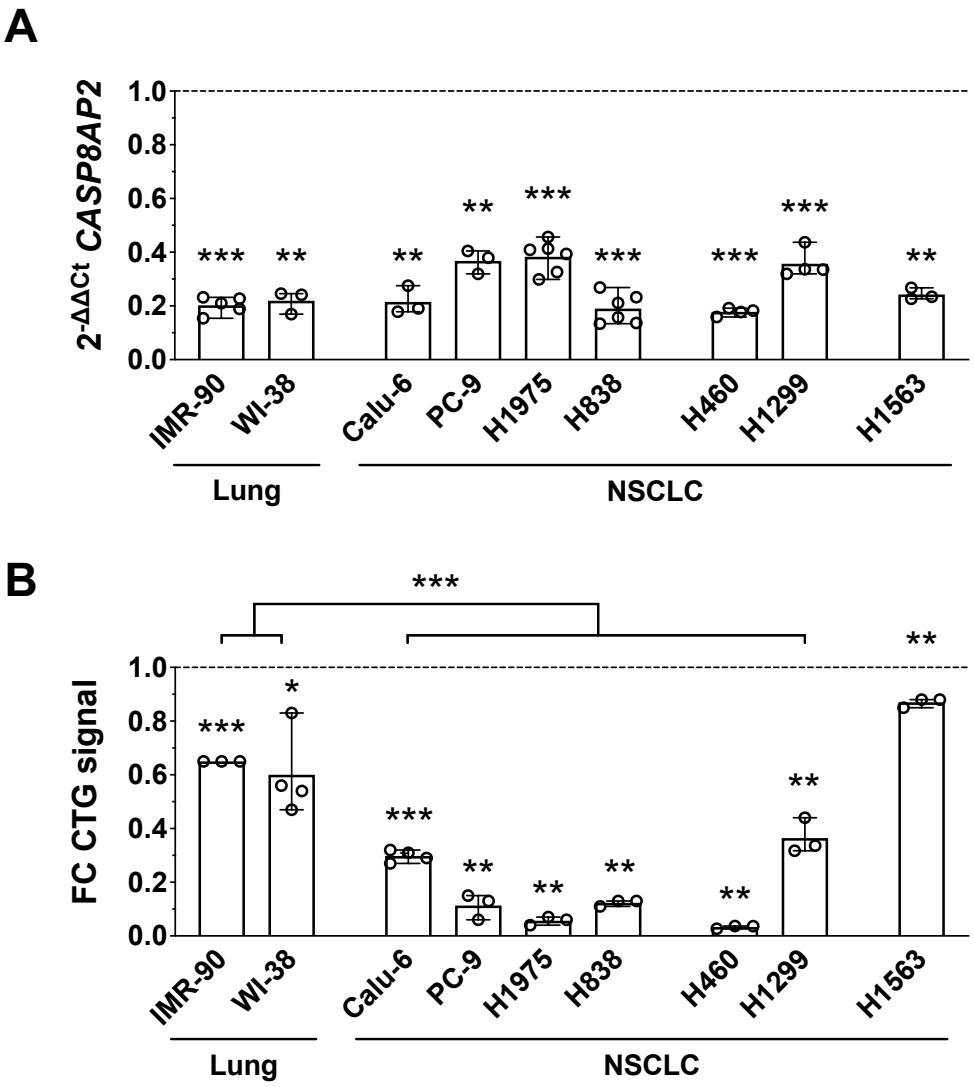


Figure 14. RNAi-based knockdown revealed differential responses to *CASP8AP2* silencing between non-transformed lung and NSCLC cell lines.

(A) RT-qPCR to confirm siPOOL-mediated *CASP8AP2* knockdown (siC8AP2, 10 nM) in all tested cell lines transfected with siC8AP2 compared to respective cell lines transfected with non-targeting negative control siPOOL (siNC, 10 nM) at 48 hours post-transfection (n = 3-6).

(B) CellTiter-Glo cell viability assay (CTG) to assess the effect of *CASP8AP2* siPOOL-mediated knockdown (siC8AP2, 10 nM) on non-transformed lung fibroblast cell lines IMR-90 and WI-38, NSCLC LUAD cell lines Calu-6, PC-9, NCI-H1975 (H1975), NCI-H838 (H838), NCI-H1299 (H1299) and NCI-H1563 (H1563), and NSCLC LCLC cell line NCI-H460 (H460) compared to respective cell lines transfected with non-targeting negative control siPOOL (siNC, 10 nM) at 72 hours post-transfection (n = 3-4).

For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed on $\log_2$-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$. Significance testing between normal and NSCLC cell lines sensitive to *CASP8AP2* silencing was performed using two-tailed unpaired Student's t-test with Welch's correction.

5.3.3 *CASP8AP2* depletion induced activation of apoptotic signaling in cancer, but not in non-transformed lung cells

Since I primarily identified *CASP8AP2* as a cell viability factor, I further examined its effects on the apoptotic signaling. The two main cellular apoptosis pathways are caspase-8-regulated extrinsic (death receptor-mediated) and caspase-9-regulated intrinsic (mitochondrial) cascades. Both extrinsic and intrinsic apoptotic pathways rely on the activation of downstream effector caspases-3 and -7 [225].

I assayed the activation of effector caspases-3/7, initiator caspase-8 and initiator caspase-9 in tolerant cell lines IMR-90 and NCI-H1563 and sensitive cell lines NCI-H1975 and NCI-H838 upon *CASP8AP2* knockdown (**Figure 15A**).

There was no detectable activation of caspases in IMR-90. However, the tolerant LUAD cell line NCI-H1563 showed significant activation of all tested caspases. Sensitive LUAD cell lines also exhibited significant activation of all caspases upon *CASP8AP2* silencing. In NCI-H1563 cells, the level of caspase-8 activation was lower compared to caspases-3/7 and caspase-9, while in the sensitive cells the level of activation of all assayed caspases was comparable within a cell line at the time of readout.

To assess the contribution of different caspase cascades to the *CASP8AP2* depletion-induced phenotype in sensitive cells, I performed a simultaneous knockdown of *CASP8AP2* in combination with caspases-3/7 (*CASP3* and *CASP7*), caspase-8 (*CASP8*) or caspase-9 (*CASP9*) and assayed the resulting cell viability in the NCI-H1975 and NCI-H838 cell lines (**Figure 15B**). Neither of the combinations resulted in a rescue of the *CASP8AP2* knockdown-induced loss of cell viability (see **Supplementary figure 6** and **Supplementary figure 7** for target knockdown validations by RT-qPCR).

Next, I used a pan-caspase inhibitor z-VAD-FMK to assess the general role of caspase-mediated cell death pathways in the *CASP8AP2* knockdown-induced loss of cell viability phenotype in NCI-H1975 cells (**Figure 16**). At 5 μM of the inhibitor, I observed an increase in the fraction of viable cells upon *CASP8AP2* knockdown from 9% to 30% of control. The fraction of viable cells plateaued with the further increase of z-VAD-FMK concentration, reaching 41% at 50 μM.

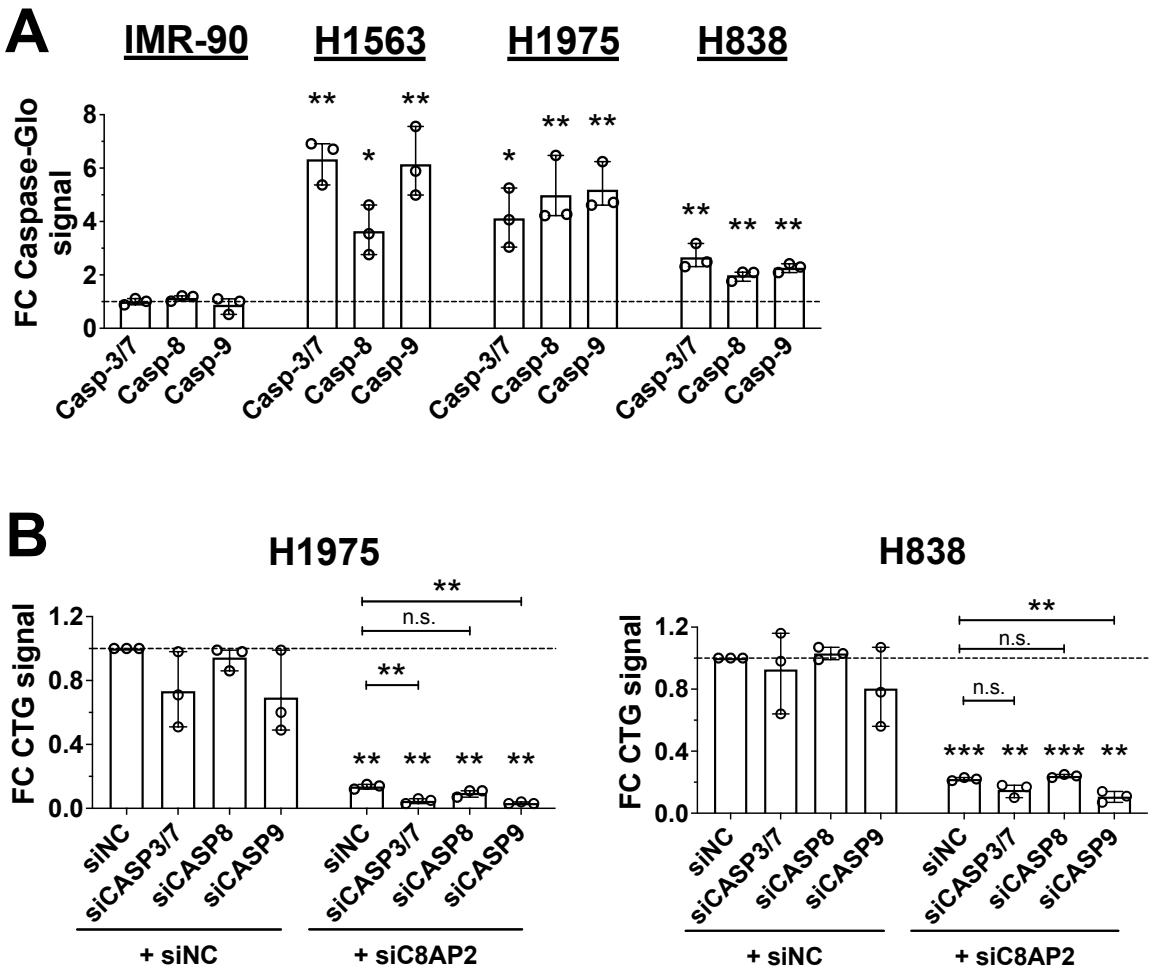


Figure 15. Loss of *CASP8AP2* expression induced apoptotic cascades in LUAD, but not in non-transformed cells.

(A) Caspase-Glo luminescent apoptosis assays to assess the activation of caspases-3 and -7 (Casp-3/7), caspase-8 (Casp-8) and caspase-9 (Casp-9) in IMR-90 and H1563 cells tolerant to *CASP8AP2* depletion and H1975 and H838 cells sensitive to *CASP8AP2* depletion transfected with siPOOL against *CASP8AP2* (siC8AP2, 10 nM) compared to respective cell lines transfected with negative control siPOOL (siNC, 10 nM) at 48 hours post-transfection (n = 3).

(B) CellTiter-Glo cell viability assay (CTG) to assess the effect of siPOOLs-mediated knockdown of *CASP3* and *CASP7* with a combined siPOOL (siCASP3/7), *CASP8* with an individual siPOOL (siCASP8) and *CASP9* with an individual siPOOL (siCASP9) on siC8AP2-mediated phenotype in H1975 and H838 cell lines compared to respective cell lines transfected with siNC at 96 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3). For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed on $\log_2$-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

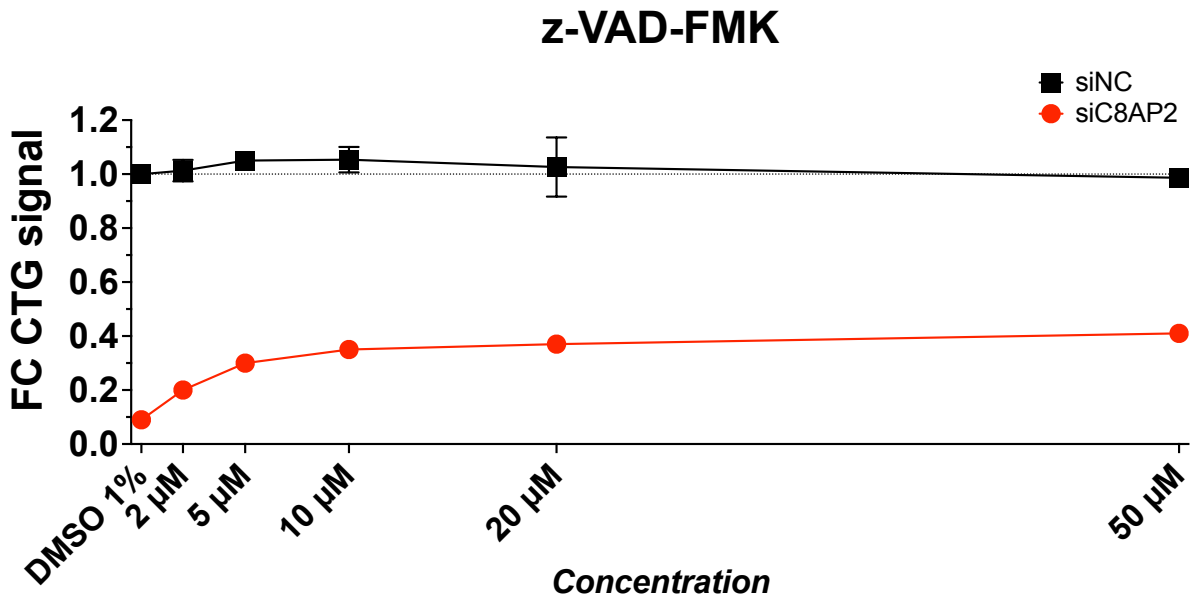


Figure 16. Pan-caspase inhibitor z-VAD-FMK partially rescued *CASP8AP2* silencing-induced loss of cell viability in the NCI-H1975 cell line.

CTG to evaluate the effect of z-VAD-FMK treatment on siC8AP2-transfected H1975 cells compared to siNC (10 nM). Cells were treated with 2-50 µM z-VAD-FMK or vehicle solvent (DMSO 1%) 90 minutes before reverse-transfection with siPOOLs and analyzed after 72h. Cells treated with vehicle solvent and transfected with siNC were used as a normalization control (n = 3). For each biological replicate, data were normalized to the respective negative control signal. Data are represented as mean per condition, error bars represent range.

5.3.4 *CASP8AP2* knockdown induced S-phase cell cycle arrest in both *CASP8AP2* depletion-tolerant and -sensitive cell lines

To further characterize the observed phenotype, I investigated the effects of *CASP8AP2* depletion on the cell cycle progression in sensitive and tolerant cell lines (**Figure 17**).

A BrdU incorporation assay showed a complete loss of a BrdU-positive cell population upon *CASP8AP2* knockdown in non-transformed IMR-90 cells (**Figure 17A, *left***). In the tolerant NSCLC cell line NCI-H1563 (**Figure 17B, *left***), as well as in the sensitive NCI-H1975 and NCI-H838 cells (**Figure 17C-D, *left***), I detected an accumulation of cells with decreased BrdU incorporation rates within the S-phase. Notably, in NCI-H1563 and NCI-H838 cells, the BrdU-positive signals accumulated towards the end of the S-phase while in NCI-H1975 the signal accumulation was most pronounced at the beginning of the S-phase.

Propidium iodide (PI) staining identified a cell cycle arrest peak at the beginning of the S-phase in NCI-H1975 cells, an early S-phase arrest peak in IMR-90 cells, a mid/late S-phase arrest peak in NCI-H838 cells and a late S-phase arrest peak in NCI-H1563 cells, recapitulating the results of the BrdU staining (**Figure 17, *middle***). Quantification of the PI signals revealed a significant decrease of cell count in G1- and a significant increase in S-phase of the cell cycle in all tested cell lines (**Figure 17, *left***).

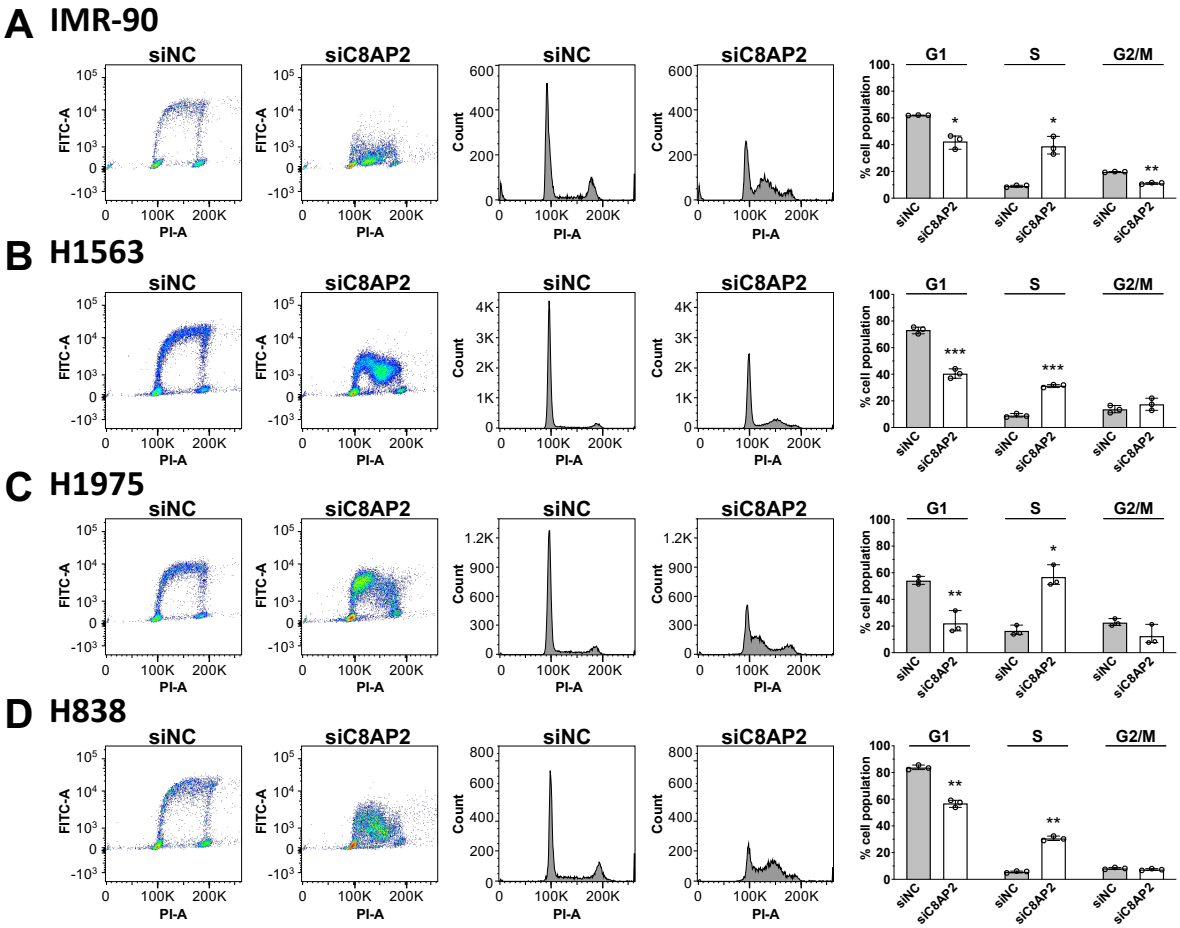


Figure 17. Loss of *CASP8AP2* induced S-phase cell cycle disruption in both tolerant and sensitive cell lines.

IMR-90 cells **(A)**, NCI-H1563 **(**H1563, **B)**, NCI-H1975 **(**H1975, **C)** and NCI-H838 **(**H838, **D)** cells were pulsed with BrdU *(left panel)* at 46 hours post-transfection with the siPOOL against *CASP8AP2* (siC8AP2, 10 nM) or the negative control siPOOL (siNC, 10 nM). Cells were harvested after 2 hours (48 hours post-transfection), fixed, stained with propidium iodide (PI) (*middle panel, quantified in the right panel)* and analyzed by flow cytometry. Single cell-derived signals were gated, plotted and quantified using FlowJo v10 software.

Quantified data are represented as individual values per biological replicate, bar heights represent mean, error bars represent range. Significance testing was performed using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

5.3.5 Systematic RNA and protein expression profiling identified *JUNB* as a differentially regulated factor between *CASP8AP2* depletion-tolerant and -sensitive cell lines

Next, I performed a systematic expression profiling of tolerant and sensitive cell lines to identify differentially regulated factors underlying their divergent responses to the loss of *CASP8AP2*.

First, I performed RNA sequencing (RNA-Seq) of IMR-90, NCI-H1975 and NCI-H838 cells at 48 hours post-transfection with the siPOOL targeting *CASP8AP2* as well as the respective cells transfected with the negative control non-targeting siPOOL. Using the DESeq2 analysis pipeline, I detected genes that were differentially expressed in each tested cell line upon *CASP8AP2* depletion compared to the negative control. Using the DESeq2 output, I identified genes that were divergently regulated upon *CASP8AP2* knockdown between the tolerant IMR-90 and the sensitive NCI-H1975 and NCI-H838 cell lines. I hypothesized that factors responsible for the *CASP8AP2* depletion-induced phenotype would be overactivated in the sensitive cell lines and therefore selected a subset of 401 genes that were significantly upregulated in both NCI-H1975 and NCI-H838 cell lines ($\log_2(FC) > 0$, $p_{adj} < 0.05$) while exhibiting a downregulation trend in IMR-90 cells ($\log_2(FC) < 0$ or NA) for further analysis. Differential expression of 30 out of 401 genes significantly ($p_{adj} < 0.05$) and divergently regulated between IMR-90 ($\log_2(FC) < 0$) and the LUAD cell lines ($\log_2(FC) > 0$) upon *CASP8AP2* silencing is depicted in **Supplementary figure 16**.

Next, I analyzed the lung cell line RNA expression dataset from the Diederichs lab [205] to identify genes with differential steady state expression levels between the tolerant IMR-90 and NCI-H1563 and the sensitive NCI-H1975 and NCI-H838 cell lines. I hypothesized that factors responsible for the divergent response to the *CASP8AP2* depletion would be underrepresented in the tolerant cells. In total, 5986 genes were underrepresented ($FC < 0.5$) at the RNA level in both IMR-90 and NCI-H1563 cell lines compared to the sensitive cell lines and taken for further analysis.

Additionally, I examined lung cell line mass spectrometry data from the Diederichs lab [205] to select for proteins that could not be detected in the tolerant cell line NCI-H1563 ($\log_{10}$(normalized intensities) = NA) but were detectably expressed in sensitive cell lines NCI-H1975 and NCI-H838 in steady state. In total, 826 proteins were selected for further analysis.

Next, the three selected datasets were merged to examine the overlap (**Figure 18A**). As a result, *JUNB* and *WNK2* were identified as genes differentially regulated between the tolerant and sensitive cell lines upon *CASP8AP2* depletion as well as in steady state at both RNA and protein expression levels.

JUNB was significantly induced in NCI-H1975 and NCI-H838 cells and weakly, albeit significantly, downregulated in the DESeq2-derived dataset upon *CASP8P2* knockdown in IMR-90 (FC of 1.6, 2.9 and 0.79, respectively). *WNK2* was also significantly upregulated in the sensitive cell lines and showed no change in expression in IMR-90 cells (FC of 3.7 for NCI-H1975, 2.3 for NCI-H838 and 0.97 for IMR-90). Notably, the DESeq2 algorithm could not perform statistical analysis of the regulation of *WNK2* in IMR-90 due to the lack of corresponding RNA-sequencing read counts. Moreover, according to the RNA expression dataset, *JUNB* was detectably expressed in steady state in all studied cell lines, while total expression levels of *WNK2* in all cell lines were under the detectable expression threshold of FPKM = 1 (**Figure 18B**). Therefore, I decided to further investigate the role of *JUNB* in the *CASP8AP2* depletion-induced phenotype.

Following this discovery, I assayed the *JUNB* mRNA expression upon siPOOLs-mediated *CASP8AP2* knockdown in the tolerant cell lines IMR-90 and NCI-H1563 and the sensitive cell lines NCI-H1975, NCI-H838, NCI-H460 and NCI-H1299. I found no significant differences in *JUNB* mRNA levels in the non-transformed lung cell line IMR-90 and a significant increase in *JUNB* mRNA levels in the tolerant NCI-H1563 as well as in all tested sensitive cell lines upon *CASP8AP2* knockdown (**Figure 18C**).

Since JUNB is a subunit of the AP-1 transcription factor, I proceeded to assess the effects of AP-1 inhibition on the cell viability of the sensitive cell lines upon *CASP8AP2* depletion.

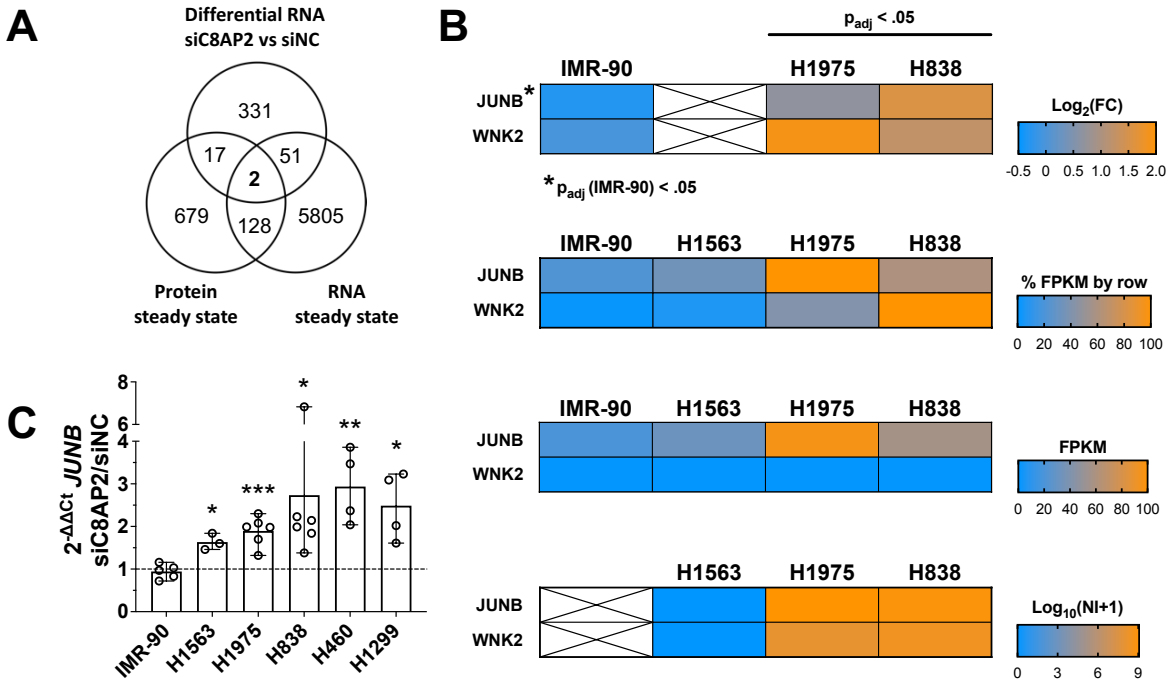


Figure 18. *JUNB* is differentially regulated between *CASP8AP2* depletion-tolerant and -sensitive cell lines.

FPKM = fragments per kb per million, NI = normalized intensities.

(A) Venn diagram representing the overlap between genes and proteins differentially expressed between *CASP8AP2* depletion-tolerant (IMR-90 and H1563) and -sensitive (H1975 and H838) cell lines in steady-state as well as upon *CASP8AP2* silencing.

"Differential RNA" dataset contains genes that were significantly upregulated ($log_2(FC) > 0$, $p_{adj} < 0.05$) by RNA-Seq in both H1975 and H838 cell lines and not upregulated ($log_2(FC) < 0$ or NA) in IMR-90 cells upon siPOOL-mediated *CASP8AP2* (siC8AP2, 10 nM) knockdown (KD) compared to respective cell lines transfected with negative control siPOOL (siNC, 10 nM) at 48 hours post-transfection (n = 3). $Log_2(FC)$ and p_{adj} were calculated using DESeq2 algorithm.

"RNA steady state" dataset contains genes with low expression in both IMR-90 and H1563 compared to mean expression between H1975 and H838 cell lines in steady state (FC < 0.5).

"Protein steady state" dataset contains proteins with no detectable expression in H1563 and detectable expression levels in both NCI-H1975 and NCI-H838 cells in steady state.

(B) ***(Top)*** heatmap containing $log_2(FC)$ in RNA expression of the two genes overlapping between the datasets in **(A)** in IMR-90, H1975 and H838 cell lines upon siC8AP2 transfection vs. siNC. ***(Middle)*** two heatmaps containing row-normalized FPKM values (100% assigned to the maximal value per comparison) and raw FPKM values, respectively, of the two genes overlapping between the three datasets in **(A)**, and ***(bottom)*** heatmap containing the log_{10}(normalized intensities+1) protein expression values of the corresponding proteins in IMR-90, H1563, H1975 and H838 cell lines. White crossed out cells represent conditions where data were not available.

(C) RT-qPCR to assess mRNA levels of *JUNB* in IMR-90, H1563, H1975, H838, H460 and H1299 cell lines 48 hours post-transfection with siPOOL against *CASP8AP2* (siC8AP2, 10 nM) compared to negative control siPOOL (siNC, 10 nM) (n = 3-6). For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed calculated on log_2-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

5.3.6 AP-1 inhibition rescued the *CASP8AP2* depletion-induced phenotype in sensitive lung cancer cell lines

I treated the sensitive cell lines NCI-H1975, NCI-H838, NCI-H460 and NCI-H1299 with the specific inhibitor of AP-1 DNA binding activity T-5224 [226–229] and compared the cell viability response between the treated and untreated cells to *CASP8AP2* knockdown. The cell viability upon *CASP8AP2* silencing was significantly restored in all T5224-treated cell lines (**Figure 19A**).

Quantification of the observed effects revealed that AP-1 inhibition had a significant antagonistic impact on the *CASP8AP2* knockdown-induced loss of cell viability in all examined sensitive cell lines (**Figure 19B**).

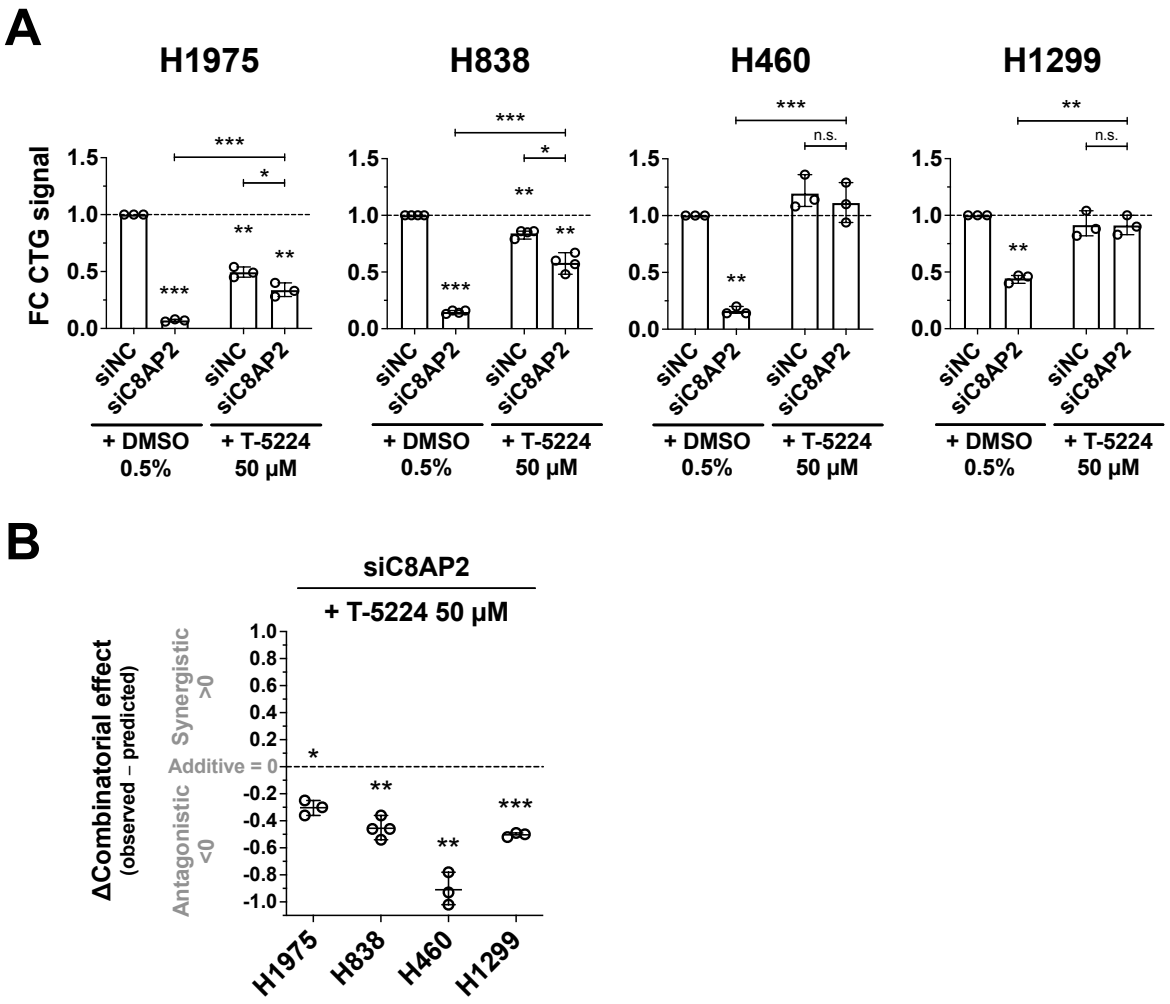


Figure 19. AP-1 inhibition rescued *CASP8AP2* depletion-induced phenotype.
(A) CTG to assess the effect of T-5224 treatment on siC8AP2-transfected NCI-H1975 (H1975), NCI-H838 (H838), NCI-H460 (H460) and NCI-H1229 (H1299) cells compared to respective cell lines transfected with negative control siPOOL (siNC, 10 nM per siPOOL). Cells were treated with 50 µM T-5224 or vehicle solvent (DMSO 0.5 %) 90 minutes before reverse-transfection with siPOOLs and analyzed after 72 hours. Cells treated with vehicle solvent and transfected with siNC were used as a normalization control (n = 3-4). For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed on $\log_2$-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.
(B) Theoretical additive effect of siPOOLs-mediated *CASP8AP2* knockdown (KD) and T-5224 treatment was calculated using the Bliss independence model and subtracted from the observed effect per condition (n = 3-4). Data are plotted as individual values per replicate, lines represent mean, bars represent range. Significance was calculated using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

5.3.7 *CDKN1A* induction upon *CASP8AP2* knockdown was regulated by p53 in non-transformed lung cells and by AP-1 in NSCLC independently of p53

During the course of my experiments, I found that *CASP8AP2* knockdown resulted in elevated mRNA levels of *CDKN1A* (p21) in both tolerant and sensitive cell lines, including the p53-null sensitive cell line NCI-H1299 and p53-deficient cell line NCI-H838 (**Table 17**, **Figure 20A**). I also discovered that simultaneous silencing of *CASP8AP2* and *TP53* resulted in the loss of *CDKN1A* expression in the non-transformed lung cell line IMR-90, but not in the tested p53-positive NSCLC cell lines. Moreover, while in NCI-H1563 and NCI-H460 cells knockdown of *TP53* partially attenuated the induction of *CDKN1A* upon *CASP8AP2* silencing, in NCI-H1975 the levels of *CDKN1A* mRNA were further elevated (**Figure 20B**). Notably, treatment of both p53-positive and p53-negative sensitive NSCLC cell lines with the AP-1 inhibitor attenuated the *CDKN1A* induction (**Figure 21**).

Table 17. *p53 status of NSCLC cell lines according to COSMIC database*

Cell line	p53 status	Mutation in *TP53* CDS	Mutation in p53 aa sequence	Mutation type	Mutation zygosity
H1563	WT	-	-	-	-
H1975	Mut	c.818G>A	p.R273H	Missense	homo
H838	Deficient	c.184G>T	p.E62*	Nonsense	homo
H460	WT	-	-	-	-
H1299	Null	c.1_954>AAG	p.?	Partial gene deletion	homo

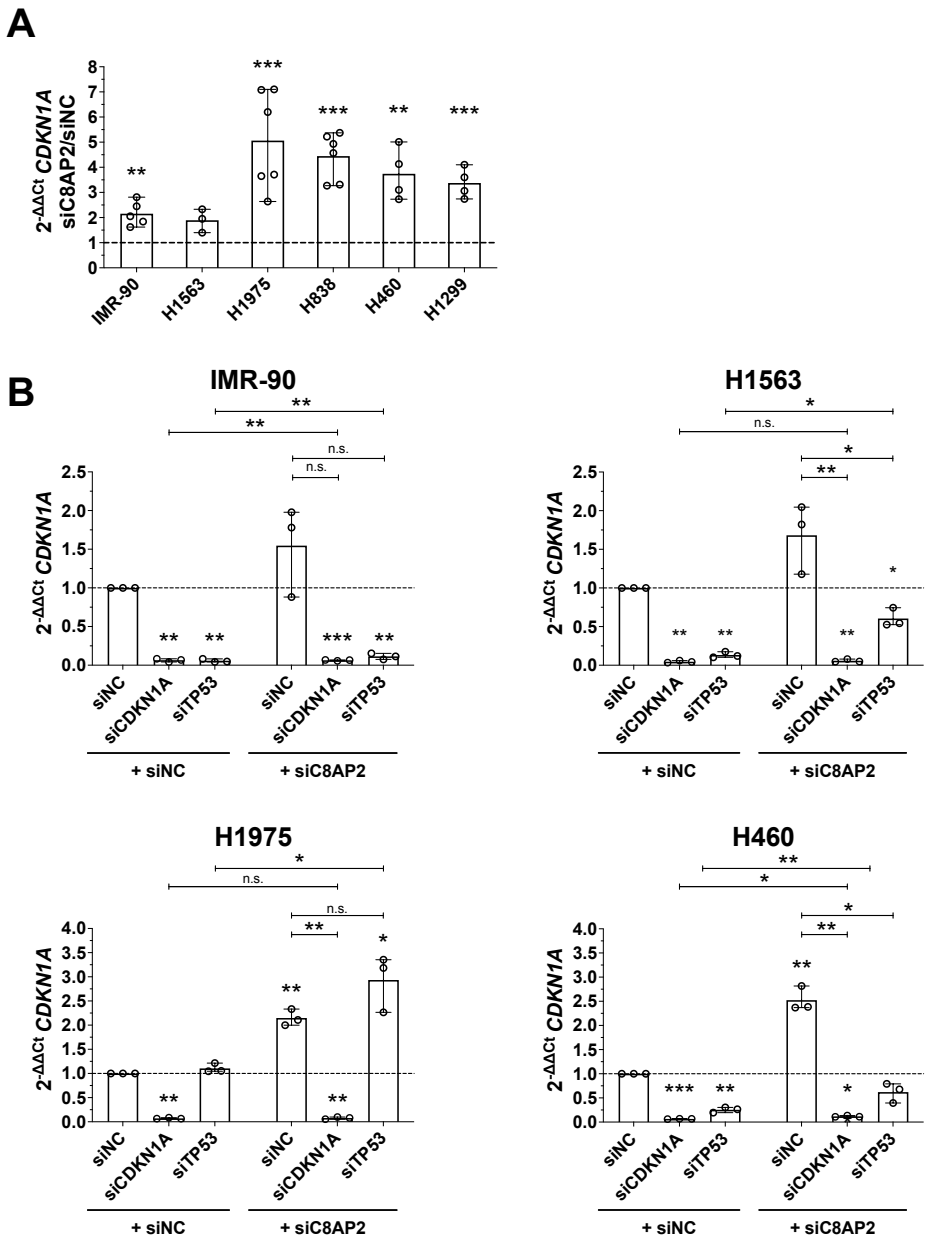


Figure 20. *CASP8AP2* knockdown (KD) induced *CDKN1A* expression in a p53-independent manner in NSCLC.

(A) RT-qPCR to assess mRNA levels of *CDKN1A* in *CASP8AP2* loss-tolerant cell lines IMR-90 and H1563 and -sensitive cell lines H1975, H838, H460 and H1299 at 48 hours post-transfection with siPOOL against *CASP8AP2* (siC8AP2, 10 nM) and negative control siPOOL (siNC, 10 nM) (n = 3-6). siNC-transfected cells were used as normalization control.

(B) RT-qPCR to assess the mRNA levels of *CDKN1A* in IMR-90, H1563, H1975 and H838 cells transfected with indicated combinations of siC8AP2, siCDKN1A, siTP53 and siNC at 48 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3). Cells transfected with siNC (20 nM) were used as normalization control.

For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed calculated on log_2-transformed fold change data using two-tailed paired Student‘s t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; “n.s.”, not significant.

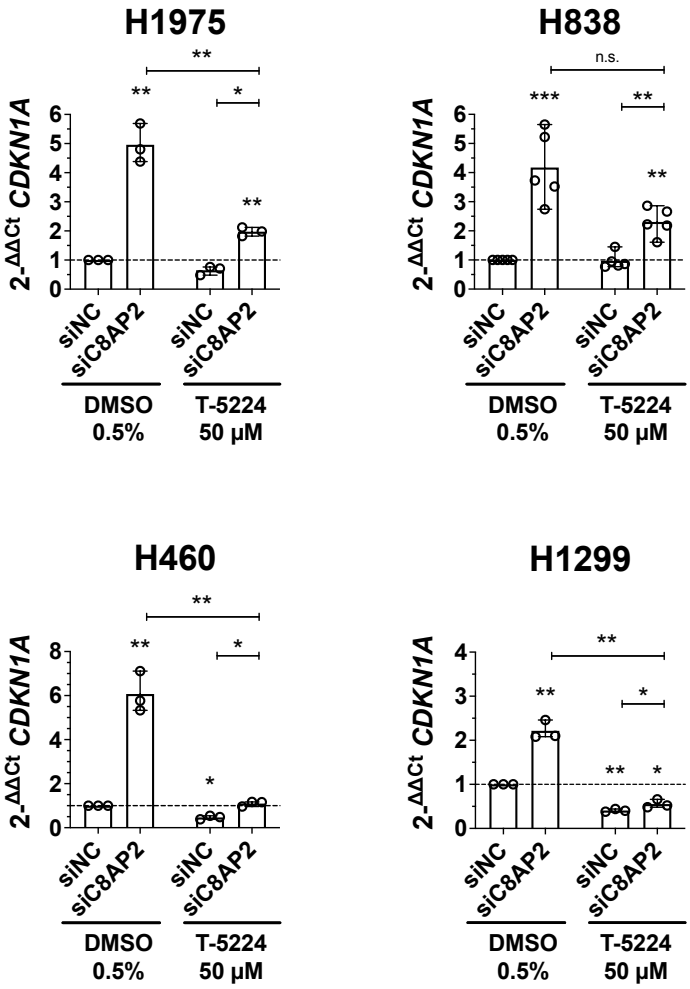


Figure 21. AP-1 regulates *CDKN1A* mRNA expression in the context of *CASP8AP2* loss.
RT-qPCR to assess the levels of *CDKN1A* mRNA in H1975, H838, H460 and H1229 cells transfected with siPOOLs against *CASP8AP2* (siC8AP2) compared to respective cell lines transfected with negative control siPOOL (siNC, 10 nM per siPOOL). Cells were treated with 50 µM T-5224 or vehicle solvent (DMSO 0.5 %) 90 minutes before reverse-transfection with siPOOLsand lysed for RNA isolation after 48 hours. Cells treated with vehicle solvent and transfected with siNC were used as a normalization control (n = 3-5).
For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance was calculated on $\log_2$-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

All in all, I identified *CASP8AP2* as a novel NSCLC viability factor that acted specifically at cancer cells. My results suggested that sensitivity to *CASP8AP2* depletion was mediated by the activity of AP-1 transcription factor and depended on the steady state expression level of *JUNB,* which could explain the divergent responses to *CASP8AP2* knockdown between cancer and non-transformed cell lines. Given that a wide panel of NSCLC cell lines exhibited substantial sensitivity to the loss of *CASP8AP2* expression, my findings indicated a promising therapeutic potential of CASP8AP2 inhibition for lung cancer treatment.

6 Discussion

6.1 Custom CRISPRi dropout screen for novel LUAD viability factors

The primary aim of my PhD project was to design and conduct a CRISPRi negative selection viability screen for general or genotype-specific lung adenocarcinoma (LUAD) vulnerabilities. To achieve that, I first developed a workflow for the screen target selection to identify putative synthetic lethality partners of genes harboring oncogenic mutations commonly detected in LUAD. I hypothesized that genes that formed synthetic lethal relationships with mutated LUAD driver genes would be overexpressed in LUAD patient samples harboring the respective mutations compared to LUAD samples with wildtype alleles of the interrogated genes, and designed an extensive pooled sgRNA library to silence those genes using the CRISPRi targeted transcriptional repression system. In order to expand the pool of novel putative cancer viability candidates, I also targeted lncRNAs overexpressed in LUAD compared to normal lung patient samples.

Since the comparison between *EGFR*-mutant and *EGFR*-wildtype LUAD patient samples yielded the highest number of differentially expressed genes, I decided to set up the screen to facilitate identification of synthetic lethal partners of mutant *EGFR.* To model a robust system, I performed the screen in four LUAD cell lines expressing activating *EGFR* mutations and four *EGFR*-wildtype LUAD cell lines. I expected that depletion of the factors that form synthetic lethal relationships with mutant *EGFR* would selectively affect the viability of *EGFR*-mutant, but not of wildtype cells.

Due to the lack of sufficient CRISPRi-tailored sgRNA design algorithms available at the beginning of my project, I developed my own pipeline according to the state-of-art characterization of the CRISPRi system by Gilbert *et al.* [164]. Since the dCas9-KRAB was shown to effectively repress transcription when targeted within a 1 kb window around a TSS [164], I mapped the off-target sites to a genome-wide list of 2 kb-long sequences surrounding all TSSs annotated in the human genome, and utilized such “TSS-wide” mapping rather than genome-wide off-target mapping as a primary off-target filter. TSS-wide off-target filtering allowed me to discard CRISPRi-irrelevant off-target sites located far away from annotated TSSs and thus provided a higher number of successful CRISPRi-tailored sgRNA designs than the genome-wide approach. However, since many lncRNA TSSs localize in close proximity to their neighbor gene TSSs [147], a limited number of TSS-wide off-targets was allowed in the

algorithm to maintain the lncRNA targets in the screen. Thus, the screen hits targeted by sgRNAs with TSS-wide off-target score of > 0 had to be manually curated after the screen performance-based hit selection process to account for potential neighbor gene off-target effects.

Since I set out to perform the screen in a panel of eight LUAD cell lines with a single CRISPRi library, I resorted to targeting every TSS annotated within each gene of interest. Such setup was chosen to allow me to (i) flexibly adjust the driver mutation context by adding more cell lines, (ii) target potential cell line-specific transcript isoforms originating from alternative TSSs and (iii) aim to distinguish the effect of different isoforms of the same gene.

CRISPRi-specific sgRNA libraries designed by the pioneering work of Horlbeck *et al*. [230], made available through Addgene and commonly employed in recently published studies [231–234], target a maximum of two TSSs/gene that were selected based on merged Cap-Analysis of Gene Expression (CAGE) data from over 800 human tissues and cell lines from the FANTOM5 database and TSS annotations from Ensembl. The resulting protein-coding genome-wide library of 10 sgRNAs/gene provides a useful tool for CRISPRi screening studies, which, however, should be used with caution across different experimental models, as cell type-specific features, e.g., nucleosome positioning and TSS usage, generally determine the CRISPRi system efficiency [230, 235, 236].

Limiting the number of target TSSs per gene allows to optimize the sgRNA/TSS rate and maintain experimentally practical library size in large-scale or genome-wide screens. As my aim was to target a focused subset of genes in an expandable panel of LUAD cell lines, I chose an opposite approach, maximizing the number of target TSSs and saturating the number of sgRNAs per target, thereby augmenting the chances of efficient target gene knockdown across all cell lines.

Two common strategies to design a gene repression experiment are to silence the most expressed transcript or the transcript originating from the annotated gene start. To find out whether targeting each TSS had any advantage compared to these conventional approaches, I tested whether the RNA expression level or the genomic position of the target TSSs could determine its performance in the screen.

In order to evaluate the level of transcriptional activity at a particular TSS, I utilized the FPKM expression value of its immediate downstream exon derived from the lung cell line RNA

expression dataset from the Diederichs lab [205]. My results indicated that TSSs corresponding to RNA expression levels higher than the empirical FPKM threshold of 1 had substantially greater chances of being selected as hits in the CRISPRi dropout viability screen. This observation suggested that TSSs expressed at levels above the set threshold were more likely to be identified as functionally relevant, justifying the use of the expression cutoff during the screen target selection process.

However, remarkably, TSSs exhibiting the highest expression within multi-TSS genes, i.e., genes with multiple annotated TSSs, were selected at a rate expected by chance, suggesting that the comparative RNA expression levels of individual gene isoforms originating from different TSSs were not predictive for biological significance. Corroborating these observations, a detailed examination of the screen hits obtained from the Calu-6 cell line demonstrated that almost half of the TSSs selected by the screen exhibited at least two-fold lower expression compared to the TSS giving rise to the strongest expression within the same multi-TSS gene. These observations suggest that RNA expression is not a good predictor of biological relevance of different gene isoforms.

Interestingly, genomic position analysis revealed that TSSs originating from the annotated gene starts were enriched among the screen hits across all cell lines. In the Calu-6 cells, more than half of the selected TSSs were first by genomic position within the respective multi-TSS genes. Such result could possibly indicate the general functional prevalence of full-length gene isoforms. However, comprehensive analysis of the individual Calu-6 screen hits suggested that my screen setup allowed to differentiate between functional and non-functional transcripts irrespective of their relative genomic position or expression level, provided they can be differentially targeted by the CRISPRi system. Overall, the multiple examples of TSSs selected as screen hits neither being first in genomic position nor being the highest expressed ones corroborates the usefulness of my screen design covering all TSSs in the targeted genes.

All in all, my results proposed that genomic position at the gene start was a more reliable prognostic factor for the performance of the TSSs in the CRISPRi viability screen than their relative expression, however, neither of them should be unambiguously utilized to predict functional transcripts within a gene, providing a word of caution for future gene silencing experiments. It is important to note that the downstream exon-based evaluation used as a proxy for the transcriptional activity did not necessarily reflect the rate of

transcription at a particular TSS, as the RNA-Seq-derived expression value assigned to a particular exon represented a cumulative expression level at the TSS of interest as well as at all upstream TSSs that gave rise to transcripts spanning the inspected exon. Thus, profiling of nascent transcription in the LUAD cell lines would allow drawing more detailed conclusions.

Additionally, the observed differential selection of TSSs within a single gene indicated that the screen setup allowed to distinguish between functional and non-functional gene isoforms. A robust validation of the TSS-specific expression silencing and the resulting phenotypic outcomes could confirm this capability of my screening setup.

Despite the efforts aimed at incorporating lncRNAs into the screen, I could not confidently identify lncRNA screen hits to follow up on. Out of seven lncRNA hits found in the screen, six were selected in only one cell line. The remaining lncRNA hit, *DDX11-AS1*, was identified in only two cell lines out of eight. Moreover, *DDX11-AS1* and five other lncRNA hits selected in the screen turned out to be expressed from bidirectional promoters shared with protein-coding genes, and thus could likely not be targeted by the CRISPRi system without affecting the expression of their neighbor. This negative result largely supported the conclusions of my previously published genome-wide analysis of lncRNA "CRISPRability" [147], suggesting that a vast majority of lncRNAs was not safely amenable to targeting by the CRISPR systems due to the high risk of neighbor gene deregulation. Biased identification of false positive hits originating from bidirectional promoters by the CRISPRi system has also been previously verified experimentally [235].

AFAP1-AS1 was the only lncRNA hit arising from a distant TSS located outside of other genes that could potentially be differentially targeted by the CRISPRi system. However, *AFAP1-AS1* has only been selected in the NCI-H838 cell line. Moreover, *AFAP1-AS1* has been extensively studied in lung cancer over the past years, confirming the validity of the screen results [237–240]. Thus, due to the absence of both common and *EGFR* status-specific novel lncRNA screen hits and high risk of neighbor TSS targeting, I did not further pursue the lncRNA line of research within this project.

6.2 Differential responses of *EGFR*-wildtype and -mutant LUAD cell lines to the *CENPE* knockdown

Contrary to my expectations, I could not identify screen hits differentially selected in the tested *EGFR*-mutant, but not in wildtype LUAD cell lines, i.e., the hypothesized synthetic lethal partners to mutant *EGFR*. Surprisingly, I detected the opposite selection pattern of relative resistance of *EGFR*-mutant compared to wildtype LUAD cell lines to the loss of a group of mitotic factors, most prominently, *CENPE*.

CENPE encodes for Centromere-associated protein-E (CENPE), a kinesin motor protein that localizes at the kinetochore and drives chromosome congression during mitosis. During prometaphase and metaphase, CENPE ensures stable capture of microtubules to kinetochores and transports mono-oriented chromosomes, i.e., chromosomes that attached to the microtubules originating from only one of the two spindle poles, to the spindle equator, ensuring proper chromosome alignment in the metaphase plate [241–243]. CENPE inhibition predominantly impairs the alignment of chromosomes positioned close to the centrosome during early prometaphase, while chromosomes located near the center of the forming spindle are exposed to a greater number of available microtubule ends and are thus able to align in the absence of CENPE activity [244, 245].

Additionally, CENPE is a key regulator of the mitotic cell cycle checkpoint, also known as the spindle assembly checkpoint (SAC), preventing aberrant chromosomal segregation [246]. During the SAC, a mitotic checkpoint complex (MCC) is formed at the unattached kinetochores. MCC inhibits the activation of the anaphase-promoting complex/cyclosome (APC/C) and thereby hinders the onset of anaphase before all chromosomes are properly attached to the spindle apparatus. Upon complete chromosome alignment, the SAC signaling is silenced and cells progress into anaphase [247]. CENPE, recruited to unattached kinetochores to facilitate correct chromosome alignment, plays an essential role in the silencing of the SAC [243, 248–252]. This function depends on its kinesin motor activity [252].

A number of studies in mammalian cells have demonstrated that disruption of CENPE function led to a mitotic delay, with the emergence of several characteristic unaligned chromosomes positioned in the proximity of the spindle poles [241, 244, 248, 249, 252–254]. In particular, a time-lapse microscopy study performed in HeLa cells upon siRNA-mediated silencing of *CENPE* revealed that in the majority of the examined cells, the chromosomes were

eventually able to align correctly, allowing the cells to proceed into anaphase [254]. Such result indicated a presence of redundant mechanisms coordinating SAC in the absence of CENPE, attributed to the activity of Aurora B kinase [254]. Additionally, a third of the examined CENPE-depleted HeLa cells proceeded to mitosis despite the presence of the misaligned chromosomes [254], corroborating earlier observations in a mouse model [249].

In several cancer models, the mitotic arrest induced by the CENPE disruption eventually led to apoptosis, suggesting CENPE as a putative target for cancer therapy [252, 255–257]. My results demonstrated that transcriptional repression of *CENPE* in *EGFR*-wildtype LUAD cell lines resulted in massive cell death, while *EGFR*-mutant cells were able to survive and continue the cell cycle. A minor decrease in the luminescent signal corresponding to the cell number of *EGFR*-mutant cell lines upon *CENPE* knockdown possibly illustrated a halt in the cell cycle due to the *CENPE* silencing-induced mitotic arrest discussed above.

The observed resistance of *EGFR*-mutant LUAD cell lines to the *CENPE* knockdown suggested that they were able to overcome the SAC, as CENPE activity is required for SAC silencing in the normal physiological context. A previous study by Maia *et al.* revealed that SAC was silenced rather than bypassed in aberrantly dividing HeLa cells upon CENPE depletion [254]. Continuous cell proliferation upon silencing of the SAC would lead to abnormal chromosome segregation and result in the accumulation of genomic aberrations. In agreement with this model, overexpression of *EGFR* has been previously linked to genomic instability [258]. Additionally, a recent study by Nahar *et al.* demonstrated a high level of genomic aberrations in *EGFR*-mutant LUAD patient samples, potentially underlying the development of drug resistant subclones [259]. The resulting accumulation of DNA damage could, in turn, sensitize *EGFR*-mutant LUAD cells to inhibition of DNA repair pathways. As such, *EGFR* mutations have been associated with sensitivity to PAPR inhibitors [260, 261]. Considering these previous observations, the ability of *EGFR*-mutant LUAD cells to overcome *CENPE* knockdown demonstrated by my results could constitute a novel tumorigenic mechanism and a therapeutic vulnerability in *EGFR*-mutant LUAD. Future characterization of this model could entail visualizing mitotic spindles and chromosome segregation, assaying SAC, genomic instability and DNA damage as well as testing the response to DNA repair inhibitors upon *CENPE* knockdown in *EGFR*-mutant and -wildtype LUAD cell lines.

In the context of this hypothesis, the observed synergistic effect of the simultaneous *EGFR* and *CENPE* silencing could suggest that the resistance of *EGFR*-mutant cell lines to

CENPE silencing-induced SAC depended directly on the expression of mutant *EGFR*. Remarkably, the synergistic effect was consistent and significant despite being weak in absolute terms, since knockdown of the *EGFR* expression alone expectedly strongly affected the viability of the PC-9 cell line, illustrating their addiction to the EGFR signaling.

Contrary to my expectations, ectopic overexpression of mutant *EGFR* did not render the *EGFR*-wildtype LUAD cell line Calu-6 resistant to the transcriptional silencing of *CENPE*. Such result, however, might be explained by the inherent impossibility to mimic the molecular context of oncogene addiction to mutant EGFR signaling in *EGFR*-wildtype cells. In order to improve the model, I am planning to overexpress mutant *EGFR* in wildtype cells upon stable knockdown of endogenous *EGFR* expression, and compare the resulting activation of molecular pathways attributed to oncogenic EGFR signaling, i.e., PI3K/AKT, MEK/ERK and JAK/STAT, to the levels exhibited by *EGFR*-mutant LUAD cell lines.

After validating the *CENPE* knockdown-mediated phenotypes, I also assayed the responses of *EGFR*-wildtype and mutant LUAD cell lines to the CENPE kinesin motor function inhibitor GSK-923295. A comprehensive study by Wood *et al*. demonstrated that GSK-923295 effectively suppressed the viability of 237 cancer cell lines, with the growth inhibiting value (GI_{50}) spanning over three orders of magnitude across the tested cell lines, from 12 nM to greater than 10 µM, with the average GI_{50} of 253 nM and the median GI_{50} of 32 nM [252]. However, neither the *EGFR*-wildtype Calu-6 and NCI-H838 LUAD cell lines nor the *EGFR*-mutant NCI-H1975 and PC-9 cell lines used in the present study have been profiled by Wood *et al*. in [252]. Corroborating the phenotypes resulting from the *CENPE* gene silencing, the *EGFR*-wildtype LUAD cell lines were sensitive to the pharmacological inhibition of CENPE with GSK-923295. Additionally, the *EGFR*-mutant LUAD cell line PC-9 was relatively tolerant to GSK-923295 from 100 nM up to 5 µM, corroborating the CRISPRi-based findings. However, the other tested *EGFR*-mutant cell line NCI-H1975 exhibited high sensitivity to the inhibitor. Additional optimization of the treatment conditions in a wider cell line panel, particularly the use of lower inhibitor concentrations as suggested by the results of Wood *et al.* [252], could derive more conclusive observations regarding the differential sensitivity of *EGFR*-wildtype and -mutant LUAD cell lines to the pharmacological inhibition of CENPE vs. the gene silencing.

A recent study by Mayes *et al*. revealed that suppression of MEK/ERK sensitized neuroblastoma, lung, pancreatic and colon cancer cell lines to CENPE inhibition with GSK-923295 [262]. Moreover, their data indicated that RAS-mutant cell lines were significantly

more resistant to GSK-923295 compared to RAS-wildtype cells, and ERK activation downstream of the RAS signaling was predictive for the resistance to GSK-923295 [262]. However, no mechanistic link between ERK activation and resistant to CENPE activation has been established.

Oncogenic *EGFR* mutations also result in an overactivation of the Ras/Raf/MEK/ERK pathway [263], which could potentially underlie the resistance of *EGFR*-mutant LUAD cell lines to *CENPE* knockdown demonstrated by my results. Additionally, the PC-9 cells harboring an oncogenic deletion in exon 19 of the *EGFR* gene appeared to be more resistant to both the CENPE inhibitor GSK-923295 and the STAT inhibitor SH-4-54, but not to the MEK/ERK inhibitor CI-1040 nor the PI3K inhibitor Wortmannin compared to the *EGFR*-wildtype cell lines, suggesting a possible implication of the JAK/STAT pathway in the mutant *EGFR*-mediated resistance to CENPE depletion (**Supplementary figure 4**).

EGFR activation has been previously shown to promote mitotic spindle aberrations and induce transcription of the Aurora kinase A gene via its interaction with STAT5 in several human cancer cell lines [264]. In turn, Aurora kinase A overexpression in mouse embryonic fibroblasts allowed the cells to override the mitotic checkpoint and resulted in cell division defects and elevated levels of chromosomal instability, also detected in HeLa cells [265, 266]. Interestingly, the activity of Aurora kinase A has been recently shown to be required for the emergence of acquired resistance phenotypes in *EGFR*-mutant LUAD cells [267]. Taken together, these observations suggest that STAT5-mediated Aurora kinase A overactivation could be one possible mechanistic link between the *EGFR* status and cellular response to CENPE disruption, providing an interesting starting point for the follow-up studies.

In summary, I suggest that my observations placed in the context of previous publications could be indicative of the following model: I propose that the oncogenic activation of EGFR allows LUAD cells to overcome SAC, which is illustrated by the observed resistance of *EGFR*-mutant LUAD cell lines to the *CENPE* knockdown. In the context of *in vivo* carcinogenesis, this likely leads to aberrant cancer cell mitosis and subsequent accumulation of genomic abnormalities, potentially driving the emergence of drug resistant subclones. On the other hand, the increased DNA damage might provide a vulnerability for therapeutic targeting of *EGFR*-mutant LUAD, e.g., via inactivation of DNA repair mechanisms with PARP inhibitors. Thus, identification of factors that mediate the resistance of *EGFR*-mutant cells to *CENPE* silencing may unravel novel mechanisms of LUAD carcinogenesis, improve our

understanding of the development of drug resistance in *EGFR*-mutant LUAD and provide new therapeutic options for lung cancer treatment.

6.3 *CASP8AP2* as a novel LUAD viability factor

CASP8AP2 was the only screen target negatively selected in all eight tested cell lines, indicating its essential role in LUAD cell viability, that was not previously annotated as a core essential gene [211]. *CASP8AP2* functionality has also not been previously described in the context of lung cancer.

The *CASP8AP2* gene encodes for the Caspase-8 Associated Protein-2, also known as FLICE-associated huge protein (FLASH). It was originally identified as a pro-apoptotic factor mediating the interaction between the activated Fas/FADD complex and the initiator caspase-8 [268] Since its initial discovery, CASP8AP2 has been found to participate in numerous biological processes outside of apoptotic signaling. However, most of its additional functions remain insufficiently characterized and often seem contradictory.

In contrast to its originally characterized pro-apoptotic function, CASP8AP2 protein levels were found to be decreased upon induction of apoptotic stimuli with Fas ligand in the human fibrosarcoma cell line HT1080 [269]. Moreover, knockdown of *CASP8AP2* expression promoted cell death in a number of experimental models [270–274]. Further *CASP8AP2* silencing experiments demonstrated its essential role in S-phase cell cycle progression in an array of both transformed and non-transformed human and murine cell lines [270, 273, 275, 276]. *CASP8AP2* has also been found to coordinate the TNF-α-induced activation of NF-κB pathway, generally associated with cell survival [277, 278]. Other functions of CASP8AP2 include regulation of histone gene expression attributed to its interaction with p73 and 3'-end processing of histone pre-mRNAs [270, 271, 276, 279], [280]. It has also been shown to participate in the regulation of the activity of several transcription factors, particularly c-Myb and glucocorticoid receptor (GR) [281–284].

My findings of *CASP8AP2* knockdown-induced loss of cell viability in NSCLC cell lines were in line with the previously characterized effects of *CASP8AP2* depletion in other cancer cell types [270–273]. However, my work was first to demonstrate the tolerance of non-transformed lung cell lines to *CASP8AP2* knockdown, to characterize the divergent

phenotypes between non-transformed and cancer cell lines of the same entity, and particularly the first to examine the involvement of *CASP8AP2* in lung cancer cell survival.

In agreement with the differences in cell viability phenotypes, I observed an activation of apoptotic signaling upon *CASP8AP2* silencing in sensitive NSCLC cell lines, but not in non-transformed IMR-90 cells. Caspase-8 and caspase-3/7 were found activated in the context of *CASP8AP2* depletion in several previous studies [269, 271], however, activation of caspase-9 had not been assessed before.

Simultaneously detected activation of initiator caspase-8, primarily mediating extrinsic apoptosis induced via death receptors, and caspase-9, regulating intrinsic mitochondria-controlled apoptotic pathways [285], could either indicate redundant apoptosis activation upon *CASP8AP2* silencing, or be the result of caspase feedback amplification recently described by McComb *et al*. [286]. The concept suggests that activation of effector caspases-3 and -7 by either extrinsic or intrinsic apoptotic pathways results in a parallel feedback activation of both upstream initiator caspases-8 and -9. Thus, I performed caspase silencing experiments in an attempt to determine the dominant apoptotic pathway mediating the loss of cell viability in lung cancer cell lines upon *CASP8AP2* depletion. Surprisingly, knockdown of neither caspase-8, nor caspase-9, nor simultaneous knockdown of caspases-3 and -7, previously shown to abolish both intrinsic and extrinsic apoptosis in human leukemia cell lines [286], had any rescue effect on cell viability of sensitive LUAD cell lines upon *CASP8AP2* silencing. Moreover, treatment with pan-caspase inhibitor z-VAD-FMK yielded only partial rescue of *CASP8AP2* depletion-induced loss of cell viability in NCI-H1975 cells. Taken together, these results suggested that *CASP8AP2* knockdown either led to excessive overactivation of a network of redundant apoptotic pathways, or resulted in the onset of non-apoptotic cell death, e.g., autophagy, necrosis, or caspase-independent cell death [287–289]. Further validation of my results using stable caspase-3/7, -8 and -9 knockout or knockdown cell lines could exclude the possible effects of residual caspase activity upon transient gene silencing at the time of readout.

Contrary to my expectations, I also observed a strong caspase-3/7 and caspase-9 activation as well as a weaker but significant activation of caspase-8 in the NCI-H1563 LUAD cell line, where knockdown of *CASP8AP2* had a relatively negligible effect on cell viability. Resistance to cleaved caspase fragments has been previously described, drug resistant cancer cells [290], but has not been documented for other cell lines. Thus, it was conceivable to

hypothesize that the level of caspase activation occurring in lung cancer cells in response to *CASP8AP2* knockdown did not bear physiological significance. Importantly, this observation suggested that mechanisms determining differential sensitivity to *CASP8AP2* loss were outside of apoptotic signaling activation.

A previous study by Sokolova *et al*. [273] demonstrated different patterns in cell cycle progression between the osteosarcoma cell line U2OS and the immortalized non-cancer epithelial cell line hTERT-RPE1 upon *CASP8AP2* silencing. According to their proposed model, knockdown of *CASP8AP2* led to histone deregulation, which in turn led to nucleosome deficiency and resulted in a disruption of cell cycle progression in the S-phase in both cell lines. As such, hTERT-RPE1 cells responded to the nucleosome depletion by directly arresting viable cells in S-phase in a p53-dependent manner, while U2OS continued the cell cycle into an elongated slow-progressing S-phase that lasted up to three days and ultimately resulted in cell death [273].

My results corroborated the previously identified involvement of *CASP8AP2* in the S-phase regulation. The absence of a BrdU-positive population of non-transformed IMR-90 cells upon *CASP8AP2* knockdown suggested a complete cell cycle arrest, however, all lung cancer cell lines tested continued to cycle despite the S-phase disruption, which occurred at different stages of the S-phase in different cell lines. Such a cell line-specific pattern could be explained by the slow S-phase progression model proposed by Sokolova *et al.* in [273]. Due to inherent differences in proliferation rates, different cell lines could be progressing through the elongated S-phase at varying rates and could therefore be stalled at different stages at the timepoint of readout. Time-course experiments could test this hypothesis. Alternatively, the heterogeneous S-phase disruption pattern could suggest that *CASP8AP2*-dependent regulation of S-phase progression involves redundant pathways acting at different stages of the S-phase that are activated differentially across lung cancer cell lines according to their molecular contexts.

Contrary to the model proposed in [273] and [269], knockdown of *TP53* did not rescue the *CASP8AP2* depletion-induced cell cycle arrest in the non-cancer lung cell line IMR-90 (**Supplementary figure 15**). Neither did the silencing of the p53 target gene *CDKN1A*, which I found to be upregulated in both non-transformed lung and NSCLC cell lines upon *CASP8AP2* knockdown. Previous studies also detected the *CASP8AP2* silencing-induced overexpression

of *CDKN1A* in different cell lines, and it was also suggested to mediate *CASP8AP2* depletion-induced phenotype in a rat glomerulonephritis model [271, 273, 274]

Silencing of neither *TP53* nor *CDKN1A* had any additional effect on cell survival upon *CASP8AP2* knockdown in sensitive NSCLC cell lines, and both had a negligible negative effect on the cell viability of tolerant cell lines (**Supplementary figures 10 - 13**). Additionally, knockdown of *TCF3*, previously suggested to reverse the *CASP8AP2* depletion-mediated loss of cell viability in rat glomerulonephritis model [274], did not have affect the viability of NSCLC cells upon *CASP8AP2* knockdown (**Supplementary figure 14A**). Moreover, in contrast to the mode of action previously proposed in [274], silencing of *TCF3* did not repress *CDKN1A* expression in NSCLC cells neither in control conditions nor upon *CASP8AP2* knockdown, suggesting that interactions between these factors were cell type- or organism-specific (**Supplementary figure 14B**).

Interestingly, my results demonstrated that the *CASP8AP2* depletion-induced *CDKN1A* upregulation was not dependent on p53 in NSCLC cell lines, as it was detected in both p53-negative and p53-positive NSCLC cells. Additionally, silencing of *TP53* did not affect overexpression of *CDKN1A* in the p53-positive NCI-H1975 cell line and only partially suppressed *CDKN1A* expression in the p53-wildtype NCI-H1563 and NCI-H460 cells upon *CASP8AP2* knockdown. The NCI-H1975 cell line bears a hotspot mutation in *TP53*, resulting in an arginine-to-histidine substitution in codon 273 (R273H), previously characterized as gain-of-function mutation promoting cancer proliferation, invasion, and drug resistance [291, 292]. Thus, it was plausible to suggest that in the context of wildtype p53, its activity contributed to the regulation of *CDKN1A* expression, but it did not determine *CASP8AP2* depletion-induced overexpression of *CDKN1A* in NSCLC cell lines.

On the other hand, *CDKN1A* expression was completely abolished in non-cancer IMR-90 cells upon *TP53* knockdown, indicating that the alternative pathway coordinating *CDKN1A* regulation upon *CASP8AP2* silencing in NSCLC cell lines was inactive in the IMR-90 cells. Additionally, the relatively lower level of *CDKN1A* upregulation in the CASP8AP2 depletion-tolerant NCI-H1563 cell line indicated that the alternative pathway was deactivated in both tolerant cell lines and possibly acted as the determinant of the divergent response to *CASP8AP2* silencing.

Examination of the lung cell line RNA and protein expression datasets [205] revealed that *JUNB* was overexpressed in the lung cancer cell lines sensitive to *CASP8AP2* knockdown

compared to the tolerant cell lines in steady state. Moreover, gene expression analysis using RNA-Seq demonstrated that *JUNB* mRNA levels were significantly upregulated in the sensitive cancer cell lines upon *CASP8AP2* silencing while being weakly but significantly downregulated in the non-transformed IMR-90 lung cell line. Validation by RT-qPCR demonstrated significant elevation of *JUNB* mRNA levels in the sensitive NSCLC cells and the lack of *JUNB* regulation in the IMR-90 cell line. The tolerant lung cancer cell line NCI-H1563 also exhibited upregulated *JUNB* mRNA levels However, in the context of the non-detectable JUNB protein expression in the steady state suggested by the proteomics dataset, the less than two-fold induction of *JUNB* mRNA levels in NCI-H1563 cells upon *CASP8AP2* silencing detected by RT-qPCR likely did not result in functionally relevant JUNB activation. Therefore, these observations suggested that differences in relative levels of JUNB expression determined the divergent responses between the sensitive and the tolerant cell lines to *CASP8AP2* silencing and identified JUNB as a potential downstream factor of CASP8AP2.

JUNB encodes for a subunit of the Activator protein-1 (AP-1) complex. AP-1 is a pluripotent transcription factor that mediates gene expression reprogramming in response to a plethora of environmental stimuli under normal as well as pathological conditions, including cancer, resulting in a range of outcomes from cell proliferation to apoptosis [293]. Due to this broad functional heterogeneity, specific modes of AP-1 action in individual entities, e.g., lung cancer, remain largely unexplored.

AP-1 is composed by members of Fos, Jun, ATF and MAF protein families that form various functional homo- and heterodimers through their basic region - leucine zipper (bZip) domains. Different dimers are regulated differentially in a cell type- and stimulus-specific manner and recognize different DNA response elements, indicating that individual AP-1 components execute distinct biological functions. [293–298]. JUNB was originally characterized as an inhibitor of cell proliferation and transformation and is therefore generally linked to tumor suppressor functionality [299]. However, it was also found to promote cell transformation and proliferation of cancers, especially in several types of lymphomas [299].

Inhibition of AP-1 transcriptional activity rescued the loss of cell viability upon *CASP8AP2* silencing in the sensitive NSCLC cell lines. These results confirmed that the CASP8AP2 – AP-1 functional axis governed the sensitivity of NSCLC cell lines to *CASP8AP2* silencing suggested by the differential expression patter of *JUNB* between sensitive and

tolerant cell lines. Moreover, inhibition of AP-1 attenuated the induction of *CDKN1A* expression in the sensitive NSCLC cells upon *CASP8AP2* silencing, fulfilling a missing link between *CASP8AP2* and *CDKN1A* regulation in NSCLC.

Future studies could entail further knockdown- or knockout-based validations of the involvement of JUNB in the modulation of CASP8AP2 – AP-1 functional axis, as the available small molecule inhibitors of AP-1, including T-5224 used in this study, are not specific to JUNB-containing AP-1 dimers [228, 229]. Additionally, further characterization of *CASP8AP2* silencing-mediated changes in the molecular landscape in NSCLC cells would throw light on the factors coordinating the interplay between CASP8AP2 and JUNB/AP-1 activity.

In summary, the present identification and characterization of *CASP8AP2* as an essential and cancer-specific factor for NSCLC viability suggests that it may be a promising novel therapeutic candidate for lung cancer treatment. Future steps towards elucidating the true therapeutic potential of CASP8AP2 inhibition could be to verify its effectiveness on cancer cells as well as the suggested tolerance of normal cells in *in vivo* lung cancer models. Besides, further definition of the molecular functions of CASP8AP2 in the context of lung cancer could enable the design of therapeutic intervention possibilities.

6.4 Conclusions

Within my PhD project, I developed and conducted a custom CRISPRi screen for novel general and genotype-specific LUAD viability factors by targeting all TSSs within the genes of interest, allowing me to differentiate between functional gene isoforms originating from individual TSSs. The screen revealed that relative position of the target TSS within the respective gene was a better predictor of its effects on the cell viability than its relative expression. However, neither the relative position nor the relative expression could universally predict the screen performance of individual TSSs, corroborating the usefulness of my screen design covering all TSSs within the target genes.

Overall, my CRISPRi screen identified several novel molecular determinants of lung cancer viability and provided starting points for understanding of their molecular modes of action and their therapeutic potential. Particularly, this study uncovered the resistance of *EGFR*-mutant LUAD cell lines to the loss of *CENPE* expression, which was essential for the viability of *EGFR*-wildtype LUAD cells. Besides, it demonstrated the synergistic effect of simultaneous *EGFR* and *CENPE* silencing on the growth suppression of *EGFR*-mutant LUAD cells and discussed a potential mechanistic link between the two factors as well as its putative clinical significance.

Moreover, I discovered *CASP8AP2* as a cancer-specific essential viability factor in NSCLC and demonstrated the tolerance of non-transformed lung cells to the *CASP8AP2* knockdown. Strikingly, my results indicated that the differential level of *JUNB* expression between the tolerant and the sensitive cell lines determined their divergent responses to *CASP8AP2* depletion. Corroborating this hypothesis, I established the involvement of the AP-1 transcription factor activity in the *CASP8AP2* silencing-induced loss of cancer cell viability and provided the basis for further elucidation of CASP8AP2 as a putative therapeutic target for lung cancer.

[This page intentionally left blank]

7 Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Dr. Sven Diederichs, for giving me a chance and supporting me all the way through this project. Thank you for the trust you have put in me, and the opportunities to learn you have created through the close supervision and the constructive criticism that you have provided. I would like to thank the members of my Thesis Advisory Committee, Prof. Dr. Georg Stoecklin and Prof. Dr. Christoph Peters, for their support and their scientific inputs to this project. I would also like to thank Prof. Dr. Stefan Wiemann and Prof. Dr. Peter Angel for agreeing to participate in my PhD examination committee.

I would like to thank the DKFZ Genomics and Proteomics Core Facility, Dr. Dietmar Pfeifer, Andreas Janes, Pia Veratti and the Freiburg Galaxy team for teaching me and helping with the NGS part of my project. I am especially grateful to Dr. Andrew Walsh for programming the sgRNA design algorithm. I would also like to thank the Lighthouse Core Facility in ZTZ Freiburg, especially Dieter Herchenbach and Dr. Marie Follo for their cooperation.

I would like to thank my colleagues in Freiburg and Heidelberg labs for the scientific discussions and the support they have provided, as well as for the fun we had throughout the process. I am grateful to Ashish, Yogita, Maïwen and Jeanette, who all have helped me at different stages of this project. I would like to thank Sonam for being an extraordinary mentor and friend. I am particularly thankful to Marisa for stepping up and taking over a lot of the pending experimental tasks during the last few months, allowing me to focus on the writing. Additionally, I would like to thank Guilia and Jane for the contributions they have made during their internships.

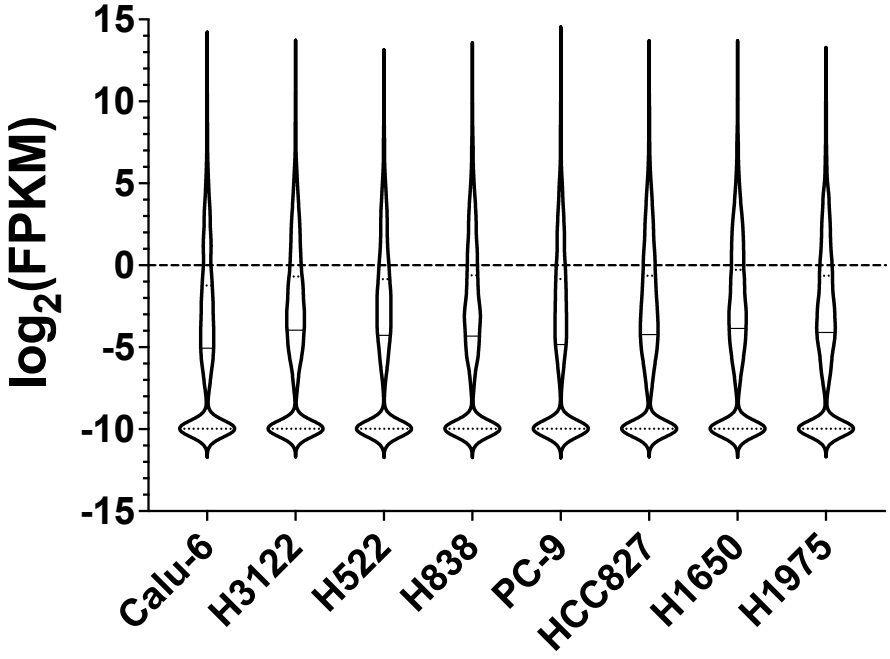
This has been an exciting journey full of discoveries and adventures, challenges and pitfalls, ends and new beginnings, in and beyond the lab. I would like to thank my parents, Anna and Alexander, for believing in me and trusting the process, and my sister Anastasia for being my closest ally. I am thankful to the rest of my family, especially my aunt Nadezhda and my grandparents, Sergey and Evgeniya, for their continuous support and contagious enthusiasm.

I am enormously grateful to all my friends for their support, care, and their patience with me throughout these past years. Quentin, without your advice and friendship this whole

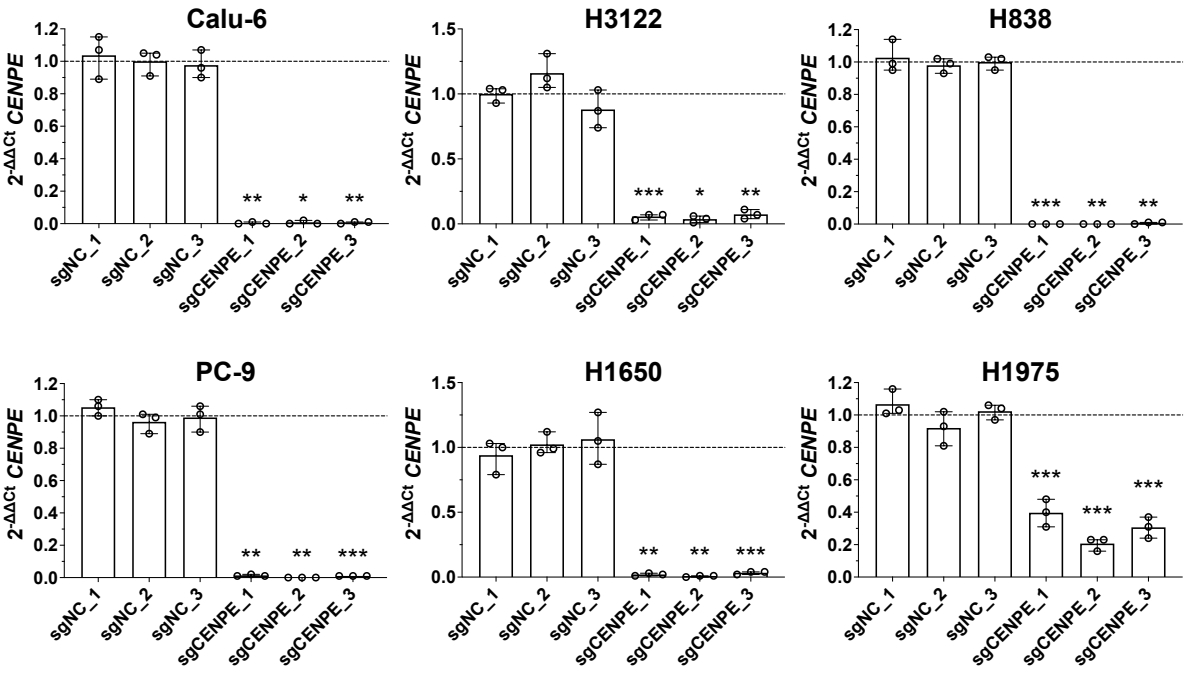
project (and so many other things) would not have been possible. Nataliya, our friendship is my comfort zone, thank you for always having my back. Veronika, Polina and Dmitry, Patrick, Angelica, Shravan, Paulo and the whole Heidelberg team, my basecamp back in Russia – I could not have done without you all. Additionally, I would like to thank the Maniacts Freiburg crew for providing a solid counterweight to that work-life balance, and Irina Golubeva for being there for me and reminding me to breathe.

Finally, I would like to thank my boyfriend, William, for his unconditional support, the motivation to keep going, and for helping me to remember what this all really is about.

8 Supplementary figures



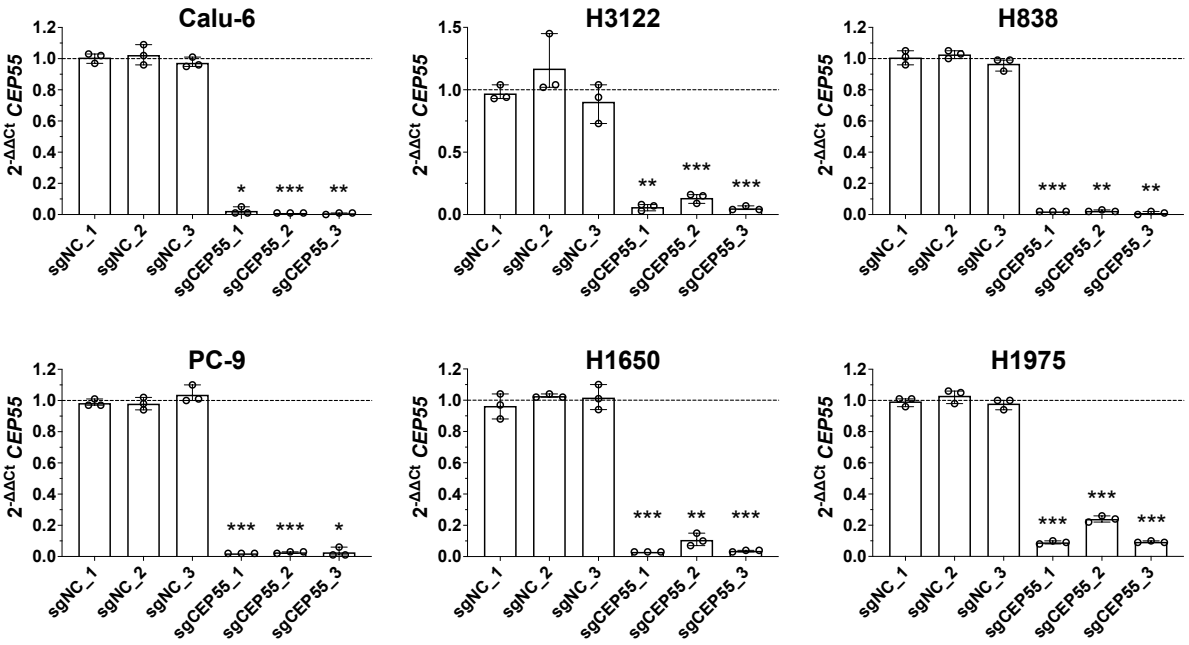
Supplementary figure 1. Target library was enriched in transcripts exhibiting low RNA expression values in the screened cell lines.
Distribution of RNA expression values corresponding to the first exon downstream of each TSS targeted by the CRISPRi screen in the eight screened cell lines. The expression data were derived from the Diederichs lab dataset [205].
RNA expression values are represented as $\log_2$(FPKM+0.001). Solid lines correspond to the mean values, dotted lines correspond to the first and third quartiles.



Supplementary figure 2. *CENPE* mRNA expression was efficiently suppressed by the CRISPRi system.

RT-qPCR was performed to assess the efficiency of target knockdown in LUAD cells transduced independently with three constructs expressing sgRNAs against *CENPE* (sgCENPE) compared to negative control sgRNAs (sgNC). Cell lysis for RNA isolation was performed at 4 days post-transduction for Calu-6 and H838, 5 days for PC-9 and H1975, 7 days for H1650 and H3122 (n = 3).

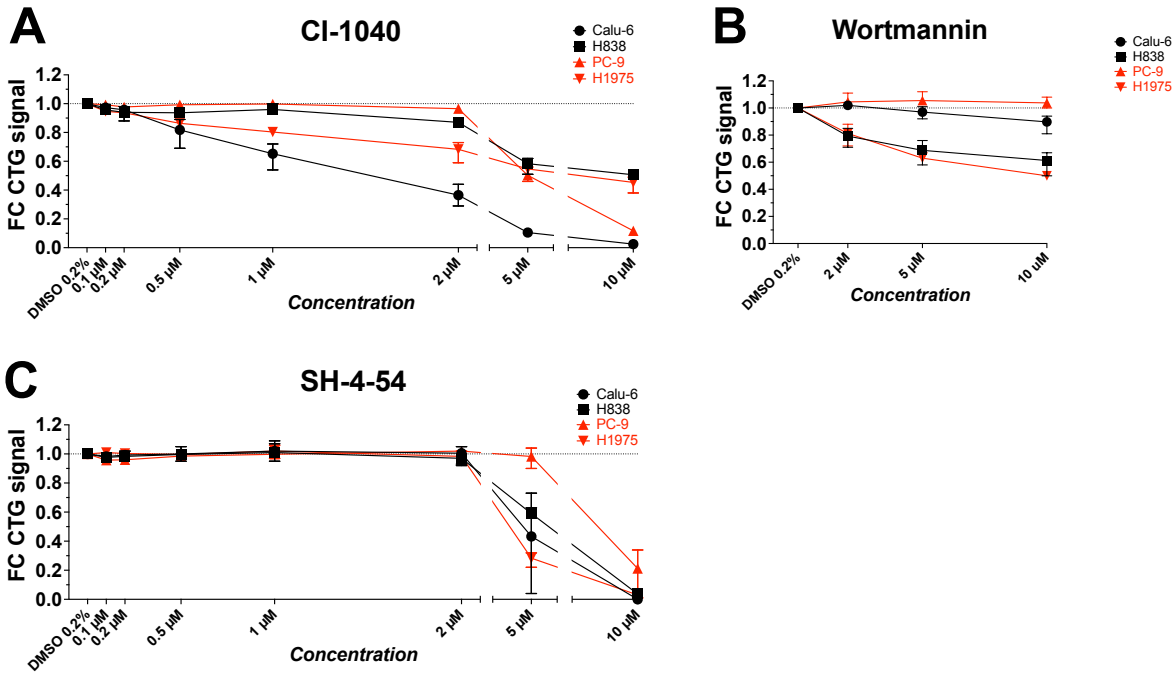
For each biological replicate, data were normalized to the average of negative control signals and plotted as individual values. Bar heights represent mean, error bars represent range. Significance was calculated on $\log_2$-transformed fold change data using two-tailed unpaired Student's t-test with Welch's correction, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.



Supplementary figure 3. *CEP55* mRNA expression was efficiently suppressed by the CRISPRi system.

RT-qPCR was performed to assess the efficiency of target knockdown in LUAD cells transduced independently with three constructs expressing sgRNAs against *CEP55* (sgCEP55) compared to negative control sgRNAs (sgNC). Cell lysis for RNA isolation was performed at 4 days post-transduction for Calu-6 and H838, 5 days for PC-9 and H1975, 7 days for H1650 and H3122 (n = 3).

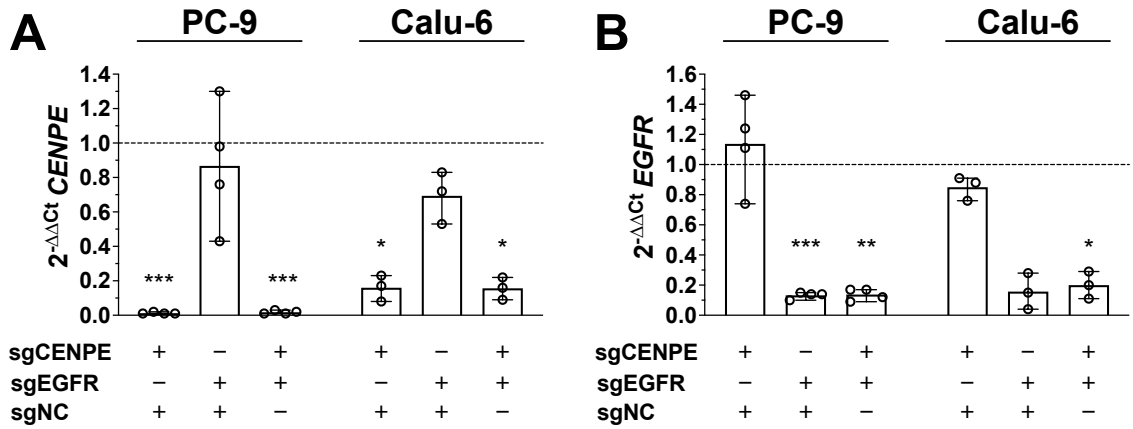
For each biological replicate, data were normalized to the average of negative control signals and plotted as individual values. Bar heights represent mean, error bars represent range. Significance was calculated on $\log_2$-transformed fold change data using two-tailed unpaired Student‘s t-test with Welch‘s correction, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.



Supplementary figure 4. Responses of *EGFR*-wildtype and -mutant LUAD cell lines to the treatments with EGFR-downstream pathway inhibitors.

CellTiter-Glo cell viability assay (CTG) was performed on *EGFR*-wildtype (WT) cell lines Calu-6 and H838 and EGFR-mutant (Mut) cell lines PC-9 and H1975 treated with 0.1-10 µM CI-1040 **(A)**, 2-10 µM Wortmannin **(B)**, 0.1-10 µM SH-4-54 **(C)** or vehicle solvent (DMSO 0.2%, normalization control) 72h post-treatment (n = 3-4).

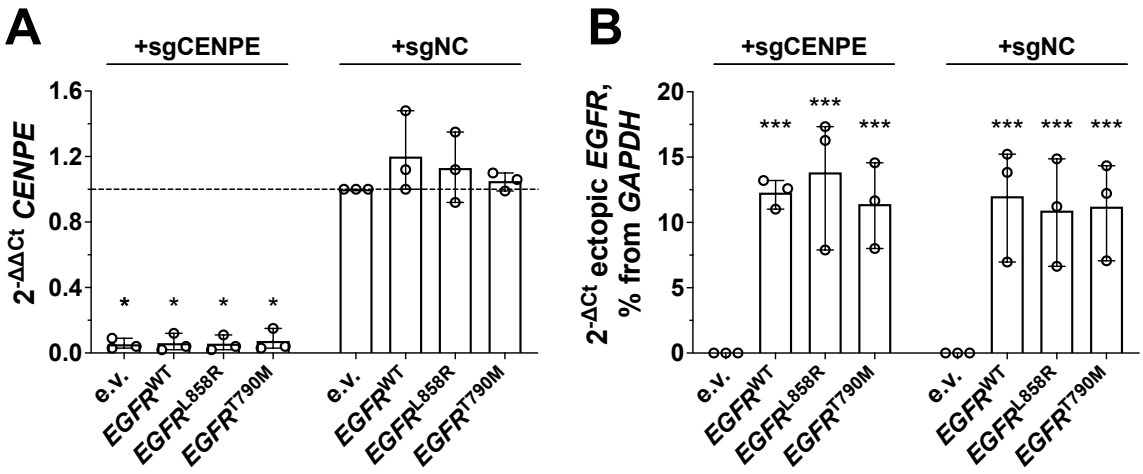
For each biological replicate, data were normalized to the respective negative control signal. Data are represented as mean per condition, error bars represent range.



Supplementary figure 5. *CENPE* and *EGFR* RNA expression was efficiently suppressed in PC-9 and Calu-6 cells transduced with combinations of CRISPRi pools.

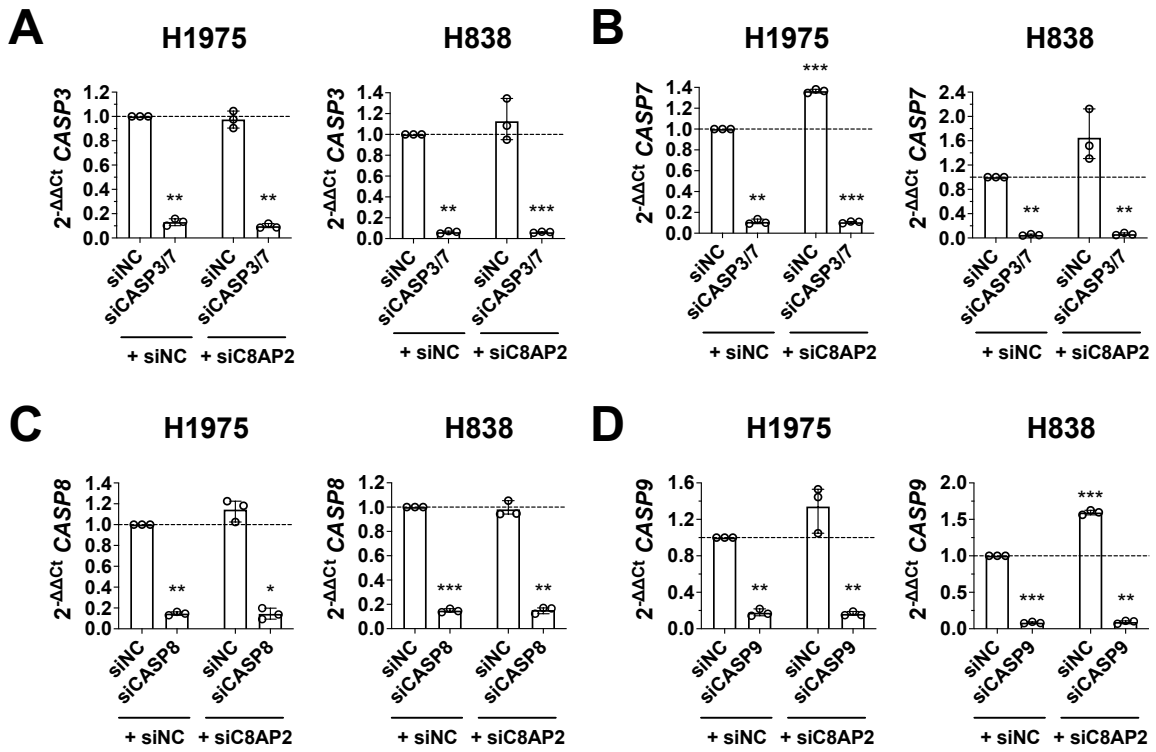
(A) RT-qPCR was performed to assess the efficiency of *CENPE* KD and (**B)** *EGFR* KD in PC-9 and Calu-6 cells (n = 3-4). Cell lysis for RNA isolation was performed at 4 days post-transduction for Calu-6 cells and 5 days for PC-9 cells. Cells transduced with sgNC were used as normalization control. Total amount of lentiviral stock used for each transduction was equal across all conditions.

For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance was calculated on $\log_2$-transformed fold change data using two-tailed paired Student‘s t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

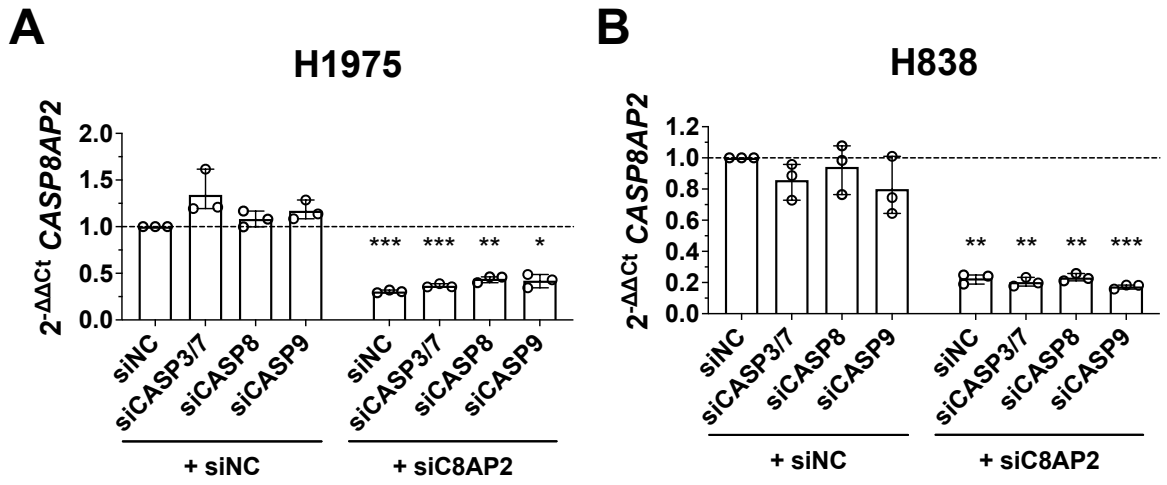


Supplementary figure 6. *CENPE* and ectopic *EGFR* mRNA expression in Calu-6 cells stably transduced with EGFR OE and sgCENPE constructs.

(A) RT-qPCR was performed to assess the efficiency of *CENPE* KD and **(B)** ectopic *EGFR* OE in Calu-6 cells (n = 3). Cell lysis for RNA isolation was performed at 4 days after transduction with the CRISPRi constructs. Cells stably transduced with e.v. and sgNC were used as normalization control. For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance was calculated on log_2-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

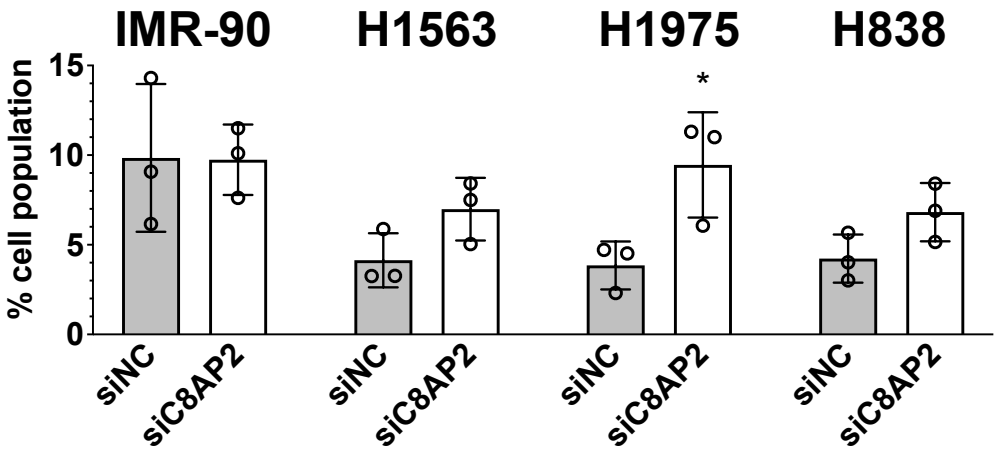


Supplementary figure 7. RT-qPCR to assess the mRNA levels of *CASP3* (A), *CASP7* (B), *CASP8* (C) and *CASP9* (D) in H1975 and H838 cell lines transfected with indicated combinations of siPOOLs against *CASP8AP2* (siC8AP2), combined siPOOL against *CASP3 and CASP7* (siCASP3/7), *CASP8* (siCASP8), *CASP9* (siCASP9) and negative control siPOOL (siNC) at 48 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3). Cells transfected with siNC (20 nM) were used as normalization control.
For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed calculated on log_2-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; "n.s.", not significant.

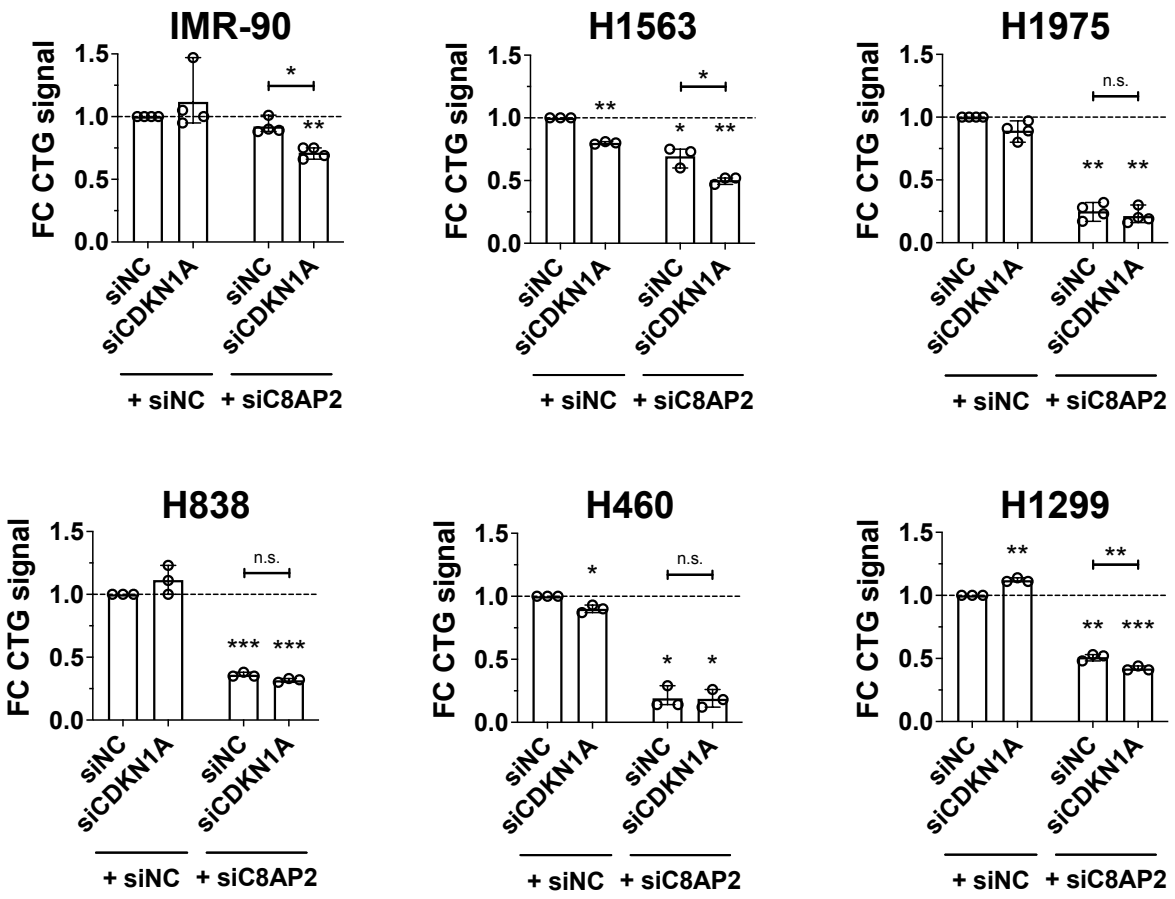


Supplementary figure 8. RT-qPCR to assess the mRNA levels of *CASP8AP2* in H1975 **(A)** and H838 **(B)** transfected with indicated combinations of siPOOLs against *CASP8AP2* (siC8AP2), combined siPOOL against *CASP3 and CASP7* (siCASP3/7), *CASP8* (siCASP8), *CASP9* (siCASP9) and negative control siPOOL (siNC) at 48 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3). Cells transfected with siNC (20 nM) were used as normalization control.

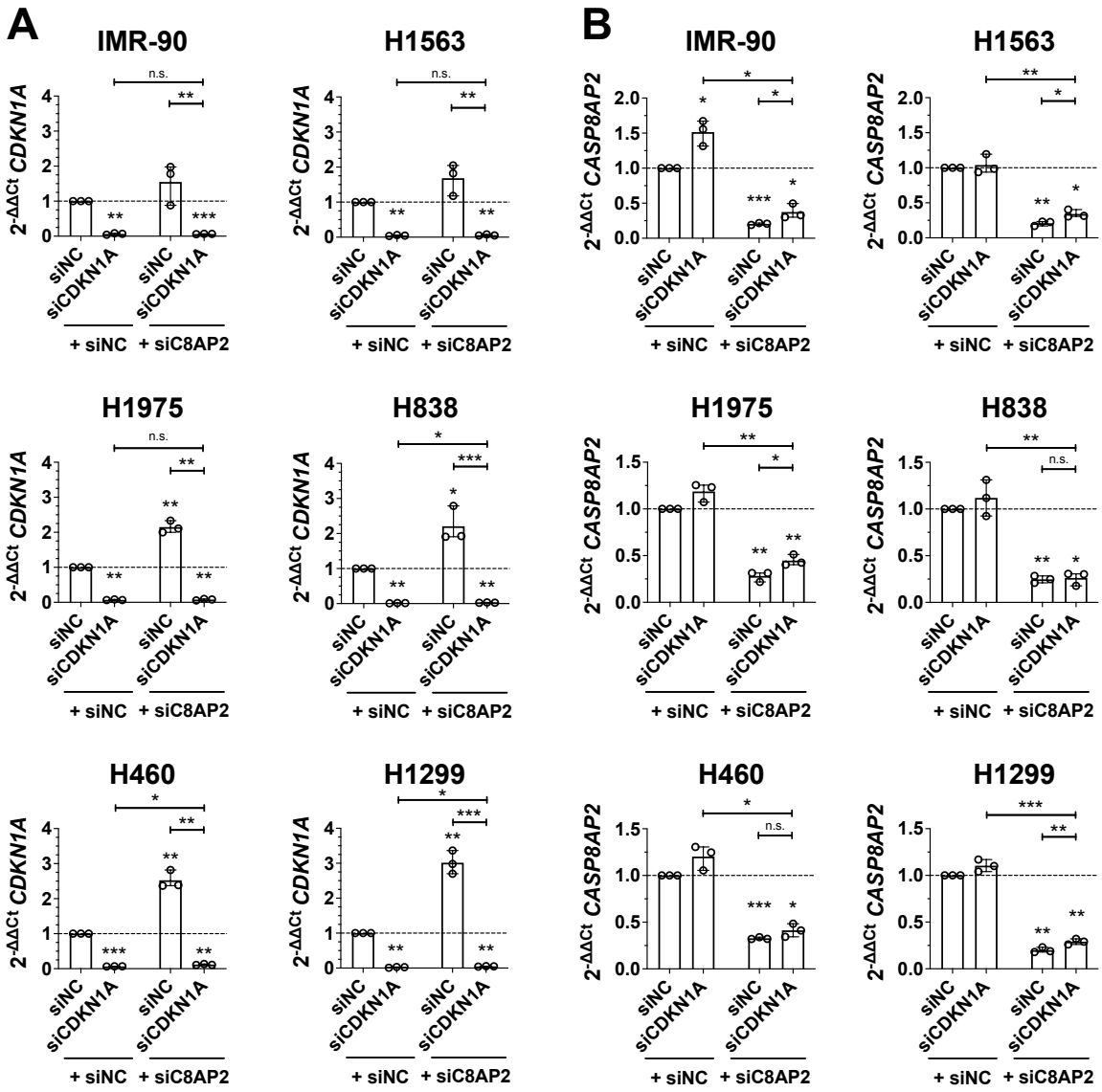
For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed calculated on log_2-transformed fold change data using two-tailed paired Student‘s t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; “n.s.”, not significant.



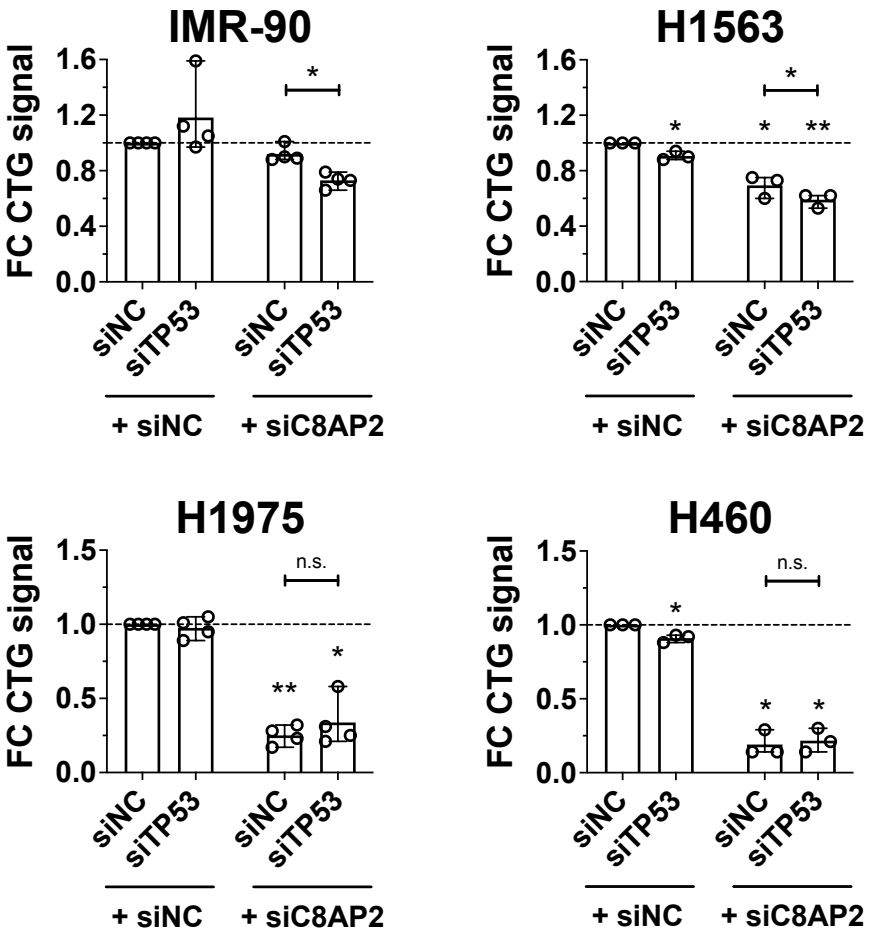
Supplementary figure 9. Ungated flow cytometry signals corresponding to subG1 cell cycle phase on the PI histogram plot were quantified using FlowJo v10 software. Data are represented as individual values per biological replicate, bar heights represent mean, error bars represent range. Significance testing was performed using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.



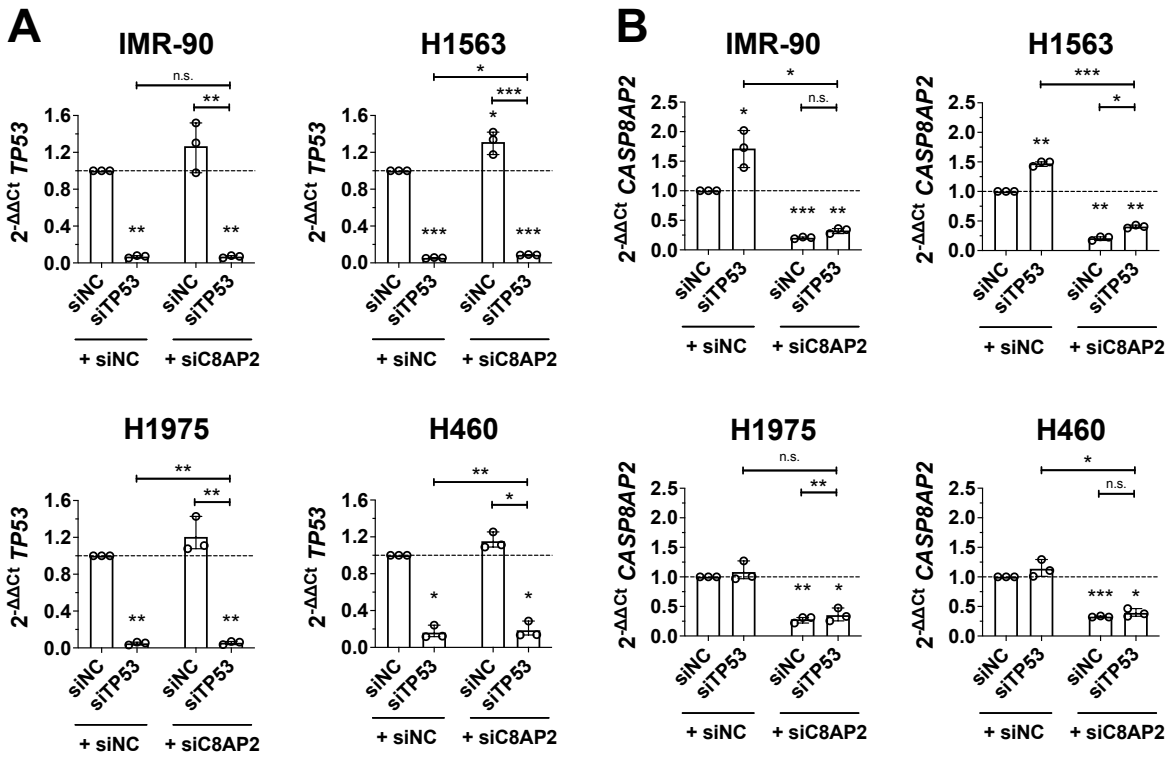
Supplementary figure 10. CellTiter-Glo cell viability assay (CTG) to assess the effect of siPOOLs-mediated knockdown of *CDKN1A* (siCDKN1A) on siC8AP2-mediated phenotype in IMR-90, H1563, H1975, H838, H460 and H1299 cell lines at 96 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3).
For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed calculated on log_2-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; "n.s.", not significant.



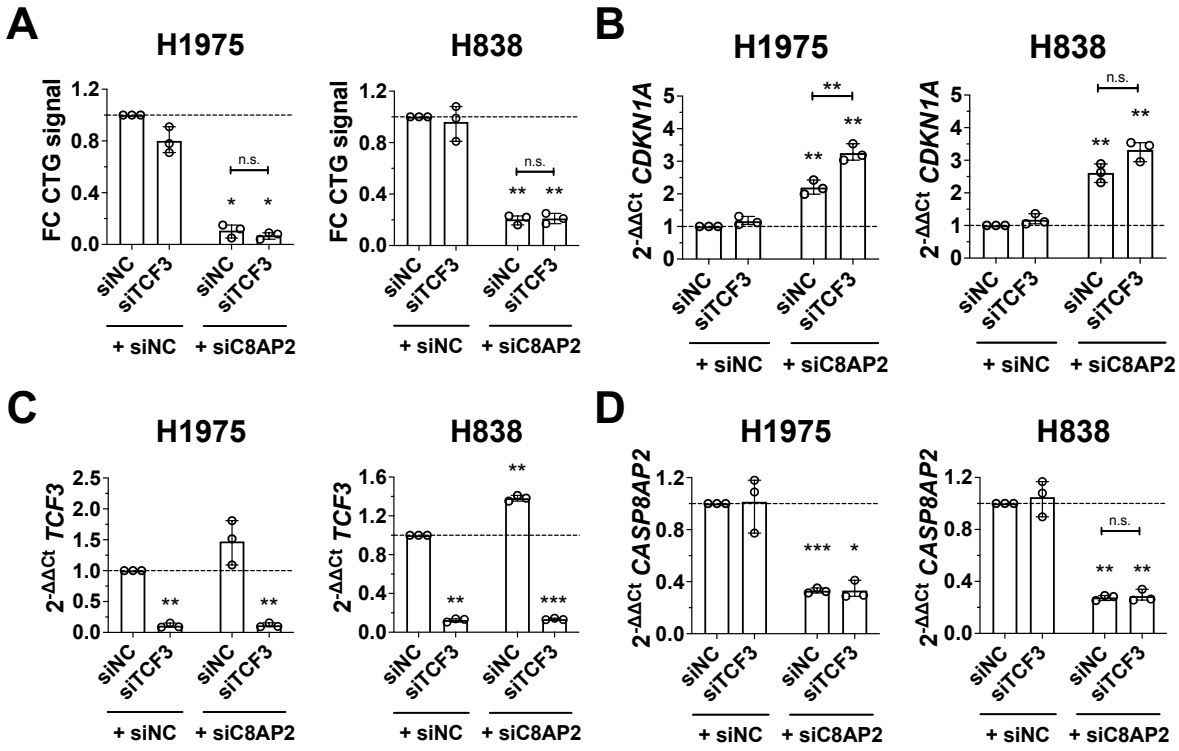
Supplementary figure 11. RT-qPCR to assess the mRNA levels of *CDKN1A* **(A)** *and CASP8AP2* **(B)** in IMR-90, H1563, H1975 H838, H460 and H1299 cells transfected with indicated combinations of siPOOLs against *CASP8AP2* (siC8AP2), *CDKN1A* (siCDKN1A) and negative control siPOOL (siNC) at 48 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3). Cells transfected with siNC (20 nM) were used as normalization control. For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed calculated on log_2-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; "n.s.", not significant.



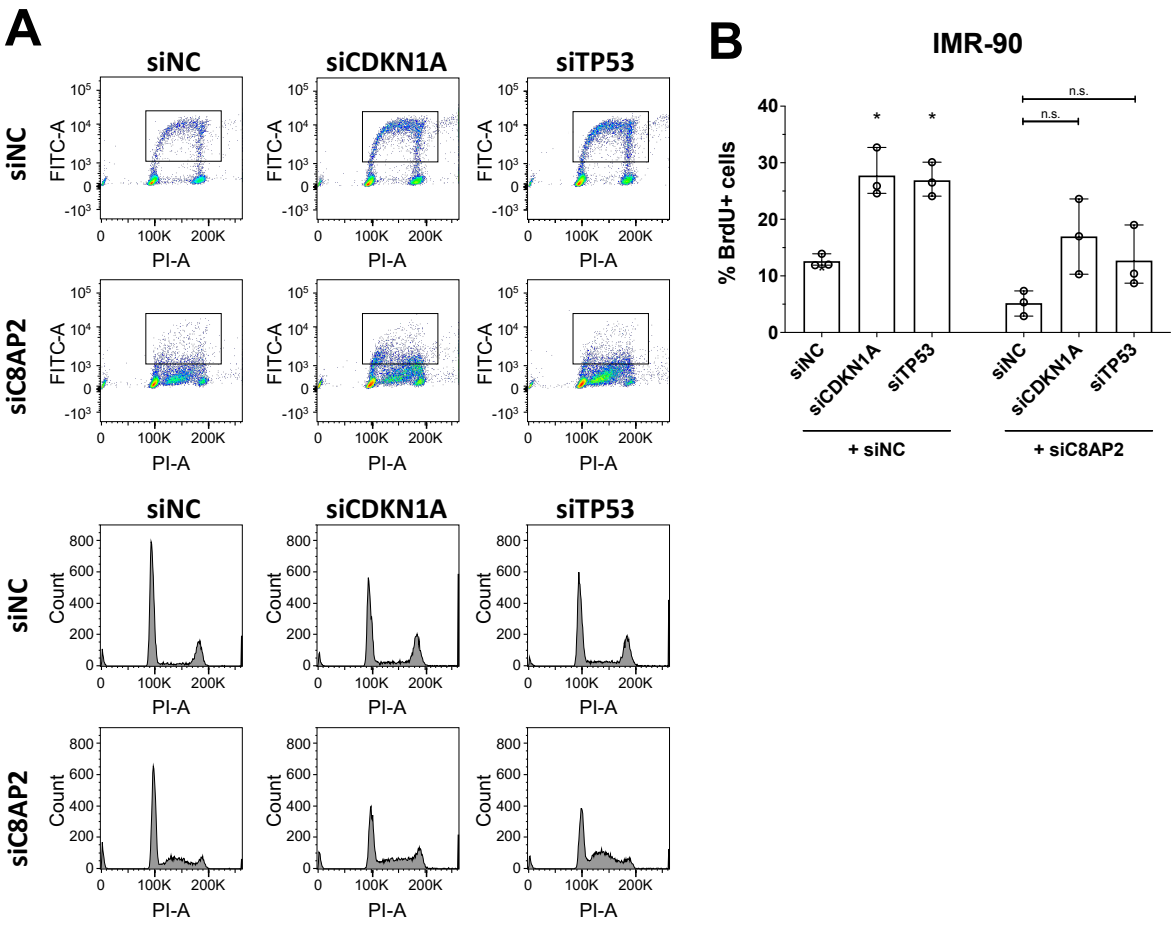
Supplementary figure 12. CellTiter-Glo cell viability assay (CTG) to assess the effect of siPOOLs-mediated knockdown of *TP53* (siTP53) on siC8AP2-mediated phenotype in IMR-90, H1563, H1975, H838, H460 and H1299 cell lines at 96 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3-4).
For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed calculated on log_2-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; "n.s.", not significant.



Supplementary figure 13. RT-qPCR to assess the mRNA levels of *TP53* **(A)** and *CASP8AP2* **(B)** in IMR-90, H1563, H1975 and H460 cells transfected with indicated combinations of siPOOLs against *CASP8AP2* (siC8AP2), *TP53* (siTP53) and negative control siPOOL (siNC) at 48 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3). Cells transfected with siNC (20 nM) were used as normalization control.
For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed calculated on log_2-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; "n.s.", not significant.



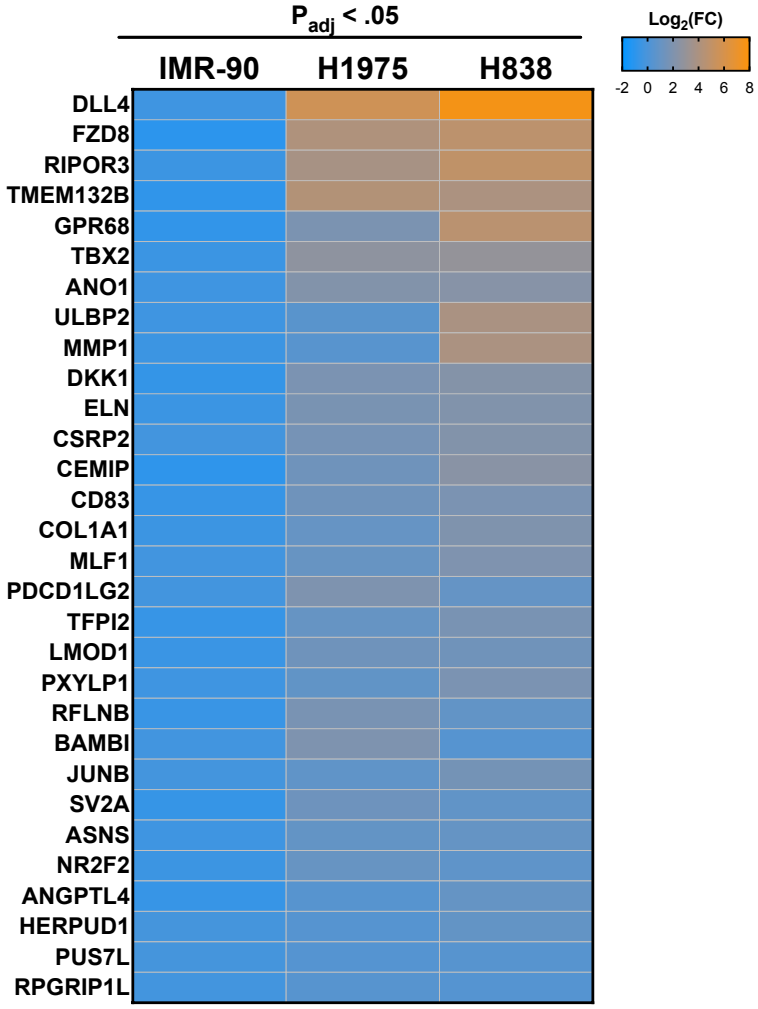
Supplementary figure 14. Cell viability and *CDKN1A*, *TCF3* and *CASP8AP2* mRNA expression levels upon simultaneous *TCF3* and *CASP8AP2* knockdown in NCI-H1975 and NCI-H838.
(A) CellTiter-Glo cell viability assay (CTG) to assess the effect of siPOOL-mediated knockdown of *TCF3* (siTCF3) on siC8AP2-mediated phenotype in H1975 and H838 cell lines at 96 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3). Cells were transfected simultaneously with indicated combinations of siPOOLs; cells transfected with siNC (20 nM) were used as normalization control.
(B, C, D) RT-qPCR to assess the mRNA levels of *TCF3*, *CDKN1A* and *CASP8AP2* in H1975 and H838 cells transfected with indicated combinations of siC8AP2, siTCF3 and siNC at 48 hours post-transfection (10 nM each siPOOL, total 20 nM siPOOLs per condition, n = 3). Cells transfected with siNC (20 nM) were used as normalization control.
For each biological replicate, data were normalized to the respective negative control signal and plotted as individual values. Bar heights represent mean, error bars represent range. Significance testing was performed calculated on log_2-transformed fold change data using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; "n.s.", not significant.



Supplementary figure 15. FACS analysis of IMR-90 cells transfected with combinations of siPOOLs against *CASP8AP2*, *CDKN1A* and *TP53*.

(A) IMR-90 cells were pulsed with BrdU ***(top panel)*** at 46 hours post-transfection with indicated combinations of siPOOLs against *CASP8AP2* (siC8AP2), *CDKN1A* (siCDKN1A) and negative control siPOOL (siNC). Cells were harvested after 2 hours (48 hours post-transfection), fixed, stained with propidium iodide (PI) ***(bottom panel)*** and analyzed by flow cytometry. Single cell-derived signals were gated, plotted and quantified using FlowJo v10 software. Rectangles represent gates used for quantification of BrdU-positive cell populations.

(B) Quantification data are represented as individual values per biological replicate, bar heights represent mean, error bars represent range (n = 3). Significance testing was performed using two-tailed paired Student's t-test, represented as ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; "n.s.", not significant.



Supplementary figure 16. Heatmap representing log_2(FC) of genes that were found significantly upregulated (log_2(FC) > 0, p_{adj} < 0.05) by RNA-Seq in both NCI-H1975 (H1975) and NCI-H838 (H838) cell lines and significantly downregulated (log_2(FC < 0, p_{adj} < 0.05) in IMR-90 cell lines upon siPOOL-mediated *CASP8AP2* (siC8AP2, 10 nM) knockdown compared to respective cell lines transfected with negative control siPOOL (siNC, 10 nM) at 48 hours post-transfection (n = 3). Log_2(FC) and p_{adj} were calculated using DESeq2 pipeline.

9 References

[1] Mattiuzzi, C., Lippi, G.: Current cancer epidemiology. *Journal of Epidemiology and Global Health*, **9** (4), 2019, p. 217–222.

[2] Ferlay, J. et al.: Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer*, **127** (12), 2010, p. 2893–2917.

[3] Schabath, M.B., Cote, M.L.: Cancer progress and priorities: Lung cancer. *Cancer Epidemiology Biomarkers and Prevention*, **28** (10), 2019, p. 1563–1579.

[4] Herbst, R.S. et al.: The biology and management of non-small cell lung cancer. *Nature*, **553** (7689), 2018, p. 446–454.

[5] Travis, W.D.: Lung Cancer Pathology. *Clinics in Chest Medicine*, **41** (1), 2020, p. 67–85.

[6] Molina, J.R. et al.: Non-Small Cell Lung Cancer: Epidemiology, Risk Factors, Treatment, and Survivorship. *Mayo Clinic Proceedings*, **83** (5), 2008, p. 584–594.

[7] U.S. Department of Health and Human Services: The Health Consequences of Smoking: A Report of the Surgeon General. *Atlanta (GA): Centers for Disease Control and Prevention (US)*, 2004.

[8] Pallis, A.G., Syrigos, K.N.: Lung cancer in never smokers: Disease characteristics and risk factors. *Critical Reviews in Oncology/Hematology*, **88** (3), 2013, p. 494–503.

[9] Sun, S. et al.: Lung cancer in never smokers — a different disease. *Nature Reviews Cancer*, **7** (10), 2007, p. 778–790.

[10] Barlesi, F. et al.: Routine molecular profiling of patients with advanced non-small-cell lung cancer: results of a 1-year nationwide programme of the French Cooperative Thoracic Intergroup (IFCT). *The Lancet*, **387** (10026), 2016, p. 1415–1426.

[11] Thompson, J.C. et al.: Detection of Therapeutically Targetable Driver and Resistance Mutations in Lung Cancer Patients by Next-Generation Sequencing of Cell-Free Circulating Tumor DNA. *Clinical Cancer Research*, **22** (23), 2016, p. 5772–5782.

[12] Dang, A.-T.H. et al.: Actionable Mutation Profiles of Non-Small Cell Lung Cancer patients from Vietnamese population. *Scientific Reports*, **10** (1), 2020, p. 2707.

[13] Li, S. et al.: Comprehensive Characterization of Oncogenic Drivers in Asian Lung Adenocarcinoma. *Journal of Thoracic Oncology*, **11** (12), 2016, p. 2129–2140.

[14] Skoulidis, F., Heymach, J.V.: Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nature Reviews Cancer*, **19** (9), 2019, p. 495–509.

[15] Campbell, J.D. et al.: Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nature Genetics*, **48** (6), 2016, p. 607–616.

[16] Unni, A.M. et al.: Evidence that synthetic lethality underlies the mutual exclusivity of oncogenic KRAS and EGFR mutations in lung adenocarcinoma. *eLife*, **4**, 2015, p. 1–23.

[17] Imielinski, M. et al.: Mapping the Hallmarks of Lung Adenocarcinoma with Massively Parallel Sequencing. *Cell*, **150** (6), 2012, p. 1107–1120.

[18] Collisson, E.A. et al.: Comprehensive molecular profiling of lung adenocarcinoma: The cancer genome atlas research network. *Nature*, **511** (7511), 2014, p. 543–550.

[19] Hirsch, F.R. et al.: Lung cancer: current therapies and new targeted treatments. *The Lancet*, **389** (10066), 2017, p. 299–311.

[20] Ellis, P.M., Vandermeer, R.: Delays in the diagnosis of lung cancer. *Journal of Thoracic Disease*, **3** (3), 2011, p. 183–188.

[21] Zarogoulidis, K. et al.: Treatment of non-small cell lung cancer (NSCLC). *Journal of thoracic disease*, **5** (Suppl 4), 2013, p. S389-96.

[22] Sjøgren, K. et al.: Timing of Severe Toxicity from Chemotherapy in Patients With Lung Cancer. *Anticancer research*, **40** (11), 2020, p. 6399–6406.

[23] Blackhall, F.H. et al.: Improving Survival and Reducing Toxicity with Chemotherapy in Advanced Non-Small Cell Lung Cancer. *Treatments in Respiratory Medicine*, **4** (2), 2005, p. 71–84.

[24] Horvath, L. et al.: Overcoming immunotherapy resistance in non-small cell lung cancer (NSCLC) - novel approaches and future outlook. *Molecular Cancer*, **19** (1), 2020, p. 141.

[25] Früh, M., Peters, S.: EGFR mutation subtype impacts efficacy of immune checkpoint inhibitors in non-small-cell lung cancer. *Annals of Oncology*, **30** (8), 2019, p. 1190–1192.

[26] Bruno, D., Dowlati, A.: Immunotherapy in EGFR mutant non-small cell lung cancer: when, who and how? *Translational Lung Cancer Research*, **8** (5), 2019, p. 710–714.

[27] Addeo, A. et al.: Immunotherapy in non-small cell lung cancer harbouring driver mutations. *Cancer Treatment Reviews*, **96**, 2021, p. 102179.

[28] Wang, H. et al.: Targeting Loss-of-Function Mutations in Tumor-Suppressor Genes as a Strategy for Development of Cancer Therapeutic Agents. *Seminars in Oncology*, **33** (4), 2006, p. 513–520.

[29] Weinstein, I.B.: Cancer. Addiction to oncogenes--the Achilles heal of cancer. *Science*, **297** (5578), 2002, p. 63–4.

[30] Weinstein, I.B., Joe, A.: Oncogene Addiction. *Cancer Research*, **68** (9), 2008, p. 3077–3080.

[31] Torti, D., Trusolino, L.: Oncogene addiction as a foundational rationale for targeted anti-cancer therapy: Promises and perils. *EMBO Molecular Medicine*, **3** (11), 2011, p. 623–636.

[32] Greulich, H.: The genomics of lung adenocarcinoma: Opportunities for targeted therapies. *Genes and Cancer*, **1** (12), 2010, p. 1200–1210.

[33] Smolle, E. et al.: Oncogene addiction and tumor mutational burden in non-small-cell lung cancer: Clinical significance and limitations. *Thoracic Cancer*, **11** (2), 2020, p. 205–215.

[34] Minuti, G. et al.: Targeted therapy for NSCLC with driver mutations. *Expert Opinion on Biological Therapy*, **13** (10), 2013, p. 1401–1412.

[35] Lackner, M.R. et al.: Mechanisms of acquired resistance to targeted cancer therapies. *Future Oncology*, **8** (8), 2012, p. 999–1014.

[36] Herbst, R.S.: Review of epidermal growth factor receptor biology. *International journal of radiation oncology, biology, physics*, **59** (2 Suppl), 2004, p. 21–6.

[37] Wang, Z.: ErbB Receptors and Cancer. *Methods in molecular biology (Clifton, N.J.)*, **1652**, 2017, p. 3–35.

[38] Sigismund, S. et al.: Emerging functions of the EGFR in cancer. *Molecular oncology*, **12** (1), 2018, p. 3–20.

[39] Brand, T.M. et al.: Nuclear EGFR as a molecular target in cancer. *Radiotherapy and Oncology*, **108** (3), 2013, p. 370–377.

[40] Hsu, W.-H. et al.: Overview of current systemic management of EGFR-mutant NSCLC. *Annals of Oncology*, **29** (suppl_1), 2018, p. i3–i9.

[41] Jänne, P.A. et al.: Epidermal Growth Factor Receptor Mutations in Non–Small-Cell Lung Cancer: Implications for Treatment and Tumor Biology. *Journal of Clinical Oncology*, **23** (14), 2005, p. 3227–3234.

[42] Shigematsu, H., Gazdar, A.: Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. *International journal of cancer*, **118** (2), 2006, p. 257–262.

[43] Chapman, A.M. et al.: Lung cancer mutation profile of EGFR, ALK, and KRAS: Meta-analysis and comparison of never and ever smokers. *Lung Cancer*, **102**, 2016, p. 122–134.

[44] Li, A.R. et al.: Clinical testing experience and relationship to EGFR gene copy number and immunohistochemical expression. *Journal of Molecular Diagnostics*, **10** (3), 2008, p. 242–248.

[45] Tokumo, M. et al.: The relationship between epidermal growth factor receptor mutations and clinicopathologic features in non-small cell lung cancers. *Clinical Cancer Research*, **11** (3), 2005, p. 1167–1173.

[46] Shigematsu, H. et al.: Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *Journal of the National Cancer Institute*, **97** (5), 2005, p. 339–346.

[47] Gazdar, A.F.: Personalized Medicine and Inhibition of EGFR Signaling in Lung Cancer. *New England Journal of Medicine*, **361** (10), 2009, p. 1018–1020.

[48] Wee, P., Wang, Z.: Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers*, **9** (5), 2017, p. 52.

[49] Jorissen, R.: Epidermal growth factor receptor: mechanisms of activation and signalling. *Experimental Cell Research*, **284** (1), 2003, p. 31–53.

[50] Sordella, R. et al.: Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science*, **305** (5687), 2004, p. 1163–1167.

[51] Takeda, M., Nakagawa, K.: First- and Second-Generation EGFR-TKIs Are All Replaced to Osimertinib in Chemo-Naive EGFR Mutation-Positive Non-Small Cell Lung Cancer? *International Journal of Molecular Sciences*, **20** (1), 2019, p. 146.

[52] Kobayashi, S. et al.: Compound EGFR Mutations and Response to EGFR Tyrosine Kinase Inhibitors. *Journal of Thoracic Oncology*, **8** (1), 2013, p. 118–122.

[53] Hayashi, H. et al.: Postprogression survival for first-line chemotherapy of patients with advanced non-small-cell lung cancer. *Annals of Oncology*, **23** (6), 2012, p. 1537–1541.

[54] Stewart, E.L. et al.: Known and putative mechanisms of resistance to EGFR targeted therapies in NSCLC patients with EGFR mutations—a review. *Translational Lung Cancer Research*, **4** (1), 2015, p. 67.

[55] Lin, Y. et al.: EGFR-TKI resistance in NSCLC patients: mechanisms and strategies. *American journal of cancer research*, **4** (5), 2014, p. 411–35.

[56] Oronsky, B. et al.: Navigating the "No Man's Land" of TKI-Failed EGFR-Mutated Non–Small Cell Lung Cancer (NSCLC): A Review. *Neoplasia*, **20** (1), 2018, p. 92–98.

[57] Santoni-Rugiu, E. et al.: Intrinsic resistance to EGFR-Tyrosine Kinase Inhibitors in EGFR-Mutant Non-Small Cell Lung Cancer: Differences and Similarities with Acquired Resistance. *Cancers*, **11** (7), 2019, p. 923.

[58] Liu, Q. et al.: EGFR-TKIs resistance via EGFR-independent signaling pathways. *Molecular Cancer*, **17** (1), 2018, p. 53.

[59] Westover, D. et al.: Mechanisms of acquired resistance to first- and second-generation EGFR tyrosine kinase inhibitors. *Annals of Oncology*, **29**, 2018, p. i10–i19.

[60] Leonetti, A. et al.: Resistance mechanisms to osimertinib in EGFR-mutated non-small cell lung cancer. *British Journal of Cancer*, **121** (9), 2019, p. 725–737.

[61] Popat, S.: Osimertinib as First-Line Treatment in *EGFR* -Mutated Non–Small-Cell Lung Cancer. *New England Journal of Medicine*, **378** (2), 2018, p. 192–193.

[62] Soria, J.-C. et al.: Osimertinib in Untreated EGFR-Mutated Advanced Non–Small-Cell Lung Cancer. *New England Journal of Medicine*, **378** (2), 2017, p. 113–125.

[63] Ramalingam, S.S. et al.: Overall Survival with Osimertinib in Untreated, EGFR-Mutated Advanced NSCLC. *New England Journal of Medicine*, **382** (1), 2019, p. 41–50.

[64] Tang, Z.-H., Lu, J.-J.: Osimertinib resistance in non-small cell lung cancer: Mechanisms and therapeutic strategies. *Cancer Letters*, **420**, 2018, p. 242–246.

[65] Schmid, S. et al.: Mechanisms of osimertinib resistance and emerging treatment options. *Lung Cancer*, **147**, 2020, p. 123–129.

[66] Planchard, D. et al.: EGFR-independent mechanisms of acquired resistance to AZD9291 in EGFR T790M-positive NSCLC patients. *Annals of Oncology*, **26** (10), 2015, p. 2073–2078.

[67] Castellano, E., Santos, E.: Functional Specificity of Ras Isoforms: So Similar but So Different. *Genes & Cancer*, **2** (3), 2011, p. 216–231.

[68] Simanshu, D.K. et al.: RAS Proteins and Their Regulators in Human Disease. *Cell*, **170** (1), 2017, p. 17.

[69] Gimple, R.C., Wang, X.: RAS: Striking at the Core of the Oncogenic Circuitry. *Frontiers in Oncology*, **9**, 2019, p. 965.

[70] Tao, J. et al.: Advancement in research and therapy of NF1 mutant malignant tumors. *Cancer Cell International 2020 20:1*, **20** (1), 2020, p. 1–8.

[71] Redig, A.J. et al.: Clinical and Molecular Characteristics of *NF1* -Mutant Lung Cancer. *Clinical Cancer Research*, **22** (13), 2016, p. 3148–3156.

[72] Hunter, J.C. et al.: Biochemical and Structural Analysis of Common Cancer-Associated KRAS Mutations. *Molecular Cancer Research*, **13** (9), 2015, p. 1325–1335.

[73] Friedlaender, A. et al.: KRAS as a druggable target in NSCLC: Rising like a phoenix after decades of development failures. *Cancer Treatment Reviews*, **85**, 2020, p. 101978.

[74] Adderley, H. et al.: KRAS-mutant non-small cell lung cancer: Converging small molecules and immune checkpoint inhibition. *EBioMedicine*, **41**, 2019, p. 711–716.

[75] Muñoz-Maldonado, C. et al.: A Comparative Analysis of Individual RAS Mutations in Cancer Biology. *Frontiers in oncology*, **9**, 2019, p. 1088.

[76] Yang, H. et al.: New Horizons in KRAS-Mutant Lung Cancer: Dawn After Darkness. *Frontiers in Oncology*, **9**, 2019, p. 953.

[77] Prior, I.A. et al.: A Comprehensive Survey of Ras Mutations in Cancer. *Cancer Research*, **72** (10), 2012, p. 2457–2467.

[78] Singh, A. et al.: A Gene Expression Signature Associated with "K-Ras Addiction" Reveals Regulators of EMT and Tumor Cell Survival. *Cancer Cell*, **15** (6), 2009, p. 489–500.

[79] Mullard, A.: Cracking KRAS. *Nature reviews. Drug discovery*, **18** (12), 2019, p. 887–891.

[80] Skoulidis, F. et al.: Sotorasib for Lung Cancers with KRAS p.G12C Mutation. *New England Journal of Medicine*, **384** (25), 2021, p. 2371–2381.

[81] Molina-Arcas, M. et al.: Drugging the Undruggable: Advances on RAS Targeting in Cancer. *Genes*, **12** (6), 2021, p. 899.

[82] Ryan, M.B., Corcoran, R.B.: Therapeutic strategies to target RAS -mutant cancers. *Nature Reviews Clinical Oncology 2018 15:11*, **15** (11), 2018, p. 709–720.

[83] Ross, S.J. et al.: Targeting KRAS-dependent tumors with AZD4785, a high-affinity therapeutic antisense oligonucleotide inhibitor of KRAS. *Science Translational Medicine*, **9** (394), 2017, p. eaal5253.

[84] Kastenhuber, E.R., Lowe, S.W.: Putting p53 in Context. *Cell*, **170** (6), 2017, p. 1062–1078.

[85] Sullivan, K.D. et al.: Mechanisms of transcriptional regulation by p53. *Cell Death & Differentiation*, **25** (1), 2018, p. 133–143.

[86] Horn, H.F., Vousden, K.H.: Coping with stress: multiple ways to activate p53. *Oncogene*, **26** (9), 2007, p. 1306–1316.

[87] Zhang, C. et al.: Gain-of-function mutant p53 in cancer progression and therapy. *Journal of Molecular Cell Biology*, **12** (9), 2020, p. 674–687.

[88] Freed-Pastor, W.A., Prives, C.: Mutant p53: One name, many proteins. *Genes and Development*, **26** (12), 2012, p. 1268–1286.

[89] Mogi, A., Kuwano, H.: TP53 Mutations in Nonsmall Cell Lung Cancer. *Journal of Biomedicine and Biotechnology*, **2011**, 2011, p. 1–9.

[90] Stein, Y. et al.: Gain-of-Function Mutant p53: All the Roads Lead to Tumorigenesis. *International Journal of Molecular Sciences*, **20** (24), 2019, p. 6197.

[91] Schulz-Heddergott, R., Moll, U.: Gain-of-Function (GOF) Mutant p53 as Actionable Therapeutic Target. *Cancers*, **10** (6), 2018, p. 188.

[92] Yuan, M. et al.: The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduction and Targeted Therapy*, **4** (1), 2019, p. 61.

[93] Chan, B.A., Hughes, B.G.M.: Targeted therapy for non-small cell lung cancer: current standards and the promise of the future. *Translational lung cancer research*, **4** (1), 2015, p. 36–54.

[94] Sehgal, K. et al.: Targeting ROS1 rearrangements in non-small cell lung cancer with crizotinib and other kinase inhibitors. *Translational cancer research*, **7** (Suppl 7), 2018, p. S779.

[95] Hallberg, B., Palmer, R.H.: Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. *Nature Reviews Cancer 2013 13:10*, **13** (10), 2013, p. 685–700.

[96] Takeuchi, K. et al.: RET, ROS1 and ALK fusions in lung cancer. *Nature Medicine*, **18** (3), 2012, p. 378–381.

[97] Katayama, R.: Drug resistance in anaplastic lymphoma kinase-rearranged lung cancer. *Cancer Science*, **109** (3), 2018, p. 572–580.

[98] Mizuta, H. et al.: Gilteritinib overcomes lorlatinib resistance in ALK-rearranged cancer. *Nature Communications*, **12** (1), 2021, p. 1261.

[99] McCoach, C.E. et al.: Resistance Mechanisms to Targeted Therapies in ROS1+ and ALK+ Non–small Cell Lung Cancer. *Clinical Cancer Research*, **24** (14), 2018, p. 3334–3347.

[100] Kawakami, H. et al.: Targeting MET Amplification as a New Oncogenic Driver. *Cancers*, **6** (3), 2014, p. 15–26.

[101] Ma, G. et al.: The Prognostic Role of MET Protein Expression Among Surgically Resected Non-small Cell Lung Cancer Patients: A Meta-Analysis. *Frontiers in Oncology*, **9**, 2019, p. 1441.

[102] Tong, J.H. et al.: MET amplification and exon 14 splice site mutation define unique molecular subgroups of non-small cell lung carcinoma with poor prognosis. *Clinical Cancer Research*, **22** (12), 2016, p. 3048–3056.

[103] Wolf, J. et al.: Capmatinib in MET Exon 14–Mutated or MET-Amplified Non–Small-Cell Lung Cancer. *New England Journal of Medicine*, **383** (10), 2020, p. 944–957.

[104] Dhillon, S.: Capmatinib: First Approval. *Drugs*, **80** (11), 2020, p. 1125–1131.

[105] Wang, Q. et al.: MET inhibitors for targeted therapy of EGFR TKI-resistant lung cancer. *Journal of Hematology & Oncology*, **12** (1), 2019, p. 63.

[106] Kim, S. et al.: Acquired Resistance of MET-Amplified Non-small Cell Lung Cancer Cells to the MET Inhibitor Capmatinib. *Cancer research and treatment*, **51** (3), 2019, p. 951–962.

[107] Leonetti, A. et al.: BRAF in non-small cell lung cancer (NSCLC): Pickaxing another brick in the wall. *Cancer Treatment Reviews*, **66**, 2018, p. 82–94.

[108] Planchard, D. et al.: Dabrafenib plus trametinib in patients with previously untreated BRAFV600E-mutant metastatic non-small-cell lung cancer: an open-label, phase 2 trial. *The Lancet Oncology*, **18** (10), 2017, p. 1307–1316.

[109] Odogwu, L. et al.: FDA Approval Summary: Dabrafenib and Trametinib for the Treatment of Metastatic Non-Small Cell Lung Cancers Harboring BRAF V600E Mutations . *The Oncologist*, **23** (6), 2018, p. 740–745.

[110] Han, J. et al.: MEK inhibitors for the treatment of non-small cell lung cancer. *Journal of Hematology & Oncology*, **14** (1), 2021, p. 1.

[111] Facchinetti, F. et al.: Molecular mechanisms of resistance to BRAF and MEK inhibitors in BRAFV600E non–small cell lung cancer. *European Journal of Cancer*, **132**, 2020, p. 211–223.

[112] Metzenmacher, M. et al.: Acquired Resistance to BRAF/MEK Inhibitor Therapy in BRAF-V600-mutated Squamous Cell Lung Cancer: Concurrent Evolvement of PTEN and MEK1 Mutations. *Clinical Lung Cancer*, **S1525-7304** (20), 2020, p. 30341–30347.

[113] O'Neil, N.J. et al.: Synthetic lethality and cancer. *Nature Reviews Genetics*, **18** (10), 2017, p. 613–623.

[114] Shimomura, I. et al.: Synthetic Lethality in Lung Cancer—From the Perspective of Cancer Genomics. *Medicines*, **6** (1), 2019, p. 38.

[115] Leung, A.W.Y. et al.: Synthetic lethality in lung cancer and translation to clinical therapies. *Molecular Cancer*, **15** (1), 2016, p. 61.

[116] Iorio, F. et al.: A Landscape of Pharmacogenomic Interactions in Cancer. *Cell*, **166** (3), 2016, p. 740–754.

[117] Luo, J. et al.: Principles of Cancer Therapy: Oncogene and Non-oncogene Addiction. *Cell*, **136** (5), 2009, p. 823.

[118] Huang, A. et al.: Synthetic lethality as an engine for cancer drug target discovery. *Nature Reviews Drug Discovery*, **19** (1), 2020, p. 23–38.

[119] Pan, R. et al.: Synthetic Lethality of Combined Bcl-2 Inhibition and p53 Activation in AML: Mechanisms and Superior Antileukemic Efficacy. *Cancer Cell*, **32** (6), 2017, p. 748-760.e6.

[120] Lord, C.J., Ashworth, A.: PARP inhibitors: Synthetic lethality in the clinic. *Science*, **355** (6330), 2017, p. 1152–1158.

[121] Lee, J.M. et al.: PARP inhibitors for BRCA1/2 mutation-associated and BRCA-like malignancies. *Annals of Oncology*, **25** (1), 2014, p. 32–40.

[122] Morice, P.-M. et al.: Identifying patients eligible for PARP inhibitor treatment: from NGS-based tests to 3D functional assays. *British Journal of Cancer*, **125** (1), 2021, p. 7–14.

[123] Kopetz, S. et al.: Encorafenib, Binimetinib, and Cetuximab in *BRAF* V600E–Mutated Colorectal Cancer. *New England Journal of Medicine*, **381** (17), 2019, p. 1632–1643.

[124] Kopetz, S. et al.: Global BRAF testing practices in metastatic colorectal cancer. *Journal of Clinical Oncology*, **39** (3_suppl), 2021, p. 128–128.

[125] Gao, S., Lai, L.: Synthetic lethality in drug development: the dawn is coming. *Future Medicinal Chemistry*, **10** (18), 2018, p. 2129–2132.

[126] Kelly, M.R. et al.: Combined Proteomic and Genetic Interaction Mapping Reveals New RAS Effector Pathways and Susceptibilities. *Cancer Discovery*, **10** (12), 2020, p. 1950–1967.

[127] Kumar, M.S. et al.: The GATA2 transcriptional network is requisite for RAS oncogene-driven non-small cell lung cancer. *Cell*, **149** (3), 2012, p. 642–655.

[128] Puyol, M. et al.: A Synthetic Lethal Interaction between K-Ras Oncogenes and Cdk4 Unveils a Therapeutic Strategy for Non-small Cell Lung Carcinoma. *Cancer Cell*, **18** (1), 2010, p. 63–73.

[129] Vicent, S. et al.: Wilms tumor 1 (WT1) regulates KRAS-driven oncogenesis and senescence in mouse and human models. *Journal of Clinical Investigation*, **120** (11), 2010, p. 3940–3952.

[130] Barbie, D.A. et al.: Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*, **462** (7269), 2009, p. 108–112.

[131] Luo, T. et al.: STK33 kinase inhibitor BRD-8899 has no effect on KRAS-dependent cancer cell viability. *Proceedings of the National Academy of Sciences*, **109** (8), 2012, p. 2860–2865.

[132] Scholl, C. et al.: Synthetic Lethal Interaction between Oncogenic KRAS Dependency and STK33 Suppression in Human Cancer Cells. *Cell*, **137** (5), 2009, p. 821–834.

[133] Lantermann, A.B. et al.: Inhibition of Casein Kinase 1 Alpha Prevents Acquired Drug Resistance to Erlotinib in EGFR-Mutant Non–Small Cell Lung Cancer. *Cancer Research*, **75** (22), 2015, p. 4937–4948.

[134] Zeng, H. et al.: Genome-wide CRISPR screening reveals genetic modifiers of mutant EGFR dependence in human NSCLC. *eLife*, **8**, 2019, p. e50223.

[135] Bivona, T.G. et al.: FAS and NF-κB signalling modulate dependence of lung cancers on mutant EGFR. *Nature 2011 471:7339*, **471** (7339), 2011, p. 523–526.

[136] Bruin, E.C. de et al.: Reduced NF1 Expression Confers Resistance to EGFR Inhibition in Lung Cancer. *Cancer Discovery*, **4** (5), 2014, p. 606–619.

[137] Sharifnia, T. et al.: Genetic modifiers of EGFR dependence in non-small cell lung cancer. *Proceedings of the National Academy of Sciences*, **111** (52), 2014, p. 18661–18666.

[138] Casás-Selves, M. et al.: Tankyrase and the Canonical Wnt Pathway Protect Lung Cancer Cells from EGFR Inhibition. *Cancer Research*, **72** (16), 2012, p. 4154–4164.

[139] Vyse, S. et al.: Exploiting Synthetic Lethality and Network Biology to Overcome EGFR Inhibitor Resistance in Lung Cancer. *Journal of Molecular Biology*, **429** (12), 2017, p. 1767–1786.

[140] Kim, J. et al.: Bioinformatics-driven discovery of rational combination for overcoming EGFR-mutant lung cancer resistance to EGFR therapy. *Bioinformatics*, **30** (17), 2014, p. 2393–2398.

[141] Liao, S. et al.: A genetic interaction analysis identifies cancer drivers that modify EGFR dependency. *Genes & Development*, **31** (2), 2017, p. 184–196.

[142] Djebali, S. et al.: Landscape of transcription in human cells. *Nature*, **489** (7414), 2012, p. 101–108.

[143] Derrien, T. et al.: The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Research*, **22** (9), 2012, p. 1775–1789.

[144] Mercer, T.R. et al.: Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nature Biotechnology 2011 30:1*, **30** (1), 2011, p. 99–104.

[145] Kopp, F., Mendell, J.T.: Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell*, **172** (3), 2018, p. 393–407.

[146] Rinn, J.L., Chang, H.Y.: Genome regulation by long noncoding RNAs. *Annual review of biochemistry*, **81**, 2012, p. 145–166.

[147] Goyal, A. et al.: Challenges of CRISPR/Cas9 applications for long non-coding RNA genes. *Nucleic Acids Research*, **45** (3), 2016, p. gkw883.

[148] Seiler, J. et al.: The lncRNA VELUCT strongly regulates viability of lung cancer cells despite its extremely low abundance. *Nucleic Acids Research*, **45** (9), 2017, p. 5458–5469.

[149] Ginn, L. et al.: LncRNAs in Non-Small-Cell Lung Cancer. *Non-Coding RNA*, **6** (3), 2020, p. 25.

[150] Chi et al.: Long Non-Coding RNA in the Pathogenesis of Cancers. *Cells*, **8** (9), 2019, p. 1015.

[151] Schmitz, S.U. et al.: Mechanisms of long noncoding RNA function in development and disease. *Cellular and Molecular Life Sciences*, **73** (13), 2016, p. 2491–2509.

[152] Schmitt, A.M., Chang, H.Y.: Long Noncoding RNAs in Cancer Pathways. *Cancer Cell*, **29** (4), 2016, p. 452–463.

[153] Huarte, M.: The emerging role of lncRNAs in cancer. *Nature Medicine*, **21** (11), 2015, p. 1253–1261.

[154] Gutschner, T., Diederichs, S.: The hallmarks of cancer. *RNA Biology*, **9** (6), 2012, p. 703–719.

[155] Ji, P. et al.: MALAT-1, a novel noncoding RNA, and thymosin β4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, **22** (39), 2003, p. 8031–8041.

[156] Gutschner, T. et al.: The Noncoding RNA *MALAT1* Is a Critical Regulator of the Metastasis Phenotype of Lung Cancer Cells. *Cancer Research*, **73** (3), 2013, p. 1180–1189.

[157] Ricciuti, B. et al.: Long noncoding RNAs: new insights into non-small cell lung cancer biology, diagnosis and therapy. *Medical Oncology*, **33** (2), 2016, p. 18.

[158] Zeng, Z. et al.: AFAP1-AS1, a long noncoding RNA upregulated in lung cancer and promotes invasion and metastasis. *Tumor Biology*, **37** (1), 2016, p. 729–737.

[159] Kung, J.T.Y. et al.: Long Noncoding RNAs: Past, Present, and Future. *Genetics*, **193** (3), 2013, p. 651–669.

[160] Wilusz, J.E. et al.: Long noncoding RNAs: functional surprises from the RNA world. *Genes & Development*, **23** (13), 2009, p. 1494–1504.

[161] Nötzold, L. et al.: The long non-coding RNA LINC00152 is essential for cell cycle progression through mitosis in HeLa cells. *Scientific Reports*, **7** (1), 2017, p. 2265.

[162] Roth, A. et al.: Targeting LINC00673 expression triggers cellular senescence in lung cancer. *RNA Biology*, **15** (12), 2018, p. 1499–1511.

[163] Larson, M.H. et al.: CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols*, **8** (11), 2013, p. 2180–2196.

[164] Gilbert, L.A. et al.: Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, **159** (3), 2014, p. 647–661.

[165] Liu, S.J. et al.: CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science*, **355** (6320), 2017, p. aah7111.

[166] Liu, S. et al.: Wnt-regulated lncRNA discovery enhanced by in vivo identification and CRISPRi functional validation. *Genome Medicine*, **12** (1), 2020, p. 89.

[167] Hartenian, E., Doench, J.G.: Genetic screens and functional genomics using CRISPR/Cas9 technology. *FEBS Journal*, **282** (8), 2015, p. 1383–1393.

[168] Schuster, A. et al.: RNAi/CRISPR Screens: from a Pool to a Valid Hit. *Trends in Biotechnology*, **37** (1), 2019, p. 38–55.

[169] Haley, B., Roudnicky, F.: Functional Genomics for Cancer Drug Target Discovery. *Cancer Cell*, **38** (1), 2020, p. 31–43.

[170] Taylor, J., Woodcock, S.: A Perspective on the Future of High-Throughput RNAi Screening. *Journal of Biomolecular Screening*, **20** (8), 2015, p. 1040–1051.

[171] Schuster, S. et al.: Antiviral RNAi in Insects and Mammals: Parallels and Differences. *Viruses*, **11** (5), 2019, p. 448.

[172] Wilson, R.C., Doudna, J.A.: Molecular Mechanisms of RNA Interference. *Annual Review of Biophysics*, **42** (1), 2013, p. 217–239.

[173] Elbashir, S.M. et al.: Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411** (6836), 2001, p. 494–498.

[174] Mohr, S.E., Perrimon, N.: RNAi screening: new approaches, understandings, and organisms. *Wiley Interdisciplinary Reviews: RNA*, **3** (2), 2012, p. 145–158.

[175] Burgess, D.J.: Shining a light on genetic screen strategies. *Nature Reviews Genetics*, **19** (1), 2018, p. 6–7.

[176] Jackson, A.L., Linsley, P.S.: Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nature Reviews Drug Discovery*, **9** (1), 2010, p. 57–67.

[177] Hannus, M. et al.: siPools: highly complex but accurately defined siRNA pools eliminate off-target effects. *Nucleic acids research*, **42** (12), 2014, p. 8049–8061.

[178] Jackson, A.L. et al.: Widespread siRNA “off-target” transcript silencing mediated by seed region sequence complementarity. *RNA*, **12** (7), 2006, p. 1179–1187.

[179] Cabili, M.N. et al.: Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biology*, **16** (1), 2015, p. 20.

[180] Lennox, K.A., Behlke, M.A.: Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. *Nucleic Acids Research*, **44** (2), 2016, p. 863–877.

[181] Zhao, Y. et al.: Challenges and Strategies in Ascribing Functions to Long Noncoding RNAs. *Cancers*, **12** (6), 2020, p. 1–21.

[182] Kornienko, A.E. et al.: Gene regulation by the act of long non-coding RNA transcription. *BMC Biology*, **11** (1), 2013, p. 59.

[183] Ørom, U.A. et al.: Long noncoding RNAs with enhancer-like function in human cells. *Cell*, **143** (1), 2010, p. 46–58.

[184] Engreitz, J.M. et al.: Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature 2016 539:7629*, **539** (7629), 2016, p. 452–455.

[185] Mali, P. et al.: RNA-guided human genome engineering via Cas9. *Science*, **339** (6121), 2013, p. 823–826.

[186] Shalem, O. et al.: Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, **343** (6166), 2014, p. 84–87.

[187] Wang, T. et al.: Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, **343** (6166), 2014, p. 80–84.

[188] Jinek, M. et al.: A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, **337** (6096), 2012, p. 816–821.

[189] LaManna, C.M. et al.: Sharing the CRISPR Toolbox with an Expanding Community. *The CRISPR Journal*, **3** (4), 2020, p. 248–252.

[190] Jinek, M. et al.: RNA-programmed genome editing in human cells. *eLife*, **2** (2), 2013, p. e00471.

[191] Cong, L. et al.: Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, **339** (6121), 2013, p. 819–823.

[192] Esvelt, K.M. et al.: Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nature Methods*, **10** (11), 2013, p. 1116–1121.

[193] Sanson, K.R. et al.: Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nature Communications*, **9** (1), 2018, p. 5416.

[194] Doench, J.G.: Am I ready for CRISPR? A user's guide to genetic screens. *Nature Reviews Genetics*, **19** (2), 2018, p. 67–80.

[195] Liu, Y. et al.: Genome-wide screening for functional long noncoding RNAs in human cells by Cas9 targeting of splice sites. *Nature Biotechnology*, **36** (12), 2018, p. 1203–1210.

[196] Ho, T.T. et al.: Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines. *Nucleic Acids Research*, **43** (3), 2015, p. e17.

[197] Aparicio-Prat, E. et al.: DECKO: Single-oligo, dual-CRISPR deletion of genomic elements including long non-coding RNAs. *BMC Genomics*, **16** (1), 2015, p. 846.

[198] Zhu, S. et al.: Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR–Cas9 library. *Nature Biotechnology*, **34** (12), 2016, p. 1279–1286.

[199] Horlbeck, M.A. et al.: Fitness effects of CRISPR/Cas9-targeting of long noncoding RNA genes. *Nature Biotechnology 2020 38:5*, **38** (5), 2020, p. 573–576.

[200] Qi, L.S. et al.: Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, **152** (5), 2013, p. 1173–1183.

[201] Gilbert, L.A. et al.: CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, **154** (2), 2013, p. 442–451.

[202] Thakore, P.I. et al.: Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nature Methods*, **12** (12), 2015, p. 1143–1149.

[203] Mirabelli et al.: Cancer Cell Lines Are Useful Model Systems for Medical Research. *Cancers*, **11** (8), 2019, p. 1098.

[204] Tate, J.G. et al.: COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research*, **47** (D1), 2019, p. D941–D947.

[205] Pal, J. et al.: Systematic analysis of migration factors by MigExpress identifies essential cell migration control genes in non-small cell lung cancer. *Molecular Oncology*, **15** (7), 2021, p. 1797–1817.

[206] Gao, J. et al.: Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Science Signaling*, **6** (269), 2013, p. pl1–pl1.

[207] Cerami, E. et al.: The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discovery*, **2** (5), 2012, p. 401–404.

[208] Howe, K.L. et al.: Ensembl 2021. *Nucleic Acids Research*, **49** (D1), 2021, p. D884–D891.

[209] Afgan, E. et al.: The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, **46** (W1), 2018, p. W537–W544.

[210] Li, J. et al.: TANRIC: An interactive open platform to explore the function of lncRNAs in cancer. *Cancer Research*, **75** (18), 2015, p. 3728–3737.

[211] Hart, T. et al.: High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell*, **163** (6), 2015, p. 1515–1526.

[212] Martin, M.: Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **17** (1), 2011, p. 10–12.

[213] Li, W. et al.: MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biology*, **15** (12), 2014, p. 554.

[214] Andrews, S., others: FastQC: a quality control tool for high throughput sequence data. *https://www.bioinformatics.babraham.ac.uk/projects/fastqc/*, 2010.

[215] Kim, D. et al.: HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, **12** (4), 2015, p. 357–360.

[216] Liao, Y. et al.: FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30** (7), 2014, p. 923–930.

[217] Love, M.I. et al.: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15** (12), 2014, p. 550.

[218] Gao, Z. et al.: Mutation of nucleotides around the +1 position of type 3 polymerase III promoters: The effect on transcriptional activity and start site usage. *Transcription*, **8** (5), 2017, p. 275–287.

[219] Arao, T. et al.: Small in-frame deletion in the epidermal growth factor receptor as a target for ZD6474. *Cancer Research*, **64** (24), 2004, p. 9101–9104.

[220] Park, M.-Y. et al.: Generation of lung cancer cell lines harboring EGFR T790M mutation by CRISPR/Cas9-mediated genome editing. *Oncotarget*, **8** (22), 2017, p. 36331–36338.

[221] Han, H.S. et al.: Detection of EGFR mutation status in lung adenocarcinoma specimens with different proportions of tumor cells using two methods of differential sensitivity. *Journal of Thoracic Oncology*, **7** (2), 2012, p. 355–364.

[222] Ng, P.K.S. et al.: Systematic Functional Annotation of Somatic Mutations in Cancer. *Cancer Cell*, **33** (3), 2018, p. 450–462.

[223] Kopp, F., Mendell, J.T.: Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell*, **172** (3), 2018, p. 393–407.

[224] Forrest, A.R.R. et al.: A promoter-level mammalian expression atlas. *Nature*, **507** (7493), 2014, p. 462–470.

[225] Shi, Y.: Mechanisms of Caspase Activation and Inhibition during Apoptosis. *Molecular Cell*, **9** (3), 2002, p. 459–470.

[226] Makino, H. et al.: A selective inhibition of c-Fos/activator protein-1 as a potential therapeutic target for intervertebral disc degeneration and associated pain. *Scientific Reports*, **7** (1), 2017, p. 16983.

[227] Kamide, D. et al.: Selective activator protein-1 inhibitor T-5224 prevents lymph node metastasis in an oral cancer model. *Cancer Science*, **107** (5), 2016, p. 666–673.

[228] Tsuchida, K. et al.: Design, synthesis, and biological evaluation of new cyclic disulfide decapeptides that inhibit the binding of AP-1 to DNA. *Journal of Medicinal Chemistry*, **47** (17), 2004, p. 4239–4246.

[229] Yin, Z.: Mechanism of Small Molecules Inhibiting Activator Protein-1 DNA Binding Probed with Induced Fit Docking and Metadynamics Simulations. *Journal of Chemical Information and Modeling*, **59** (12), 2019, p. 5276–5280.

[230] Horlbeck, M. et al.: Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife*, **5**, 2016, p. e19760.

[231] Vasseur, M. le et al.: Genome-wide CRISPRi screening identifies OCIAD1 as a prohibitin client and regulatory determinant of mitochondrial Complex III assembly in human cells. *eLife*, **10**, 2021, p. e67624.

[232] Bassaganyas, L. et al.: New factors for protein transport identified by a genome-wide CRISPRi screen in mammalian cells. *The Journal of cell biology*, **218** (11), 2019, p. 3861–3879.

[233] Tian, R. et al.: CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-Derived Neurons. *Neuron*, **104** (2), 2019, p. 239-255.e12.

[234] Tian, T. et al.: Genome-wide CRISPRi/a screens in human neurons link lysosomal failure to ferroptosis. *Nature neuroscience*, **24** (7), 2021, p. 1020–1034.

[235] Rosenbluh, J. et al.: Complementary information derived from CRISPR Cas9 mediated gene deletion and suppression. *Nature Communications*, **8** (1), 2017, p. 15403.

[236] Horlbeck, M. et al.: Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *eLife*, **5**, 2016, p. e12677.

[237] Yin, D. et al.: Long noncoding RNA AFAP1-AS1 predicts a poor prognosis and regulates non–small cell lung cancer cell proliferation by epigenetically repressing p21 expression. *Molecular Cancer*, **17** (1), 2018, p. 92.

[238] Huang, N. et al.: LncRNA AFAP1-AS1 Supresses miR-139-5p and Promotes Cell Proliferation and Chemotherapy Resistance of Non-small Cell Lung Cancer by Competitively Upregulating RRM2. *Frontiers in Oncology*, **9**, 2019, p. 1103.

[239] Tang, X. et al.: lncRNA AFAP1-AS1 Promotes Migration and Invasion of Non-Small Cell Lung Cancer via Up-Regulating IRF7 and the RIG-I-Like Receptor Signaling Pathway. *Cellular Physiology and Biochemistry*, **50** (1), 2018, p. 179–195.

[240] Zhang, F. et al.: AFAP1-AS1: A novel oncogenic long non-coding RNA in human cancers. *Cell proliferation*, **51** (1), 2018, p. e12397.

[241] Putkey, F.R. et al.: Unstable kinetochore-microtubule capture and chromosomal instability following deletion of CENP-E. *Developmental Cell*, **3** (3), 2002, p. 351–365.

[242] Kapoor, T.M. et al.: Chromosomes can congress to the metaphase plate before biorientation. *Science*, **311** (5759), 2006, p. 388–391.

[243] Craske, B., Welburn, J.P.I.: Leaving no-one behind: how CENP-E facilitates chromosome alignment. *Essays in Biochemistry*, **64** (2), 2020, p. 313–324.

[244] McEwen, B.F. et al.: CENP-E is essential for reliable bioriented spindle attachment, but chromosome alignment can be achieved via redundant mechanisms in mammalian cells. *Molecular Biology of the Cell*, **12** (9), 2001, p. 2776–2789.

[245] Tovini, L., McClelland, S.: Impaired CENP-E Function Renders Large Chromosomes More Vulnerable to Congression Failure. *Biomolecules*, **9** (2), 2019, p. 44.

[246] Wood, K.W. et al.: Centromere-Associated Protein E: A Motor That Puts the Brakes on the Mitotic Checkpoint: Fig. 1. *Clinical Cancer Research*, **14** (23), 2008, p. 7588–7592.

[247] Raaijmakers, J.A. et al.: BUB1 Is Essential for the Viability of Human Cells in which the Spindle Assembly Checkpoint Is Compromised. *Cell Reports*, **22** (6), 2018, p. 1424–1438.

[248] Yao, X. et al.: CENP-E forms a link between attachment of spindle microtubules to kinetochores and the mitotic checkpoint. *Nature Cell Biology*, **2** (8), 2000, p. 484–491.

[249] Weaver, B.A.A. et al.: Centromere-associated protein-E is essential for the mammalian mitotic checkpoint to prevent aneuploidy due to single chromosome loss. *Journal of Cell Biology*, **162** (4), 2003, p. 551–563.

[250] Foley, E.A., Kapoor, T.M.: Microtubule attachment and spindle assembly checkpoint signalling at the kinetochore. *Nature Reviews Molecular Cell Biology*, **14** (1), 2013, p. 25–37.

[251] Abrieu, A. et al.: CENP-E as an essential component of the mitotic checkpoint in vitro. *Cell*, **102** (6), 2000, p. 817–826.

[252] Wood, K.W. et al.: Antitumor activity of an allosteric inhibitor of centromere-associated protein-E. *Proceedings of the National Academy of Sciences of the United States of America*, **107** (13), 2010, p. 5839–5844.

[253] Schaar, B.T. et al.: CENP-E function at kinetochores is essential for chromosome alignment. *Journal of Cell Biology*, **139** (6), 1997, p. 1373–1382.

[254] Maia, A.F. et al.: Aurora B kinase cooperates with CENP-E to promote timely anaphase onset. *Chromosoma*, **119** (4), 2010, p. 405–413.

[255] Iegiani, G. et al.: CENPE inhibition leads to mitotic catastrophe and DNA damage in medulloblastoma cells. *Cancers*, **13** (5), 2021, p. 1028.

[256] Ohashi, A. et al.: A novel time-dependent CENP-E inhibitor with potent antitumor activity. *PLoS ONE*, **10** (12), 2015, p. e0144675.

[257] Balamuth, N.J. et al.: Serial transcriptome analysis and cross-species integration identifies centromere-associated protein e as a novel neuroblastoma target. *Cancer Research*, **70** (7), 2010, p. 2749–2758.

[258] Hisamatsu, Y. et al.: Effect of EGFR and p-AKT Overexpression on Chromosomal Instability in Gastric Cancer. *Annals of Surgical Oncology*, **23** (6), 2016, p. 1986–1992.

[259] Nahar, R. et al.: Elucidating the genomic architecture of Asian EGFR-mutant lung adenocarcinoma through multi-region exome sequencing. *Nature Communications*, **9** (1), 2018, p. 216.

[260] Marcar, L. et al.: Acquired Resistance of EGFR-Mutated Lung Cancer to Tyrosine Kinase Inhibitor Treatment Promotes PARP Inhibitor Sensitivity. *Cell Reports*, **27** (12), 2019, p. 3422–3432.

[261] Pfäffle, H.N. et al.: EGFR-activating mutations correlate with a fanconi anemia-like cellular phenotype that includes PARP inhibitor sensitivity. *Cancer Research*, **73** (20), 2013, p. 6254–6263.

[262] Mayes, P.A. et al.: Mitogen-activated protein kinase (MEK/ERK) inhibition sensitizes cancer cells to centromere-associated protein e inhibition. *International Journal of Cancer*, **132** (3), 2013, p. E149.

[263] Roberts, P.J., Der, C.J.: Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene*, **26** (22), 2007, p. 3291–3310.

[264] Hung, L.Y. et al.: Nuclear epidermal growth factor receptor (EGFR) interacts with signal transducer and activator of transcription 5 (STAT5) in activating Aurora-A gene expression. *Nucleic Acids Research*, **36** (13), 2008, p. 4337–4351.

[265] Anand, S. et al.: AURORA-A amplification overrides the mitotic spindle assembly checkpoint, inducing resistance to Taxol. *Cancer Cell*, **3** (1), 2003, p. 51–62.

[266] Meraldi, P. et al.: Aurora-A overexpression reveals tetraploidization as a major route to centrosome amplification in p53-/- cells. *EMBO Journal*, **21** (4), 2002, p. 483–492.

[267] Shah, K.N. et al.: Aurora kinase A drives the evolution of resistance to third-generation EGFR inhibitors in lung cancer. *Nature Medicine 2018 25:1*, **25** (1), 2018, p. 111–118.

[268] Imai, Y. et al.: The CED-4-homologous protein FLASH is involved in Fas-mediated activation of caspase-8 during apoptosis. *Nature*, **398** (6730), 1999, p. 777–785.

[269] Chen, S. et al.: FLASH Knockdown Sensitizes Cells To Fas-Mediated Apoptosis via Down-Regulation of the Anti-Apoptotic Proteins, MCL-1 and Cflip Short. *PLoS ONE*, **7** (3), 2012, p. e32971.

[270] Barcaroli, D. et al.: FLASH is required for histone transcription and S-phase progression. *Proceedings of the National Academy of Sciences of the United States of America*, **103** (40), 2006, p. 14808–14812.

[271] Hummon, A.B. et al.: Systems-wide RNAi analysis of CASP8AP2/FLASH shows transcriptional deregulation of the replication-dependent histone genes and extensive effects on the transcriptome of colorectal cancer cells. *Molecular Cancer*, **11** (1), 2012, p. 1.

[272] Kittler, R. et al.: An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature*, **432** (7020), 2004, p. 1036–1040.

[273] Sokolova, M. et al.: Genome-wide screen of cell-cycle regulators in normal and tumor cells identifies a differential response to nucleosome depletion. *Cell Cycle*, **16** (2), 2017, p. 189–199.

[274] Hirano, T. et al.: A novel interaction between FLICE-Associated Huge Protein (FLASH) and E2A regulates cell proliferation and cellular senescence via tumor necrosis factor (TNF)- Alpha-p21WAF1/CIP1 Axis. *PLoS ONE*, **10** (7), 2015, p. e0133205.

[275] Kiriyama, M. et al.: Interaction of FLASH with arsenite resistance protein 2 is involved in cell cycle progression at S phase. *Molecular and cellular biology*, **29** (17), 2009, p. 4729–4741.

[276] de Cola, A. et al.: FLASH is essential during early embryogenesis and cooperates with p73 to regulate histone gene transcription. *Oncogene*, **31** (5), 2012, p. 573–582.

[277] Choi, Y. et al.: FLASH coordinates NF-kappa B activity via TRAF2. *The Journal of Biological Chemistry*, **276** (27), 2001, p. 25073–25077.

[278] Jun, J. et al.: Role of FLASH in caspase-8-mediated activation of NF-kappaB: dominant-negative function of FLASH mutant in NF-kappaB signaling pathway. *Oncogene*, **24** (4), 2005, p. 688–696.

[279] Minamida, Y. et al.: FLASH/casp8ap2 Is Indispensable for Early Embryogenesis but Dispensable for Proliferation and Differentiation of ES Cells. *PLoS ONE*, **9** (9), 2014, p. e108032.

[280] Yang, X. cui et al.: FLASH, a Proapoptotic Protein Involved in Activation of Caspase-8, Is Essential for 3′ End Processing of Histone Pre-mRNAs. *Molecular Cell*, **36** (2), 2009, p. 267–278.

[281] Kino, T. et al.: FLASH interacts with p160 coactivator subtypes and differentially suppresses transcriptional activity of steroid hormone receptors. *Journal of Steroid Biochemistry and Molecular Biology*, **92** (5), 2004, p. 357–363.

[282] Kino, T., Chrousos, G.: Tumor necrosis factor alpha receptor- and Fas-associated FLASH inhibit transcriptional activity of the glucocorticoid receptor by binding to and

interfering with its interaction with p160 type nuclear receptor coactivators. *The Journal of biological chemistry*, **278** (5), 2003, p. 3023–3029.

[283] Alm-Kristiansen, A.H. et al.: PIAS1 interacts with FLASH and enhances its co-activation of c-Myb. *Molecular Cancer*, **10** (1), 2011, p. 21.

[284] Alm-Kristiansen, A. et al.: FLASH acts as a co-activator of the transcription factor c-Myb and localizes to active RNA polymerase II foci. *Oncogene*, **27** (34), 2008, p. 4644–4656.

[285] Parrish, A.B. et al.: Cellular Mechanisms Controlling Caspase Activation and Function. *Cold Spring Harbor Perspectives in Biology*, **5** (6), 2013, p. a008672.

[286] McComb, S. et al.: Efficient apoptosis requires feedback amplification of upstream apoptotic signals by effector caspase-3 or -7. *Science Advances*, **5** (7), 2019, p. eaau9433.

[287] Tait, S.W.G., Green, D.R.: Caspase-independent cell death: leaving the set without the final cut. *Oncogene*, **27** (50), 2008, p. 6452–6461.

[288] Tang, D. et al.: The molecular machinery of regulated cell death. *Cell Research*, **29** (5), 2019, p. 347–364.

[289] Kroemer, G. et al.: Classification of cell death: recommendations of the Nomenclature Committee on Cell Death 2009. *Cell Death & Differentiation*, **16** (1), 2009, p. 3–11.

[290] Janson, V. et al.: Resistance to caspase-8 and -9 fragments in a malignant pleural mesothelioma cell line with acquired cisplatin-resistance. *Cell Death and Disease*, **1** (9), 2010, p. 78.

[291] Li, J. et al.: Mutants TP53 p.R273H and p.R273C but not p.R273G Enhance Cancer Cell Malignancy. *Human Mutation*, **35** (5), 2014, p. 575–584.

[292] Sun, S. et al.: Hotspot mutant p53-R273H inhibits KLF6 expression to promote cell migration and tumor metastasis. *Cell Death and Disease*, **11** (7), 2020, p. 1–9.

[293] Shaulian, E., Karin, M.: AP-1 in cell proliferation and survival. *Oncogene*, **20** (19), 2001, p. 2390–2400.

[294] Bakiri, L. et al.: Promoter Specificity and Biological Activity of Tethered AP-1 Dimers. *Molecular and Cellular Biology*, **22** (13), 2002, p. 4952–4964.

[295] van Dam, H., Castellazzi, M.: Distinct roles of Jun : Fos and Jun : ATF dimers in oncogenesis. *Oncogene*, **20** (19), 2001, p. 2453–2464.

[296] Angel, P. et al.: Function and regulation of AP-1 subunits in skin physiology and pathology. *Oncogene*, **20** (19), 2001, p. 2413–2423.

[297] Karin, M. et al.: AP-1 function and regulation. *Current Opinion in Cell Biology*, **9** (2), 1997, p. 240–246.

[298] Bejjani, F. et al.: The AP-1 transcriptional complex: Local switch or remote command? *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, **1872** (1), 2019, p. 11–23.

[299] Shaulian, E.: AP-1 — The Jun proteins: Oncogenes or tumor suppressors in disguise? *Cellular Signalling*, **22** (6), 2010, p. 894–899.

[This page intentionally left blank]