Dissecting oncogenic signaling networks through combinatorial and single-cell CRISPR screens

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften

(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I

– Biowissenschaften –

der Martin-Luther-Universität

Halle-Wittenberg

vorgelegt

von Herr M.Sc. Ghanem El Kassem

Halle (Saale), 2025

Supervisor: Prof. Dr. Michael Böttcher

First reviewer: Prof. Dr. Sonja Keßler

Second reviewer: Prof. Dr. Michael Boettcher

Third reviewer: Prof. Dr. Manuel Kaulich

Date of public defense: 22.01.2026

Declaration

I hereby declare that I have completed this work independently, without any third-party assistance, except where specific references are made to the contributions of others. All sources used, whether literally or through content, are acknowledged and appropriately cited. This work has not been submitted, in whole or in part, for any academic qualification at another institution.

Halle (Saale), 09.07.2025

Ghanem El Kassem

List of publications

El Kassem, G., Hillmer, J. & Boettcher, M. Evaluation of Cas13d as a tool for genetic interaction mapping. Nat Commun 16, 1631 (2025). https://doi.org/10.1038/s41467-025-56747-4

El Kassem, G., Sieber, A., Klinger, B., Uhlitz, F., Steinbrecht, D., van Bentum, M., et al. Targeted Perturb-seq Reveals EGR1 and FOS as Key Regulators of the Transcriptional RAF-MAPK Response. BioRxiv. (2024) Jan 15. https://doi.org/10.1101/2024.01.13.575500

Acknowledgment

This thesis marks the end of a long and challenging journey that I could not have completed without the support and encouragement of many people. Reflecting on this path, I am deeply grateful for the mentors, colleagues, friends, and family who stood by me, offering their guidance, expertise, and encouragement through the obstacles and the growth.

First and foremost, I am grateful to Prof. Michael Böttcher for trusting me with his first PhD position. His persistence, goal-oriented guidance, and leadership have laid the groundwork for this work. His mentorship, both in the lab and beyond, will always inspire me in the future. I also want to thank Constanze Westphal, Jasmin Hillmer, Erik Haußner, and other current and former members of the Böttcher group for their invaluable support. Our engaging scientific discussions and their continuous feedback have shaped this thesis, and I am equally thankful for the conversations and the many moments of laughter that made the long hours in the lab more enjoyable. I am also profoundly grateful to Prof. Sonja Keßler for supervising and reviewing this PhD thesis.

I extend my gratitude to Prof. Nils Blüthgen and his research group for their valuable computational insights during this work and for collaborating on numerous projects. I would also like to sincerely thank Dr. Alex Navarrete Santos, Annika Weißenborn, Dr. Nadine Bley, and Alexander Rausch for their diligent efforts and assistance with the flow cytometry experiments. Special thanks go to Dr. Stefen Panzner, Dr. Christian Reinsch, Dr. Sandra Lagauzère, and Dr. Johannes Stuttmann, who initially provided me with the research experience that led me to this point, along with their invaluable advice and coaching.

Words can never truly express how much I am in debt to my wife, Mayssam. I could not have completed this journey without her motivation, strength, and belief in me. To my friends, Mohammad Abu Khalaf, Mohammad Rafa'a, Ali Kanish, Ibrahim Morgan, Mohammad Saoud, and Khaled Al Sakran, I will never forget your encouragement and companionship during the best and worst times. Finally, with deepest gratitude, I dedicate this work to my family. I owe everything to my parents, who have always been a source of unconditional love and unwavering presence. To my brother Nabil and sister Lina, I would also like to express my appreciation to you for your ongoing encouragement and faith in me.

Table of contents

Table of figures

Table of tables

Table of abbreviations

°C	Degree Celsius	M	Molar
4OHT	4-Hydroxytamoxifen	MAPK	Mitogen-Activated Protein Kinase
AAV	Adeno-Associated Virus	MAPKK	Mitogen-Activated Protein Kinase Kinase
ABL1	Abelson Murine Leukemia Viral Oncogene Homolog 1	MAPKKK	Mitogen-Activated Protein Kinase Kinase Kinase
ACTB	Actin Beta	MAX	MYC-Associated Factor X
ADP	Adenosine Diphosphate	MEK	MAPK/ERK Kinase, Mitogen-Activated Protein Kinase Kinase
AKT	Protein Kinase B	mg	milligram
AP-1	Activator Protein 1	min	Minutes
APC	Allophycocyanin	miRNA	microRNA
ARID3B	AT-Rich Interaction Domain 3B	mL	milliliter
ATCC	American Type Culture Collection	MLE	Maximum Likelihood Estimation
ATP	Adenosine Triphosphate	mM	millimolar
BCR	B Cell Receptor	MMEJ	Microhomology-Mediated End Joining
BHLHE40	Basic Helix-Loop-Helix Family Member E40	MOI	Multiplicity of Infection
BICDL	BICD Family-Like Cargo Adaptor 1	mRNA	messenger Ribonucleic Acid
BRCA	Breast Cancer Gene	mTOR	Mechanistic Target of Rapamycin Kinase
BSA	Bovine Serum Albumin	mU6	mouse U6 promoter
CALCB	Calcitonin-Related Polypeptide Beta	MXD1	MAX Dimerization Protein 1
CAR	Chimeric Antigen Receptor	MYC	Myelocytomatosis Oncogene
Cas	CRISPR-associated Protein	ncRNA	non-coding RNA
CASCADE	CRISPR-associated Complex for Antiviral Defense	NF1	Neurofibromin 1
CCLE	Cancer Cell Line Encyclopedia	NGS	Next-Generation Sequencing
cDNA	complementary Desoxyribonucleic Acid	NHEJ	Non-Homologous End Joining
CML	Chronic Myeloid Leukemia	NMD	Nonsense-Mediated Decay
CR	Conserved Regions	NPAS2	Neuronal PAS Domain Protein 2

CREB	cAMP Response Element-Binding Protein	NR4A1	Nuclear Receptor Subfamily 4 Group A Member 1
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats	NTC	Non-Targeting Control
CRISPRa	CRISPR activation	NUC	Nuclease Lobe
CRISPRi	CRISPR interference	PAM	Protospacer Adjacent Motif
crRNA	CRISPR RNA	PARP	Poly (ADP-ribose) Polymerase
CSRNP1	Cysteine And Serine Rich Nuclear Protein 1	PBS	Phosphate-Buffered Saline
DAPI	4',6-Diamidino-2-Phenylindole	PCR	Polymerase Chain Reaction
dCas9	Nuclease-dead Cas9	PEG	Polyethylene Glycol
DEG	Delayed-Early Gene	PFS	Protospacer Flanking Site
DepMAP	Cancer Dependency Map	PI	PAM-Interacting Domain
DKK1	Dickkopf-related Protein 1	PI3K	Phosphoinositide 3-kinase
DMEM	Dulbecco's Modified Eagle Medium	Pol II	RNA Polymerase II Promoter
DNA	Deoxyribonucleic Acid	Pol III	RNA Polymerase III Promoter
DR	Direct Repeat	pre-crRNA	precursor crRNA
DSB	Double-Strand Break	PTPN1	Protein Tyrosine Phosphatase, Non-Receptor Type 1
dsDNA	double-stranded Deoxyribonucleic Acid	PTPN11	Protein Tyrosine Phosphatase Non-Receptor Type 11
E-MAP	Epistatic Mini-array Profiling	QC	Quality Control
EB	Elution Buffer	qPCR	quantitative Polymerase Chain Reaction
EGFP	Enhanced Green Fluorescent Protein	RAF	Rapidly Accelerated Fibrosarcoma
EGFR	Epidermal Growth Factor Receptor	Ras	Rat Sarcoma Virus Protein
EGR	Early Growth Response	Ras-GAP	Ras-GTPase Activating Protein
EGR1	Early Growth Response 1	Ras-GEF	Ras Guanine Nucleotide Exchange Factor
EGR2	Early Growth Response 2	REC	Recognition Lobe
EGR3	Early Growth Response 3	RNA	Ribonucleic Acid
EGR4	Early Growth Response 4	RNA-Seq	RNA Sequencing
ELF4	E74 Like ETS Transcription Factor 4	RNAi	RNA interference

ELK1	ETS Like-1 Transcription Factor	RNAse	Ribonuclease
EN2	Engrailed Homeobox 2	RNP	Ribonucleoprotein
ER	Estrogen Receptor	rpm	Revolutions Per Minute
ERK	Extracellular Signal-Regulated Kinase	RPMI	Roswell Park Memorial Institute Medium
ETS-1	E26 Transformation-Specific 1	RT	Reverse Transcription
ETV5	ETS Variant Transcription Factor 5	RTK	Receptor Tyrosine Kinase
FACS	Fluorescence Activated Cell Sorting	SAM	Synergistic Activation Mediator
FAH	Fumarylacetoacetate Hydrolase	scCRISPR-seq	single-cell CRISPR sequencing
FBS	Fetal Bovine Serum	scRNA-seq	single-cell RNA sequencing
FDR	False Discovery Rate	SD	Standard Deviation
FLASH	Fast Length Adjustment of SHort reads	SDS	Sodium Dodecyl Sulfate
FOS	Fos Proto-Oncogene, AP-1 Transcription Factor Subunit	SGA	Synthetic Genetic Array
FOSB	FosB Proto-Oncogene	sgRNA	single-guide RNA
FOSL1	FOS Like 1	SHC	Src Homology 2 Domain Containing
FOSL2	FOS Like 2	SHERLOCK	Specific High-sensitivity Enzymatic Reporter unLOCKing
g	Relative Centrifugal Force	shRNA	short hairpin RNA
GAB2	GRB2-Associated Binding Protein 2	siRNA	small interfering RNAs
GAPDH	Glyceraldehyde 3-Phosphate Hydrogenase	SOS1	Son of Sevenless Homolog 1
gDNA	genomic DNA	SPRED2	Sprouty-Related EVH1 Domain-Containing Protein 2
GDP	Guanosine Diphosphate	SRA	Sequence Read Archive
GEF	Guanine nucleotide Exchange Factor	SRF	Serum Response Factor
GEM	Gel Bead-in-Emulsion	SRG	Secondary Response Gene
GI	Genetic Interaction	SSA	Single-Strand Annealing
GPRIN1	G Protein-Regulated Inducer of Neurite Outgrowth 1	ssRNA	single-stranded RNA
GRB2	Growth factor receptor bound protein 2	STAR	Spliced Transcripts Alignment to a Reference

gRNA	guide RNA	STAT	Signal Transducer and Activator of Transcription
GSK3β	Glycogen Synthase Kinase 3 Beta	TALEN	Transcription Activator-like Effector Nuclease
GTP	Guanosine Triphosphate	TAPseq	Targeted Perturb-seq
h	Hours	TBE	Tris-Borate-EDTA
HDR	Homology-Directed Repair	TCF	Ternary Complex Factor
HEPN	Higher Eukaryotes and Prokaryotes Nucleotide-Binding Domain	TCF7	Transcription Factor 7
HES6	Hairy and Enhancer of Split 6	TF	Transcription Factor
HNH	His-Me Finger Endonucleases Domain	TKI	Tyrosine Kinase Inhibitor
HSF1	Heat Shock Factor 1	TME	Tumor Microenvironment
hU6	human U6 Promoter	TPM	Transcripts Per Million
ID4	Inhibitor of DNA Binding 4	tracrRNA	trans-activating CRISPR RNA
IEG	Immediate Early Gene	TSC	Tuberous Sclerosis
ILG	Immediate Late Gene	TSS	Transcription Start Site
INDEL	Insertion and Deletion Mutation	U	Units
JAK	Janus kinase	UMAP	Uniform Manifold Approximation and Projection
JNK	c-Jun N-terminal Kinase	UMI	Unique Molecular Identifier
JUN	Proto-oncogene c-JUN	UTR	Untranslated Region
JUNB	JunB Proto-Oncogene	VPR	VP64-p65-Rta Synthetic Transcriptional Activator
KLF10	KLF Transcription Factor 10	WED	Wedge Domain
KRAB	Krüppel-Associated Box	WPRE	Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog	ZFN	Zinc-finger Nucleases
L	Liter	ZNF26	Zinc Finger Protein 26
LB	Luria-Bertani Medium	µg	Microgram
lncRNAs	long non-coding RNA	µL	Microliter
LTR	Long Terminal Repeat	µM	Micromolar

Abstract

Understanding how genes function together to guide cellular behavior is central to studying disease mechanisms and identifying therapeutic targets. The development of CRISPR screening methods enabled the systematic, high-throughput perturbation of genes, and facilitated the mapping of functional relationships and uncovering of key regulators of cellular signaling pathways. This thesis presents the characterization and application of two CRISPR-based functional genomics platforms aimed at studying genetic interactions (GIs) and transcriptional regulation in human cells.

In the first part of the thesis, we evaluate the RNA-targeting CRISPR effector Cas13d as a tool for quantitative genetic interaction mapping. Unlike DNA-targeting systems, Cas13d enables reversible gene silencing at the transcript level without introducing double-stranded breaks. We demonstrate that Cas13d can achieve efficient and uniform dual-gene perturbations by comparing it to the DNA-targeting nucleases Cas9 and Cas12a. We address key challenges in designing dual-gene perturbation libraries, such as gRNA-gRNA interference, through a dual-promoter gRNA expression strategy. After enhancing the phenotypic effects of gene perturbations using same-gene targeting arrays, we apply this system in the chronic myeloid leukemia cell line K562 and perform a dual-gene targeting combinatorial CRISPR screen to study the GIs between six genes that modulate the cell's response to the tyrosine kinase inhibitor imatinib. Finally, we demonstrate the potential of Cas13d for reproducible, quantitative GI mapping in oncogenic signaling pathways.

In the second part, we utilize single-cell CRISPR sequencing (scCRISPR-seq) screens to reconstruct the transcriptional network downstream of the RAF-MAPK signaling pathway. Although the upstream regulators of this pathway are well-studied, a knowledge gap persists in its downstream transcriptional regulation. By perturbing 22 transcription factors (TFs) in an inducible RAF1 HEK293 model cell line, we capture high-dimensional transcriptomic responses at single-cell resolution. We leverage the transcriptomic data to reconstruct a *de novo* TF interaction network model, revealing a positive feedback loop between EGR1 and TCF7. This feedback loop allows Wnt signaling gene expression to be enhanced when both pathways are activated, thus establishing a nexus that connects the MAPK and Wnt signaling pathways.

Together, these approaches offer scalable and high-resolution platforms for mapping genetic dependencies and unraveling the regulatory landscape in complex cell systems. The results and methods presented here have broad implications for the identification of new approaches for cancer therapy.

Keywords: CRISPR, Cas13d, Genetic interaction mapping, Single-cell CRISPR screening, Perturb-seq, Transcriptional networks, Transcription factors, RAF-MAPK signaling, Combinatorial CRISPR screening, RNA targeting, Cas9, Cas12a.

Zusammenfassung

Das Verständnis darüber, wie Gene zusammenwirken, um das Verhalten von Zellen zu steuern, ist zentral für das Entschlüsseln von Krankheitsmechanismen und das Identifizieren therapeutischer Zielstrukturen. Die Entwicklung von CRISPR-Screening-Methoden ermöglichte die systematische Störung von Genen im Hochdurchsatz und erleichterte die Kartierung funktioneller Zusammenhänge und die Aufdeckung von Schlüsselregulatoren zellulärer Signalwege. Diese Arbeit stellt die Charakterisierung und Anwendung zweier funktioneller Genomik-Plattformen auf Basis von CRISPR vor, die darauf abzielen, genetische Interaktionen (GI) und transkriptionelle Regulation in menschlichen Zellen zu untersuchen.

Im ersten Teil der Arbeit evaluieren wir den RNA-targeting CRISPR-Effektor Cas13d als Werkzeug für die quantitative Kartierung genetischer Interaktionen. Im Gegensatz zu DNA-targeting-Systemen ermöglicht Cas13d eine reversible Inhibition auf Transkriptebene, ohne DNA Doppelstrangbrüche zu verursachen. Wir zeigen, dass Cas13d eine effizientere und gleichmäßigere Inhibition mehrerer Gene ermöglicht, als die DNA-schneidenden Nukleasen Cas9 und Cas12a. Zudem gehen wir zentrale Herausforderungen beim Design von kombinatorischen gRNA Bibliotheken an, wie etwa gRNA-gRNA-Interferenzen, die wir durch eine duale Promotorstrategie lösen. In der chronisch myeloischen Leukämie-Zelllinie K562 zeigen wir, dass Phänotypen durch Konkatenierung von gRNAs zu sogenannten „arrays“ verstärkt werden. Mittels eines kombinatorischen GI-Screens untersuchten wir genetische Interaktionen zwischen sechs Genen, die die zelluläre Antwort auf den Tyrosinkinase-Inhibitor Imatinib modulieren. Abschließend demonstrieren wir das Potenzial von Cas13d für die reproduzierbare, quantitative Kartierung genetischer Interaktionen in onkogenen Signalwegen.

Im zweiten Teil nutzen wir Einzelzell-CRISPR-Sequenzierung (scCRISPR-seq) mittels Perturb-seq, um das transkriptionelle Netzwerk unterhalb des RAF-MAPK-Signalwegs zu rekonstruieren. Obwohl die vorgelagerten Regulatoren dieses Signalwegs gut erforscht sind, bleibt eine Wissenslücke bei der nachgelagerten Transkriptionsregulation. Durch die gezielte Inhibition von 22 Transkriptionsfaktoren in einer RAF1-induzierbaren HEK293-Modellzelllinie erfassen wir hochdimensionale transkriptomische Reaktionen auf Einzelzellebene. Die so gewonnenen Daten verwenden wir zur Konstruktion eines *de novo* TF-Interaktionsnetzwerkmodells, in dem wir eine positive Rückkopplungsschleife zwischen EGR1 und TCF7 identifizieren. Diese Rückkopplungsschleife ermöglicht es, die Genexpression des Wnt-Signalwegs zu verstärken, wenn beide Wege aktiviert sind, und so einen Nexus zu schaffen, der die MAPK- und Wnt-Signalwege integriert.

Insgesamt stellen beide Ansätze skalierbare und hochauflösende Plattformen zur Kartierung genetischer Abhängigkeiten und zur Entschlüsselung regulatorischer Netzwerke in komplexen Zellsytemen dar. Die hier vorgestellten Ergebnisse und Methoden haben weitreichende Implikationen für die Identifikation neuer Ansätze zur Krebstherapie.

Stichwörter: CRISPR, Cas13d, Kartierung genetischer Interaktionen, Einzelzell-CRISPR-Screening Perturb-seq, Transkriptionelle Netzwerke, Transkriptionsfaktoren, RAF-MAPK-Signalweg, Kombinatorisches CRISPR-Screening, RNA Targeting, Cas9, Cas12a.

Introduction

Mapping genetic interactions and transcriptional networks in cellular systems

1.1 Genetic networks and cancer

Understanding the complex architecture of biological systems is important for uncovering the mechanisms of diseases, such as cancer, and finding new targets for therapy. Genetic networks are regulatory systems that govern the interactions between genes and their regulatory elements. They are responsible for the function of critical cellular processes such as metabolism, growth, differentiation, and response to stimuli (1). They are usually characterized by their redundancy, interdependence, complexity, and robustness which enables the cells to respond to a variety of environmental and genetic changes. However, studying these networks is complicated by several factors such as cell-type-specific protein functions, interactions among multiple protein complexes, compensatory mechanisms, and diverse gene expression patterns (2).

Cancer is one of the leading causes of death worldwide. More than 20 million cases of cancer are diagnosed annually, resulting in over 10 million deaths from the disease, which accounts for about one in six deaths globally. The global burden of cancer is expected to exceed 35 million cases annually by 2050. Therefore, there is an urgent need for novel strategies to better understand and treat this disease (3). The transition from a normal cell to a tumor cell involves major changes in the cellular regulatory networks. These changes are referred to as “the Hallmarks of Cancer" and are characterized by cellular resistance to cell death, genomic instability, and deregulated transcription (4–6). Mutations in two classes of genes are often characteristic of tumor cells and are responsible for the changes that cause the transition to malignancy. Activating gain-of-function mutations in oncogenes and inactivating loss-of-function mutations in tumor suppressor genes promote cell growth, proliferation, and survival. Tumor suppressor genes are further classified into three categories: caretaker genes that maintain genomic stability, gatekeeper genes that regulate growth and apoptosis, and finally, landscaper genes that modify the tumor microenvironment (Fig. 1) (7–9). Cell signaling pathways are among the key processes driving these genetic and subsequent functional cellular changes.

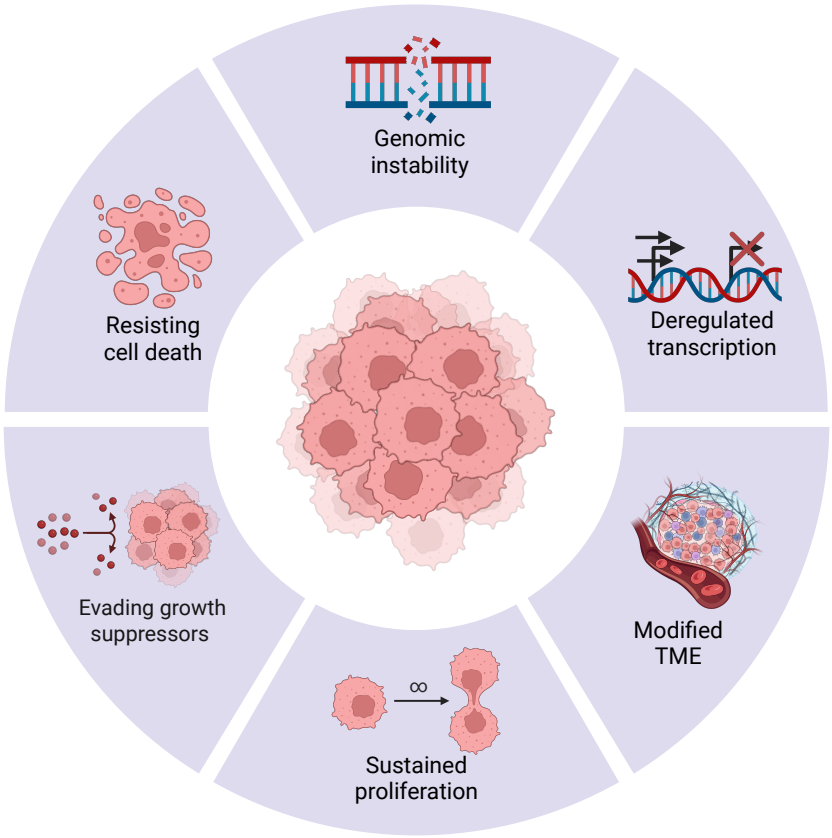


Figure 1 - The Hallmarks of Cancer.

This diagram illustrates six essential changes in cell physiology that lead to cellular malignant transformation. Together, these features enable cancer initiation, progression, and persistence. TME = Tumor microenvironment. The figure was created using Biorender.com.

1.2 Intracellular signaling pathways regulate cell function

Cell signaling pathways are intricate networks comprised of multiple players that enable cells to communicate with their environment and coordinate essential biological functions. They begin at the cell surface with extracellular receptors that bind to a variety of ligands, such as growth factors, cytokines, and hormones, initiating an intracellular signaling cascade. This cascade is conveyed through multilevel phosphorylation of proteins and secondary messengers, activating distinct transcriptional programs that regulate cell proliferation, survival, apoptosis, and differentiation (10).

Several key signaling pathways have been extensively studied over the last decade for their roles in regulating vital cellular processes. These include the mitogen-activated protein kinase (MAPK), phosphoinositide 3-kinase (PI3K)/AKT, Janus kinase/signal transducer and activator of transcription (JAK/STAT), Notch, and Wnt/β-catenin signaling pathways. While these pathways diverge and interconnect with one another, they each control fundamental cellular functions. For example, the MAPK pathway and the PI3K/Akt

pathway are responsible for cell growth and survival, respectively, while JAK/STAT signaling is involved in immune regulation and hematopoiesis (11). The Notch pathway controls cell fate and the maintenance of stem cells (12), whereas Wnt/β-catenin signaling is critical in embryogenesis and tissue homeostasis (13). Accumulating mutations in key components of these pathways are associated with a wide range of diseases, such as cancer, neurodegeneration, and immune disorders (5,14). Of these pathways, the MAPK pathway is particularly significant due to its central role in controlling diverse cell fate decisions, especially in the context of cancer, and therefore, it is the primary focus of this thesis.

1.2.1 The MAPK signaling pathway and its role in cellular regulation

The MAPK pathway is a conserved cell signaling pathway that controls a wide range of cellular processes such as proliferation, apoptosis, differentiation, migration, and metabolism. Since dysregulation in these processes is central to cancer formation, it is considered a model for studying signal-induced transcriptional programs in both normal and diseased conditions (15,16). It is composed of four major signaling modules in eukaryotic cells with the ERK (extracellular signal-regulated kinase) signaling cascade being the most studied because of its involvement in cell proliferation, growth, survival, and differentiation (17). The remaining three non-canonical MAPK signaling modules include the JNK (c-Jun N-terminal kinase)/stress-activated protein kinase, the P38 MAPK, and finally the ERK5 cascade, which are involved in stress responses and apoptosis (18–20).

The ERK-MAPK pathway starts with cell surface receptors responding to various extracellular stimuli and converging the signal through the intracellular space by a three-tiered kinase cascade consisting of a MAPK kinase kinase (MAP3K), a MAPK kinase (MAP2K), and the terminal effector MAPK (21,22). In the case of the ERK/MAPK pathway, receptor tyrosine kinases (RTKs), such as the epidermal growth factor receptor (EGFR), are activated in response to extracellular signals. Activated receptors recruit guanine nucleotide exchange factor (GEF) adaptor proteins such as SHC, GRB2, and SOS to form a receptor complex that converts the rat sarcoma virus protein (Ras), a small GTPase and a product of the Ras oncogene, from its inactive GDP-bound form to its active GTP-bound form (23).

Acting as the upstream activator of the pathway, Ras activates the MAP3K rapidly accelerated fibrosarcoma protein (RAF), initiating the cascade (24). Three subtypes of RAF proteins are expressed in human cells: A-RAF, B-RAF, and RAF1. These proteins are serine/threonine kinases composed of three conserved regions (CR) named CR1,

CR2, and CR3 (25). CR1 spans amino acids 61 to 194 and serves as the primary site for activated Ras binding to the RAF1 protein kinase. CR2 is located near the amino terminus, covering amino acids 254 to 269. Finally, CR3 is located at the carboxyl terminus from amino acids 335 to 627, and constitutes the catalytic kinase domain of the RAF1 protein (26). Activated RAF phosphorylates and activates the two isoforms of MEK (MAP2K), MEK1 and MEK2. MEK is a dual-specificity kinase that subsequently phosphorylates and activates ERK (MAPK) via its N-terminal region, completing the cascade (27). Activated ERK dimerizes and translocates from the cytoplasm to the nucleus, where it phosphorylates and activates key transcription factors (TFs), such as ELK1, CREB, AP-1, ETS-1, and MYC, which regulate gene expression programs controlling cell metabolism, proliferation, and differentiation (28) (Fig. 2).

1.2.2 MAPK pathway dysregulation in cancer

Pathological dysregulation of the tightly controlled MAPK pathway has been involved in a broad range of diseases such as cancer, autoimmune diseases, and neurological disorders (29,30). The MAPK pathway has the highest median frequency of mutations across all cancer types, with these mutations promoting irregular proliferation and survival signaling leading to tumorigenesis (31). KRAS mutations are the most common with approximately 30% of all cancers carrying driver Ras mutations (32). These are followed by B-RAF mutations which account for around 8% of cancer cases, especially in melanoma and thyroid cancers (31). Finally, dysregulations in downstream MAPK effectors such as MEK and ERK are less frequent and are mainly present in specific cancer contexts (33). As a result, therapeutic strategies have increasingly targeted MAPK components to correct dysfunctional signaling and control disease progression (34,35).

Mutations in MAPK signaling components often lead to the activation of other signaling pathways through processes such as pathway cross-talk, redundancy, and multiple feedback loops which enable tumor cells to survive targeted therapy by activating compensatory parallel pathways (36). For example, Ras activates PI3K which in turn activates AKT initiating a parallel survival signaling axis. AKT in turn inhibits glycogen synthase kinase 3 beta (GSK3β) thereby promoting Wnt-driven transcription. On the other hand, activated ERK1/2 inhibits the TSC1/2 complex and thus promotes mTOR signaling downstream of PI3K (17,37) (Fig. 2). Dissecting the complexity of the MAPK signaling pathway and its role in gene expression regulation requires experimental methods that can perturb individual components in a precise manner and map the downstream transcriptional effects. In this thesis, we apply one of these methods, single-cell CRISPR sequencing (scCRISPR-seq), to study the transcriptional network downstream of RAF kinase, offering new insights into this complex signaling landscape.

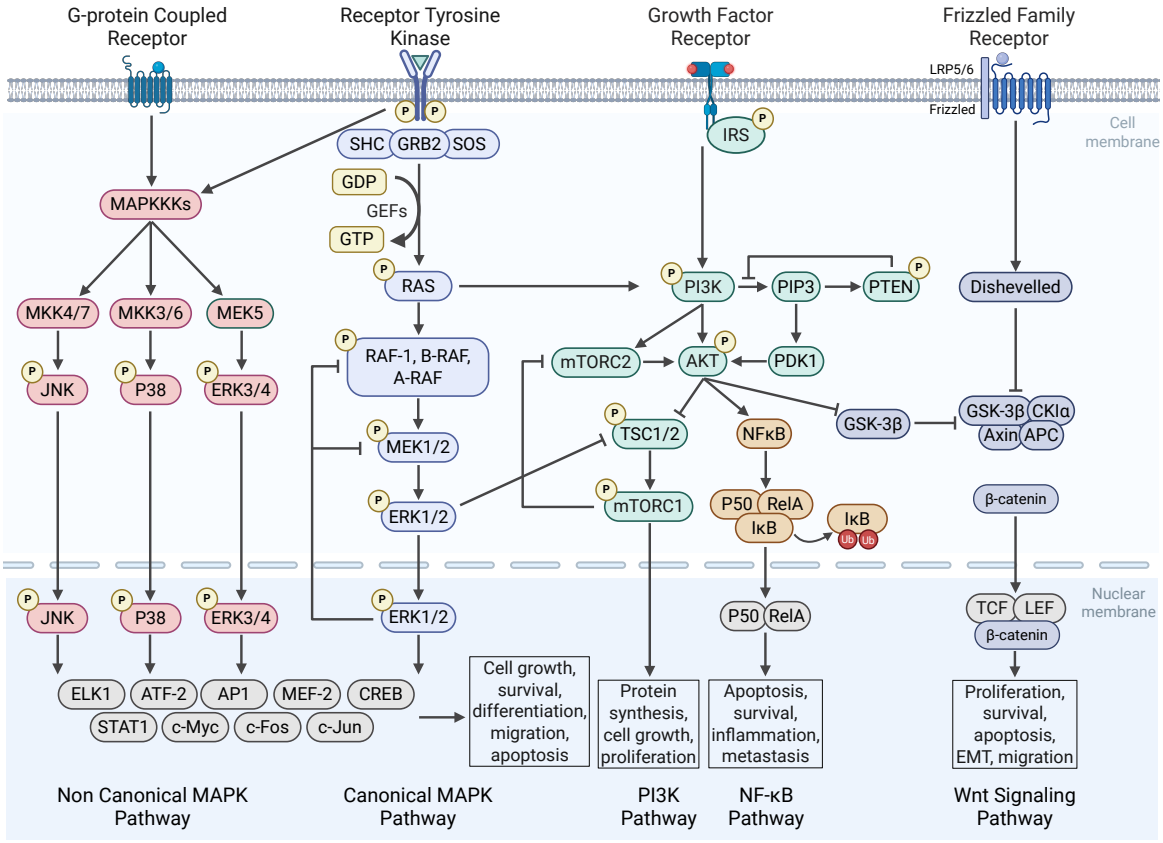


Figure 2 - Overview of the MPAK cell signaling pathway a cross-talk with other signaling pathways.

This schematic illustrates the canonical and non-canonical MAPK signaling cascade components and their interactions with other signaling pathways, such as the PI3K pathway, NF-κB pathway, and Wnt signaling pathway. This integrated network highlights the complexity and interconnectivity of cellular signaling mechanisms essential for maintaining cellular homeostasis and responding to extracellular cues. The figure was created using Biorender.com.

1.2.3 MAPK pathway dysregulation in chronic myeloid leukemia

CML (Chronic myeloid leukemia) is caused by the reciprocal translocation between chromosomes 9 and 22 leading to the creation of the Philadelphia chromosome containing the constitutively active, multidomain, chimeric tyrosine kinase BCR::ABL1 fusion protein (38,39). The main signaling pathways that are activated in CML are the Ras MAPK and the PI3K/AKT signaling pathways but the MAPK remains the core driver of CML transformation (40,41). The Ras/MAPK pathway is activated by the autophosphorylation of BCR::ABL1 which recruits adaptor proteins such as GRB2. Binding of GRB2 to the BCR region recruits GAB2 and SOS (42,43). The multi-protein complex BCR-ABL1-GRB2-SOS then activates Ras, initiating the ERK, JNK and the P38 signaling cascades (44–46). GAB2 recruits PTPN11 as well which enhances Ras-MAPK signaling by inhibiting Ras-GAP and sustaining the signaling cascade (47,48). Together with PI3K/AKT, these signaling pathways regulate cell proliferation, survival, and growth factor independence (45).

Following the mechanistic understanding of the signaling networks involved in the development of CML, tyrosine kinase inhibitors (TKIs) were developed to inhibit BCR::ABL1 activity. By targeting the ATP-binding site of the BCR::ABL1 fusion gene kinase domain, they block the signaling process and reduce the proliferation and survival of CML cells (49). Imatinib was the first TKI developed by Novartis Pharmaceuticals and it marked a major advancement in targeted cancer therapy. Despite the success of Imatinib in treating CML, targeting a single driver mutation is not always sufficient, particularly in cancers with redundant pathways or compensatory mechanisms that can diminish therapeutic efficacy. This highlights a broader challenge in functional genomics: the robustness of biological systems where genetic redundancy and functional buffering can mask the effects of individual gene disruptions.

1.3 Genetic redundancy and synthetic lethality

1.3.1 Genetic redundancy and functional buffering

Through evolution, the human genome acquired a high degree of genetic redundancy caused by processes such as diploidy, gene duplication, and functional redundancy of signaling pathways (50,51). This provides a buffer against genetic mutations that cause the loss or disruption of particular genes, enabling cells to maintain critical biological functions that are important for survival and confer cellular robustness and adaptability under normal and stressful conditions. This is more relevant in tumor cells, where genetic redundancy enables their survival despite having highly rearranged genomes.

Gene paralogues, which constitute about two thirds of the human genome, are an example of genetic redundancy. They are usually derived from a common ancestral gene and are located in different genomic regions (52,53). Paralogs compensate for one another to provide functional buffering and compensatory mechanisms, decreasing the chances of single-gene essentiality. However, they also provide opportunities for synthetic lethality, a phenomenon where the simultaneous loss of two genes leads to cell death, whereas the loss of a single gene is tolerated (54,55).

1.3.2 Synthetic lethality and its applications in cancer therapy

Synthetic lethality was initially established in model organisms such as yeast and *Drosophila* by observing that mutations in pairs of genes specifically killed cells, while corresponding single-gene mutations had no effect (56–58). This concept has now been adapted for cancer therapy by utilizing tumor-specific genetic alterations to exploit vulnerabilities that can be targeted therapeutically (59). Vulnerabilities related to malignant transformation, such as genomic instability, defective DNA repair, and disordered

transcription, which weaken the tumor cell's buffering capacity, can be specifically targeted by selectively exploiting the interrelationships between genetic mutations (9) (Fig. 3). One successful clinical example of exploiting synthetic lethal interactions involves the use of poly (ADP-ribose) polymerase (PARP) inhibitors in tumors that carry mutations in either the *BRCA1* or *BRCA2* genes. With impaired homologous recombination and compromised DNA repair, PARP inhibitors selectively target these tumor cells while sparing normal cells (60–62).

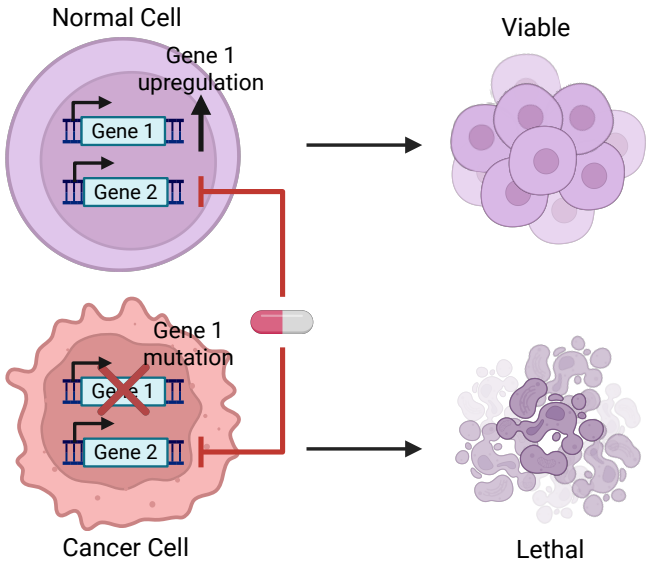


Figure 3 - Schematic representation of synthetic lethality in cancer therapy.

In normal cells, two genes perform overlapping functions through genetic redundancy, enabling one gene to compensate if the other is inhibited. However, in cancer cells, a loss-of-function mutation in one of these genes makes them vulnerable to therapies targeting the second gene, resulting in selective cancer cell death. The figure was created using Biorender.com.

While synthetic lethality has proven to be a strong therapeutic concept, the context-dependent nature of these interactions in human cells remains a significant burden. More broadly, synthetic lethality is one type of genetic interaction (GI) that can be systematically explored through GI mapping to create genome-wide maps of gene-gene dependencies (63). Systematic and scalable high-throughput combinatorial screening strategies that can capture the complexity of cellular systems are needed for conveying information about functional associations, buffering capacity, and pathway redundancy.

1.4 Genetic interactions: Mechanisms and applications

1.4.1 Defining genetic interactions

A GI between two genes occurs when the combined effect of perturbing two genes deviates from the expected phenotypes of the single-gene perturbations (64). GI assays

perturb multiple genes within the same cell to identify nonlinear combinatorial phenotypes, revealing dependencies that are not apparent from single-gene perturbations. Quantitative GIs are classified into two classes: negative (synergistic) and positive (buffering) interactions. Negative interactions, like synthetic sickness or lethality, represent those interactions where the combined disruption of two genes has a more severe phenotype than what might be expected from the sum of single gene perturbations (2,65). In turn, positive interactions may weaken one mutation against the other, as occurs in suppressive and epistatic interactions, in which the effect of one gene mutation suppresses or masks the effect of the second gene, respectively. In positive interactions, the double mutant is less harmful than the expected sum of effects of the single mutants (66) (Fig. 4).

GIs do not necessitate direct physical interaction between gene products nor their co-expression but rather they reflect the functional relationship between genes such as their involvement in the same biological processes or pathways (67). By providing insights into gene function and network organization, GI maps shed light on relevant cellular processes by systematically linking genotype to phenotype. They use the full interaction profile signature of a gene to group genes with similar functions based on their involvement in the same protein complex, signaling pathway, or biological process (68–70). Creating GI maps requires dense data to classify genes by function in an unbiased manner and has so far been instrumental in shedding light on the mechanism behind synergistic or buffering phenotypes (2).

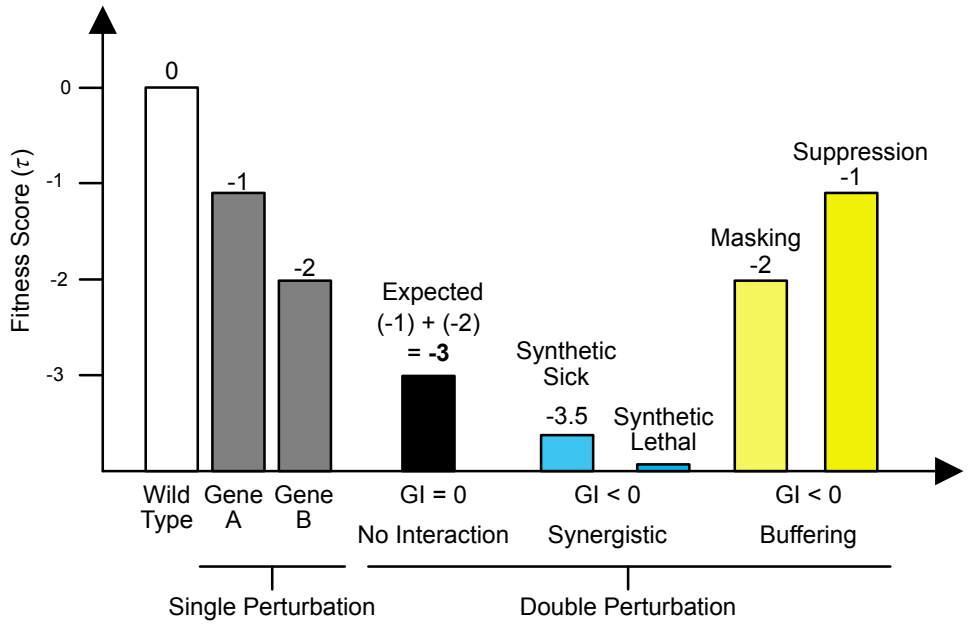


Figure 4 - A graphical representation of a genetic interaction (GI) screen showing cell fitness phenotypes resulting from single and double gene perturbations.

The fitness effects of single-gene perturbations of Gene A and Gene B (grey bars) are compared to the expected additive effect of their combined perturbation (black bar) to determine GIs. A measured fitness score equal to the expected value indicates no interaction (GI = 0). A greater-than-expected fitness defect from the dual perturbation of both genes reveals synergistic interactions (GI < 0), such as synthetic sickness and synthetic lethality (blue bars). Conversely, buffering interactions (GI > 0), such as masking and suppression, occur when the double perturbation results in higher fitness than expected (yellow bars).

1.4.2 Experimental mapping of genetic interactions

GI mapping has been extensively studied in model organisms where foundational work has elucidated gene function, protein complexes, signaling pathways, and evolutionary principles (64,71). The availability of comprehensive mutant libraries and genetic traceability in yeast has made it the ideal system for the systematic study of GIs. High throughput methods such as synthetic genetic array (SGA) screening and epistatic mini-array profiling (E-MAP) have accelerated the large-scale mapping of GIs (72–75) and allowed the mapping of ~75% of the yeast genome in pairwise fitness screens covering more than 5.4 million gene pairs in *Saccharomyces cerevisiae* (74).

Despite these advances, the findings from model systems are hard to translate to more complex organisms such as mammalian cells. Human cells exhibit a much higher degree of complexity and context-dependent GIs. To date, the largest GI map study has pairwise interactions across 472 genes (68), representing only 0.05% of all possible interactions of human protein-coding genes. Therefore, novel methods are needed to expand this GI network. As such, high-throughput tools such as RNA interference (RNAi) and the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas offer scalable and versatile platforms for perturbing gene function at a genome-wide level. When combined with predictive models, which can reduce the number of gene combinations that need to be tested, they allow deeper and more efficient comprehensive mapping of GIs in human cells (67).

1.5 High-Throughput screens for genetic interaction mapping: From RNAi to CRISPR

1.5.1 RNAi

Initial synthetic lethality screens in mammalian cells were limited to chemical compound-based approaches which lacked in scalability and target specificity (76,77). Introduction of RNAi technology allowed the first genome-wide programmable loss-of-function screens in human cells (2,78–81). They use small interfering RNAs (siRNAs) or short hairpin RNAs (shRNAs) as a programmable effector molecule that can be designed to target specific mRNAs simply by achieving sequence complementarity and thus silencing gene

expression. These screens have been instrumental in the identification of synthetic lethal interactions and systematic probing of gene function in mammalian cells without prior assumptions about their function or interactions (82). But despite its versatility and ease of use, RNAi had remarkable limitations.

Knockdown efficiency from RNAi targeting is highly variable due to differences in RNAi constructs (83,84) resulting in incomplete or inconsistent phenotypic outcomes (85). Off-target effects add a second layer of complexity by increasing false positives and complicating data interpretation. These challenges have been partially overcome by careful RNAi construct design strategies such as using multiple constructs to target the same gene. Despite this progress, conducting effective and reliable genome-wide RNAi screens remained very hard making the results from RNAi-based applications less reliable and reproducible (86). Nevertheless, RNAi screens successfully identified multiple synthetic lethal interactions and gene functions, especially in cancer research. However, the limitations of RNAi paved the way for the adoption of more precise genome-editing technologies.

1.5.2 Evolution of genome editing technologies

Technologies targeting the genome, such as zinc-finger nucleases (ZFNs) (87,88), transcription activator-like effector nucleases (TALENs) (87,89,90), and homing meganucleases (91), represent significant advancements in targeted genome modification. However, these approaches require a costly, fully customized protein engineering for each target site. This increases their complexity and limits their scalability and adoption (2,79). The advent of the CRISPR system revolutionized the field by addressing these limitations.

1.5.3 From prokaryotic immune systems to genome editing

The CRISPR system is a multistep adaptive immunity mechanism in bacteria and archaea that is composed of three steps: recognition, incorporation, and degradation of the foreign genetic material, thus protecting them against nucleic acid-based infections such as viruses and phages (92–94). A functional CRISPR system mainly consists of two components. First, the CRISPR array, which is located in the host genome and is composed of stretches of hypervariable fragments of foreign DNA or spacers interspersed between repetitive sequences. Each spacer represents a genetic record of a previous infection. The second component is a heterogeneous combination of genes that translate CRISPR-associated proteins (Cas), which are nucleases that are responsible for processing spacer arrays and degrading the invading genetic material (95) (Fig. 5).

The prokaryotic CRISPR-Cas immune response can be separated into three different stages:

1. Adaptation: This step includes the recognition of the foreign DNA as non-self and integrating it as a new spacer, which is a short nucleotide sequence that is complementary to the invading foreign genetic material, between two adjacent repeat units within the CRISPR locus. The spacer is integrated mainly at the beginning or the leader end of the array, giving a chronological record for the acquisition events from previous infections (94,96,97). Short conserved sequences adjacent to the spacer, known as the protospacer adjacent motif (PAM), are often critical for this step (98,99). Spacer acquisition is accomplished by Cas1 and Cas2 which are universally present in CRISPR-containing genomes (100).
2. Expression: The CRISPR locus is transcribed into a long precursor CRISPR RNA (pre-crRNA) which is then processed into individual mature crRNAs by endoribonucleases. Together with trans-acting crRNAs (tracrRNAs), these crRNAs provide the sequence information to guide Cas nucleases to target foreign nucleic acids. Therefore, they are also called guide RNAs (gRNAs) (101,102).
3. Interference: gRNAs direct Cas nucleases towards complementary foreign DNA or RNA sequences (101). These proteins range from single effectors to multiprotein complexes, depending on their classification. Class 1 Cas systems which include types II, V, and VI, are composed of multiprotein complexes such as the CASCADE (CRISPR-associated complex for antiviral defense) in organisms like *E. coli* (101). Class 2 Cas systems such as Cas9, however, are typically single effector proteins and include types I, III, and IV (103). Finally, the target DNA or RNA is cleaved, thus neutralizing the invading genetic element (94,98).

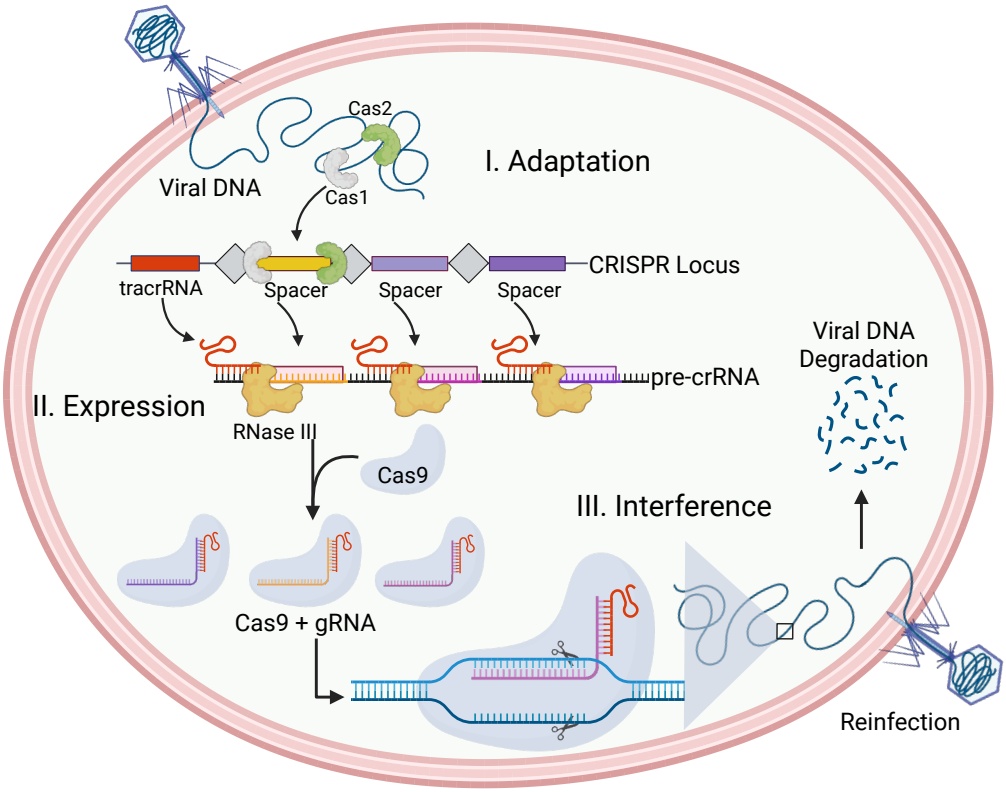


Figure 5 - Overview of the CRISPR-Cas9 mediated adaptive immune system in prokaryotes.

The defense mechanism consists of three stages. ***Adaptation****: Upon viral infection, Cas1 and Cas2 capture fragments of viral DNA and integrate them as spacers into the CRISPR locus within the host genome.* ***Expression****: The CRISPR locus is transcribed into a long pre-crRNA, which is processed by RNase III into individual crRNAs. These associate with tracrRNA and Cas9 to form the functional gRNA-Cas9 ribonucleoprotein (RNP) complex.* ***Interference****: Upon reinfection, the gRNA-Cas9 RNP complex recognizes and binds to the matching viral DNA sequence, resulting in its cleavage and degradation, thus preventing viral reinfection. The figure was created using Biorender.com.*

In recent years, a wide range of CRISPR-Cas systems have been discovered from multiple microorganisms, but the type II system encoding Cas9 from *S. pyogenes* has been the most studied and widely applied in genome editing and other applications (103–108). Cas9 was the first characterized single-effector nuclease which gave unprecedented simplicity to genome-editing applications. Understanding the mechanism of CRISPR-Cas9 for genome editing and the discovery of the tracrRNA as a core component in crRNA biogenesis marked a turning point in genetic engineering. The impact of this discovery on the scientific field was recognized with the awarding of the 2020 Nobel Prize in Chemistry to Emmanuelle Charpentier and Jennifer Doudna for their contributions to its development (109).

1.5.4 Applications of CRISPR in genetic screens

Rapidly after being characterized as an adaptive immune system in bacteria (92–94), the CRISPR-Cas9 system was adapted for precise genome editing in mammalian cells (103,106,107). This two-component system relies on a single-guide RNA (sgRNA) to direct the Cas9 nuclease to specific endogenous genomic loci, where it induces double-stranded breaks (DSBs) (93). These DSBs are then repaired by either the error-prone non-homologous end-joining (NHEJ) pathway, microhomology-mediated end joining (MMEJ), single-strand annealing (SSA), or the homology-directed repair (HDR) pathway (110–113). The error prone NHEJ is the most frequent and efficient repair pathway often resulting in insertion and deletion mutations (INDELs) leading to a frameshift in the gene coding sequence and the creation of premature stop codons triggering nonsense-mediated decay (NMD) of the subsequent mRNA transcript leading to gene knockout especially when early exons of protein coding regions are targeted (114,115).

Compared to previous gene-editing tools, CRISPR-Cas9 offers significant advantages in precision, scalability, and ease of use. Direct comparisons between CRISPR and RNAi have shown that CRISPR consistently outperformed RNAi-based methods in terms of precision and reproducibility (86,116). For example, a side-by-side comparison of next-generation shRNA and Cas9 sgRNA libraries showed that the Cas9 based library was more accurate in identifying more essential genes than its complementary shRNA counterpart was (117,118). Shortly after Cas9 was established, other systems such as Cas12a and Cas13d were developed for gene perturbation further expanding the range of CRISPR applications. Each of these systems works through different mechanisms and offers a set of advantages and limitations that make them useful tools in different research areas.

1.6 The CRISPR toolbox

1.6.1 Cas9

1.6.1.1 S. pyogenes Cas9 gene knockout

The *Streptococcus pyogenes* SF370 type II CRISPR locus encodes the Cas9 nuclease and two noncoding CRISPR RNAs: the tracrRNA and a pre-crRNA array. The pre-crRNA includes spacer sequences that guide the nuclease to its target DNA sequence and is interspersed with identical direct repeats (DRs), forming the core of the RNA-guided targeting mechanism (119).

The Cas9 protein has a bilobed structure consisting of the nuclease (NUC) lobe, which contains the HNH and RuvC-like nuclease domains, a PAM-interacting (PI) domain, a

wedge domain (WED), and an α-helical recognition (REC) lobe that facilitates the recognition of the sgRNA and target DNA (112,120,121). Upon binding to the sgRNA, Cas9 forms an active RNP complex that scans the genome for the PAM sequence (5′-NGG-3′) and binds to it, creating an active tertiary complex. This complex activates the Cas9 nuclease domains, cleaving the noncomplementary strand of the targeted DNA and generating a blunt-end DSB (112) (Fig. 6).

The *S. pyogenes* CRISPR-Cas9 system has become the most widely used tool for gene editing applications, but with some limitations. For example, Cas9 has a strict requirement for an NGG PAM, which restricts the targetable sites in the human genome to about one site per 8 base pairs. Other limitations include reduced on-target activity due to crRNA secondary structure, chromatin accessibility, and DNA methylation states (122). Addressing these limitations has involved utilizing Cas enzymes from other microorganisms and developing various engineered variants that differ in size, PAM requirements, on-target specificity, and off-target profiles, thus expanding the range of targetable genomic loci. For instance, SpCas9-HF1, eSpCas9, and HypaCas9 introduce point mutations that drastically reduce off-target effects while maintaining robust on-target activity. xCas9 and SpCas9-NG expand PAM compatibility beyond NGG, enabling access to a broader set of genomic sites (123). Additionally, smaller Cas versions, such as mini-Cas9 and split-Cas9, can fit into a single lentiviral vector together with the gRNA sequence while maintaining high transducibility, at the expense of reduced on-target efficacy (124–126). Furthermore, Cas9 can be further engineered into a nicking enzyme that induces single-strand breaks and facilitates homology-directed repair with minimal mutagenic activity (103).

1.6.1.2 CRISPR interference and CRISPR activation

Besides the standard method of CRISPR-based gene knockout, adaptations of the core CRISPR technology by combining the Cas9 effector protein with various fusion proteins has expanded the utility of CRISPR beyond the DSB mediated gene knockout making it possible to activate and silence target genes. Catalytically inactive Cas9 (dCas9), which retains the sgRNA-programmable ability to search but not cut the genome, can be engineered by the introduction of D10A and H840A mutations within the RuvC and HNH nuclease domains, respectively (127). This allows it to be used for RNA-directed recruitment of other molecular functions to specific loci in the genome (128–132).

CRISPR interference (CRISPRi) uses a fusion protein of a dCas9 and a transcriptional repression domain Krüppel-associated box (KRAB). By targeting the gene transcription start site (TSS), this system represses gene expression by inhibiting transcription enabling gene downregulation without DSB formation (133,134) (Fig. 6). CRISPRi is less toxic

compared to CRISPRko since it does not cause DSB formation. In addition, it is less prone to activate compensatory mechanisms (134,135) however, it requires continuous expression of the effector molecules to maintain gene silencing and is usually associated with increased off-target effects.

CRISPR activation (CRISPRa), on the other hand, uses a dCas9 fusion with an activation domain such as VP64 to achieve gene overexpression. Upon binding to the gene TSS, CRISPRa systems recruit TFs to increase gene expression (Fig. 6). Next-generation CRISPRa systems include the Suntag system, which consists of repeating 10 to 24 peptide arrays of a fusion between a single-chain fragment antibody and VP64 (136). The VPR system features a tri-component activation domain consisting of VP64, p65, and Rta (137). Finally, the synergistic activation mediator (SAM) system recruits the two TFs HSF1 and p65 to enhance gene activation on top of a dCas9-VP64 fusion (138).

1.6.2 Cas12a

The type V CRISPR Cas12a family, previously referred to as Cpf1, is an alternative to Cas9 for genome editing with multiple key differences and advantages. Like Cas9, the Cas12a protein features a bi-lobed protein structure composed of a REC lobe and a NUC lobe. However, unlike Cas9, which has an HNH domain, the NUC lobe of Cas12a has two RuvC nuclease domains positioned in a configuration that allows them to be superimposed (139). Second, Cas12a recognizes a 5'-TTTV-3' PAM sequence and generates staggered cuts instead of blunt cuts in the DNA which can be useful for HDR applications. Third, Cas12a has a single RNA requirement represented by a 20-nucleotide DR sequence (140). However, the main distinguishing feature is the ability to process an array of gRNAs separated by DR sequences expressed from a single transcript into single effector gRNAs (140) (Fig. 6). Because of their compact design, these arrays offer substantial advantages during DNA synthesis and sequencing, enhancing their utility in high-throughput applications such as multiplexed gene perturbation screens.

The activity of various Cas12a orthologues derived from different organisms, including *Acidaminococcus* (AsCas12a) and *Lachnospiraceae* (LbCas12a) has been demonstrated in human cells (140). The nucleotide preferences of these orthologues have been broadly characterized, enabling more refined gRNA design for specific applications (141,142). Protein engineering of the enhanced AsCas12a variant (enAsCas12a) has increased its on-target activity, broadening the spectrum of its usable applications (143). Cas12a has also been adapted to CRISPRi and CRISPRa systems similar to Cas9 (144–146). Finally, Cas12a shows collateral cleavage activity where it can nonspecifically cut single-stranded DNA after recognizing its target. This characteristic was used in diagnostics like the

SHERLOCK system to detect viral and bacterial infections, as well as genetic mutations (147,148).

1.6.3 Cas13d

Type II Cas9 and Type V Cas12a CRISPR-Cas systems induce permanent gene disruption through the introduction of DSB in the genomic DNA (gDNA) which is irreversible and associated with cellular toxicity. To address these limitations, a novel class of RNA-targeting Type VI CRISPR-Cas13 family of endonucleases was introduced in 2016 capable of programmable single-stranded RNA (ssRNA) cleavage without altering the gDNA (149). The type VI Cas13 system is subclassified into four subtypes with each subtype consisting of a distinctive Cas13 enzyme such Cas13a (C2c2) for type VI-A, Cas13b for type VI-B, Cas13c for type VI-C, and Cas13d for type VI-D. All of these effectors share two conserved Higher Eukaryotes and Prokaryotes Nucleotide-binding (HEPN) domains that harbor RNase motifs that are responsible for RNA cleavage with high efficiency (139).

The most recently discovered Cas13d effector, particularly the RfxCas13d from *Ruminococcus flavefaciens* XPD3002 (also known as CasRx), has several advantages over other Cas13 variants and has been adopted for RNA-targeting applications in mammalian systems (139,150). Compared to other Cas13 enzymes, Cas13d demonstrates minimal off-target effects and no cytotoxicity, with superior specificity and knockdown efficacy. It is the smallest of all Cas13 enzymes which allows it to be efficiently packaged into adeno-associated virus (AAV) vectors, facilitating *in vivo* delivery. In addition, Cas13d does not require a protospacer flanking sequence (PFS), allowing for more flexible and unconstrained target selection compared with other systems with specific PAM requirements (139,150,151) (Fig. 6). Similar to Cas12a, Cas13d possesses dual nuclease activity which enables both precise cleavage of target ssRNA and processing of gRNA arrays into single gRNAs capable of targeting multiple target sites enabling combinatorial gene perturbation, increased knockdown efficiency, and smaller library sizes for high throughput screens (150,152).

Cas13d is particularly well-suited for targeting non-coding RNAs (ncRNAs), such as microRNAs (miRNAs) and long non-coding RNAs (lncRNAs), which play critical roles in gene regulation and disease pathology (153–155). Unlike protein-coding genes, lncRNAs lack open reading frames and are thus they are less prone to disruption via traditional CRISPR-mediated INDEL formation (2). While CRISPRa and CRISPRi approaches can modulate lncRNA expression, off-target effects from lncRNAs overlapping with or being

adjacent to protein-coding genes complicate the interpretation of results coming from these systems (156,157).

Cas13d can induce collateral cleavage activity where it can nonspecifically cut ssRNA in close proximity after recognizing its target. Several publications have reported cytotoxicity in eukaryotic cells associated with Cas13d collateral activity depending on factors such as cell type, target expression levels, and delivery method (158–161). Although initially RfxCas13d associated collateral RNA degradation was not detected (150), recent studies have shown that collateral effects can arise when targeting highly expressed genes, particularly in transfection-based delivery systems such as piggyBac transposons or liposome-mediated transfection (159,160,162–164). In contrast, viral delivery systems, including lentiviral vectors, exhibit significantly lower collateral activity (152,165–167). In addition, newly engineered high fidelity Cas13d variant (163) as well as Cas13d orthologous such as *Dj*Cas13d (160) have shown minimal cytotoxicity even when targeting highly expressed transcripts.

RfxCas13d has been employed for several RNA-targeting studies involving gene function and developmental changes in model organisms (168,169), metabolic and cell engineering (165,170), as well as cancer research (171–173). In addition, the system has also been used for RNA engineering and mRNA alternative splicing modulation by engineering dCas13d-splice factor fusion proteins or targeting cis-elements of pre-mRNA (150,174). Despite its advantages for perturbing multiple targets per cell, no study has investigated the utility of RfxCas13d for quantitative GI mapping in human cells. In this thesis, we compare it to the DNA targeting Cas9 and Cas12a CRISPR systems and evaluate its applicability for quantitative GI mapping screening.

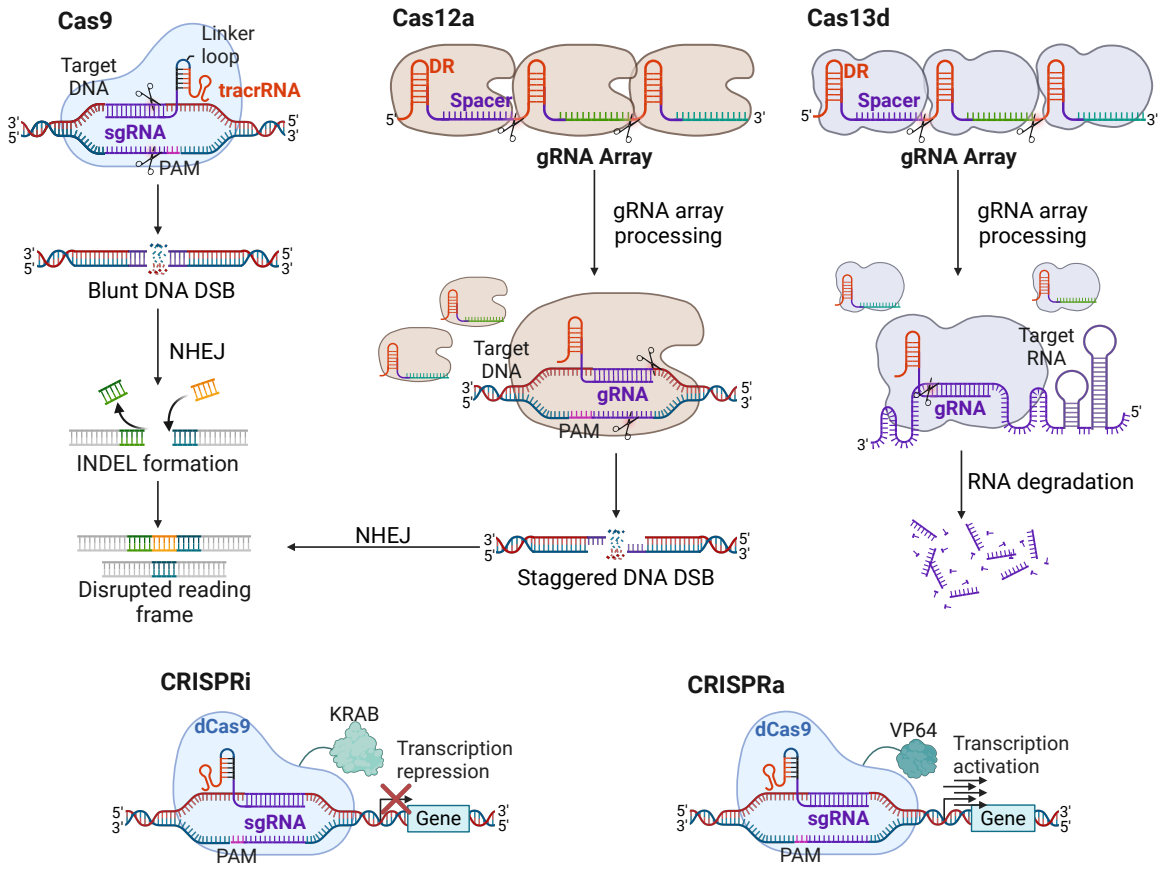


Figure 6 - Overview of various CRISPR-Cas systems and their gene perturbation mechanisms.

***Cas9**: Utilizes a sgRNA composed of crRNA and tracrRNA fused by a linker loop to induce blunt DSB at specific DNA targets. DNA repair via NHEJ can result in INDEL formation, leading to disruption of the gene reading frame and gene knockout. **Cas12a**: Processes multi-gRNA arrays releasing single gRNAs that target multiple locations. It introduces staggered DSBs at target DNA sites, resulting in gene disruption via NHEJ. **Cas13d**: Targets RNA instead of DNA. Similar to Cas12a, it processes multi-gRNA arrays against multiple RNA targets. The cellular RNA degradation machinery degrades the cleaved RNAs, leading to gene expression knockdown. **CRISPRi**: A catalytically inactive dCas9 is fused to a KRAB repressor domain, which binds to a gene promoter and blocks transcription, resulting in gene silencing. **CRISPRa**: dCas9 is fused to a transcriptional activator (e.g. VP64), promoting gene expression by recruiting transcriptional machinery to the target gene's promoter region. The figure was created using Biorender.com.*

1.7 CRISPR screens in cancer functional genomics

Forward genetics is a type of open-ended, hypothesis generating research focusing on studying genes, pathways and mechanisms involved in a given phenotype or biological process (175). Since the CRISPR specificity is dictated by a 20-base pair sequence in the sgRNA, this ease of target selection and modification allowed CRISPR to rapidly be used in genome-wide and multiplexed loss-of-function high-throughput genetic screens (114,176,177), replacing RNAi-based libraries. This made CRISPR screens a powerful source of biological discovery in PERTomics applications, enabling the unbiased

interrogation of gene function and an indispensable tool in the identification of cancer drug targets and in the investigation of mechanisms of drug resistance (54,85,115,178–180).

1.7.1 Pooled CRISPR screens for single-gene functional analysis

Pooled CRISPR screening approach combines thousands of genetic perturbations in a single experiment. It is a forward genetics approach where a gRNA library is applied to a bulk of cells to introduce various genetically encoded perturbations. They are relatively inexpensive and easily scalable, making them ideal for high-throughput phenotypic discovery (2). The underlying principle of such screens is based on the ability to introduce a single gRNA expression vector per target cell via lentiviral transduction at a low multiplicity of infection (MOI). In addition to the gRNAs, the CRISPR-Cas protein is expressed in the target cells either transiently by transfecting or electroporating the plasmid, mRNA or protein into the cells, or stably by lentiviral transduction and genomic integration. Stable expression can be achieved either together with the gRNA library on all-in-one constructs or in isolation with subsequent clonal selection for efficient knockout levels. The cells are then allowed to proliferate under a biological challenge, such as cell competition or drug treatment, where cells containing gene knockouts which sensitize them under the selective condition are depleted, while at the same time the cells that contain gene knockouts that confer resistance are enriched. Each gRNA in the library will specifically perturb one target gene, enabling genome-scale assessment of gene function, with each cell acting as an independent experiment. Since the gRNA sequences are integrated into the genome, they serve as barcodes that can be read with next-generation sequencing (NGS) enabling inference of the respective gene knockouts (114,176,177). At the end of the screen, enrichment and depletion analysis is performed by calculating the difference in gRNA distribution between different experimental conditions or different time points. The resulting data enables the identification of genes that confer sensitivity or resistance to the applied challenge (114,175,176) (Fig. 7).

Viability-based CRISPR screening methods can be divided into three categories: negative-selection, positive-selection, and marker-based screens. In negative-selection screens, perturbed genes that are essential for cell survival and growth are depleted to study context-specific dependencies and synthetic lethal interactions that can guide targeted cancer treatments. In positive-selection screens, a strong selective pressure is applied through drug treatments or environmental challenges to identify genetic perturbations that enable the cells to survive otherwise lethal conditions. They are especially useful for studying drug resistance mechanisms and survival factors. Marker-based screens on the other hand do not depend on changes in cell viability as a result of genetic perturbation but rather they track the expression of marker genes that can be detected through

fluorescent tagging or antibody labeling which allows the cells with altered marker expression to be isolated using fluorescence-activated cell sorting (FACS) to identify key regulators of specific pathways or gene expression (175,181).

Decreasing costs of NGS have made pooled CRISPR screens highly accessible for most research labs. Initial efforts focused on optimizing gRNA design (2), and foundational genome scale screens demonstrated that it is feasible to perturb most genes in the human genome in a single experiment (114,115,176). These studies identified essential genes for proliferation and drug response and showcased CRISPR as a functional genomics platform (176,179). Currently, the Cancer Dependency Map (DepMap) which is an open-access initiative led by the Broad Institute, contains over 1,000 human cancer cell lines with systematic maps of background specific genetic vulnerabilities using genome-scale loss-of-function screens, primarily CRISPR-Cas9 and RNAi (85). Later screens have identified genes involved in transcription regulation (182), epigenetics (183), cell signaling (184,185), proliferation (186), and differentiation (187,188). In addition, researchers have been using pooled CRISPR screening in more sophisticated models to investigate more context-dependent biological phenomena such as in primary cells, in tissue explants or in stem-cell-derived cultures including organoids as well as *in vivo* at a smaller scale (189).

Subsequent screens have employed both genome-wide and targeted libraries in defined genetic contexts, including isogenic cell lines or panels with specific mutations, to uncover GIs (77,190). However, one major limitation of single-gene knockout screens is that they cannot evaluate functional redundancy among paralogous genes, which frequently retain partial overlapping functions and their inactivation compensates for each other. A systematic analysis of screen data in cancer cell lines generated by the Cancer Dependency Map observed that half of all constitutively expressed genes are never detected in any single gene perturbation CRISPR screen (191). To address these limitations, higher order combinatorial screening methods have been developed to reveal interactions otherwise obscured by traditional single-gene screening approaches.

1.7.2 Pooled combinatorial CRISPR screens

Combinatorial CRISPR screens are a type of pooled CRISPR screens in which, instead of a single perturbation per cell, multiple perturbations are introduced to study GIs between genes. These screens typically focus on gene pairs but can also be applied to study higher-order combinations where multiple genes are targeted in each cell. To date, pairwise perturbation GI screens have mapped over 200,000 gene pairs in human cells (68).

A common strategy for pairwise combinatorial screens involves custom lentiviral libraries where each vector expresses two gRNAs targeting distinct genes (68–70,188,192–199).

Researchers may use either a single CRISPR enzyme or orthogonal Cas proteins to support such multiplexing, allowing for tighter and more independent control of gain- and loss-of-function perturbations (198,200). A less frequently used approach relies on high MOI with single gRNA lentiviruses, generating random pairwise and higher-order perturbations (201).

Because the number of gRNAs per gene and the number of target genes exponentially increase the number of possible combinations, combinatorial CRISPR screens are typically performed in two modes: square combinatorial screens, where every element in one position is tested against every element in the other, or anchored combinatorial screens, where one gene or perturbation remains fixed while all others are systematically varied (68,70,200,202,203). While most combinatorial CRISPR GI screens are squared, it is infeasible to experimentally test all pairwise interactions between ~20,000 human genes, resulting in ~400 million combinations even without considering that multiple gRNAs per gene are required in such a library (2). Furthermore, the GI landscape is sparse, with pairwise studies finding less than 0.1% of GIs between unrelated genes (70). As a result, these screens have focused until now on small subsets of target genes involved in critical biological processes or disease pathways (2).

Combinatorial CRISPR screens have been widely used to explore GIs, where gene pairs interact either synergistically or antagonistically, and to identify synthetic lethal interactions that have significant implications for cancer therapy (68,70,191,193,195,198,199,203–207). These screens have been conducted using either paired or single gRNA expression cassettes, benefiting from CRISPR systems like Cas12a (208,209). Furthermore, combining gain-of-function and loss-of-function perturbations within the same screen adds directionality information to GI data, which can be accomplished using orthogonal Cas enzymes or sgRNA scaffolds with effector recruitment domains (198,210). Recent advancements in these methods have facilitated systematic GI mapping to reveal novel cancer dependencies and vulnerabilities using single and orthogonal Cas systems (197,211,212).

Several computational models have been developed to calculate and interpret GIs. The additive and multiplicative models assume that the expected phenotype of a double mutant is the sum or product of the single-mutant phenotypes, respectively. In contrast, the Log model assumes a logarithmic fitness scale, while the Min model assumes that for non-interacting genes, the fitness of the double mutant should be similar to that of the more severe single mutant. Although these models often agree with each other, they can diverge under specific conditions; hence, the need for context-sensitive approaches (213).

While most GI screens focus on viability as a readout, this study parameter is not suitable for studying complex regulatory mechanisms where multiple genes work together to regulate cellular function (2). Advances in single-cell CRISPR screening methods, such as Perturb-seq, now allow the investigation of gene regulatory functions and combinatorial effects in a transcriptome-wide context (69,214–216).

1.7.3 Single-cell CRISPR-seq screens (scCRISPR-seq)

In the past decade, bulk RNA-seq has been widely used to study gene expression at the population level, since it averages signals across many cells. The development of single-cell RNA sequencing (scRNA-seq) has enabled gene expression analysis at the single-cell level, making it especially valuable for studying cellular heterogeneity (217). scCRISPR-seq screening combines the scale of pooled CRISPR approaches with scRNA-seq to enable high-resolution measurement of high-dimensional phenotypes in response to gene perturbation, allowing the construction of comprehensive genotype-phenotype maps for studying transcriptional networks and GIs at unprecedented resolution. These maps often reveal subtle regulatory effects behind complex phenotypes such as changes in transcriptional programs, RNA splicing and processing, cell differentiation, transcriptional heterogeneity, and cell-cycle states (218,219).

Large-scale functional screens aim to organize genes into pathways or complexes. As such, a major strength of scCRISPR-seq screens lies in their ability to capture complex phenotypes that can be either data-driven, allowing clustering genes into shared programs, or hypothesis-driven, such as studying gene splicing defects. Once a phenotype of interest is defined, researchers can use the same dataset to trace its genetic underpinnings, much like a forward genetic screen, while also gaining the depth and interpretability typical of reverse genetics (218). When scCRISPR-seq data is combined with network modeling, it allows reverse-engineering of transcriptional networks and identifying regulatory hubs or critical nodes that are key mediators of phenotypic outcomes, thus complementing other high-throughput GI screens for advancing our understanding of the molecular basis of GIs and their implications for cellular function and disease. Several technologies have been developed to implement scCRISPR-seq screens, each offering different strategies for linking genetic perturbations to single-cell transcriptomic profiles.

1.7.3.1 Perturb-seq

Perturb-seq was the first method to integrate CRISPR-based genetic perturbations with scRNA-seq to quantify transcriptomic responses at single-cell resolution (215). It is based on droplet-based scRNA-seq platforms such as 10x Genomics, which encapsulate

individual cells in droplets containing barcoded oligo(dT) primers that capture cellular polyadenylated transcripts. Reverse transcription generates uniquely barcoded cDNA from each cell, which is then pooled, amplified, and sequenced (214,215,220).

During traditional pooled CRISPR screens, the sgRNAs are transcribed by RNA polymerase III (Pol III) rather than RNA polymerase II (Pol II). Therefore, they lack polyadenylation and cannot be captured by standard scRNA-seq methods (221). To address this problem, the gRNA expression vector was modified by inserting a unique guide barcode downstream of the reporter, which is then transcribed together with the guide barcode from a Pol II promoter. This enables the barcode to be captured alongside the cell's transcriptome during sequencing and subsequently identify the sgRNA identity (215). However, Perturb-seq has a risk of barcode-guide recombination due to template switching during lentiviral production or transduction limiting the use of this technology. This recombination increases with higher physical distance between the gRNA and barcode loci, causing potential misassignment of perturbations to cells and confounding downstream analysis (222).

1.7.3.2 CROP-seq

CROP-seq was developed to overcome the technical issues of Perturb-seq. Instead of relying on a separate barcode to identify perturbations, CROP-seq enables direct detection of the sgRNA itself in the scRNA-seq readout, by using a lentiviral vector in which the sgRNA expression cassette is placed within the 3′ long terminal repeat (LTR) of the viral genome and expressed by a Pol III promoter. This cassette is then duplicated into both the 5′ and 3′ LTRs upon integration into the host genome, increasing guide expression levels (223). The upstream reporter gene, expressed from a Pol II promoter such as EF-1α, includes the sgRNA cassette within its 3′ untranslated region (UTR) and as a result, the sgRNA sequence is transcribed as part of a polyadenylated mRNA and captured by standard poly(dT)-primed scRNA-seq protocols, making the sgRNA both the perturbation and its own identifier (223,224).

However, there are two disadvantages with CROP-seq. First, the 3′ LTR has a constrained capacity for insert size, which for example, can restrict the number of sgRNA cassettes used in combinatorial perturbation experiments. Long inserts into the 3' LTR are associated with decreased lentiviral titers and inefficient viral integration. Second, CROP-seq tends to show slightly lower recovery of sgRNAs compared to barcode-based methods, especially for weakly expressed guides. To overcome this issue, hemi-nested PCR is used to enrich the sgRNA-containing Pol II transcript, allowing for accurate guide identification and assignment (222).

1.7.3.3 Direct capture Perturb-seq

Direct capture Perturb-seq eliminates the need for barcodes or inserting sgRNAs in polyadenylated transcripts. Instead, it enables direct hybridization and amplification of modified gRNA sequences during library preparation by adding a specifically designed capture sequence to either the loop or the 3' end of the sgRNA tracr sequence without disrupting its CRISPR functionality (225). The capture sequence serves as an annealing site for cDNA oligos that are co-inserted onto the same barcoded beads used for mRNA capture, allowing the gRNA and transcriptome to be linked through the same cell barcode (Fig. 7).

Compared to previous methods, direct capture Perturb-seq has higher sensitivity and gRNA capture efficiency, which enables robust perturbation assignment across thousands of cells. This leads to reduced background noise by minimizing the misassignment of gRNAs to non-target cells. It is also supported by popular commercial workflows and platforms such as 10x Genomics. However, it has some limitations such as the possibility of lower sgRNA editing efficiency caused by the use of modified gRNA scaffolds. Therefore, it usually requires testing and optimization of different capture sequences and insertion positions to maintain high on-target and sgRNA capture efficiency. In addition, libraries cloned with these modified tracrRNA sequences are often not cross-compatible across different single-cell platforms which limits their flexibility (226,227).

One limitation of traditional scRNA-seq is that it requires sequencing the entire transcriptome, which can be inefficient and expensive when only specific or low expressed genes are of interest. To address this, Tap-seq was introduced which uses semi-nested gene specific PCR primers combined with a general adapter to amplify both gRNA sequences and up to 1,000 selected transcripts per cell. This enrichment of transcripts improves sensitivity and cost effectiveness and offers a much higher signal-to-noise ratio compared with previous scCRISPR-seq methods (228).

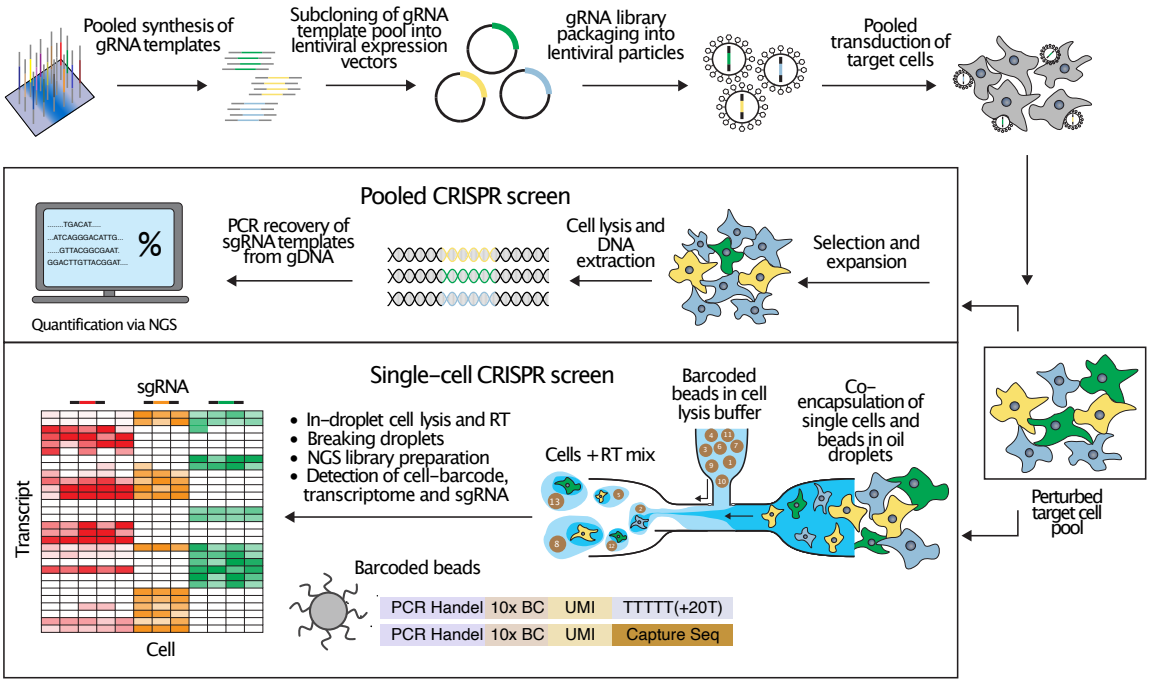


Figure 7 - Overview of pooled and single-cell CRISPR screening workflows.

*A pooled library of gRNA templates is synthesized as ssDNA oligos and cloned into lentiviral expression vectors, which are then packaged into lentiviruses used to transduce target cells. In the **pooled CRISPR screen** (top), transduced cells are selected and expanded for a predefined time frame, followed by gDNA extraction and PCR-based recovery of gRNA sequences. NGS is then used to quantify the guide representation to identify genes involved in the phenotype of interest. In the **scCRISPR screen** (bottom), the perturbed cell pool is encapsulated with barcoded beads in water-in-oil droplets on a microfluidic chip. Cell lysis and reverse transcription (RT) occur within these droplets, enabling the capture of mRNA and gRNA information on the barcoded beads. This permits NGS-based detection of cell barcodes, transcripts, and sgRNAs, facilitating high-resolution mapping of genetic networks at the single-cell level.*

1.8 Aim of the study

The development of CRISPR systems for gene perturbation enabled the investigation of single gene function at an unprecedented scale. However, the complexity of cellular systems necessitates higher-order, high content, perturbation technologies to study how genes work together to regulate cell function. This thesis explores two CRISPR-based functional genomics platforms for mapping GIs and reconstructing transcriptional regulatory networks in oncogenic signaling pathways: (1) the use of the Cas13d RNA-targeting CRISPR system for GI mapping, and (2) the application of scCRISPRseq to investigate transcriptional responses in the RAF-MAPK signaling pathway.

In Chapter 1, we explore Cas13d, an RNA-targeting CRISPR platform, as a tool for quantitative GI mapping. GI mapping in human cells is essential for understanding the functional relationships between genes and identifying therapeutic vulnerabilities in disease-relevant pathways. We begin by comparing the performance of Cas13d to the

more common DNA-targeting CRISPR systems, Cas9 and Cas12a, to establish Cas13d's efficiency as a tool for dual gene perturbation. We hypothesized that Cas13d's RNA-targeting capabilities would enable precise and uniform gene perturbations, overcoming the limitations of DNA-targeting CRISPR systems that perturb genes by introducing DSBs at genomic target sites. Next, we use Cas13d to perform multiple GI screens to systematically study genetic dependencies between six genes affecting drug sensitivity in the CML cell line K562 under both imatinib-treated and untreated conditions. In these screens, we take advantage of Cas13d's gRNA array processing ability to investigate the utility of concatenated dual and triple gRNA arrays for dual gene and same-gene targeting, respectively, and test different dual gRNA expression strategies to avoid gRNA-gRNA sequence-dependent interference. This study aims to establish Cas13d as a fast and reproducible platform for high-fidelity GI mapping and offers methodological insights for designing future combinatorial Cas13d CRISPR screens.

In Chapter 2, we employ direct capture Perturb-seq, which integrates CRISPR perturbation with single-cell transcriptomics to investigate transcriptional networks downstream of RAF-MEK-ERK signaling. While the upstream regulators of key signaling pathways are well-characterized, the downstream transcriptional responses that influence cell fate decisions remain poorly understood. To address this issue, we study the transcriptional effects of perturbing 22 transcription factors downstream of RAF1 in an inducible RAF-MAPK signaling model in HEK293 cells. Using a topology-based modeling approach, we aim to construct a *de novo* network model that identifies central regulators of the MAPK signaling pathway downstream of RAF, as well as feedback loops between our selected transcription factors. This study shows how scCRISPR-seq screening can provide a mechanistic view of how external signaling is transformed into structured gene expression programs and highlights key nodes that govern transcriptional output.

Collectively, these chapters demonstrate how distinct experimental and computational strategies, such as GI mapping using Cas13d and regulatory network inference through scCRISPR-seq, can independently enhance our understanding of gene interactions in mammalian systems. While methodologically separate, they contribute to the wider goal of decoding functional GIs and pathway cross-talk.

Parts of this thesis have been published at the time of writing. Chapter 1 is based on the article El Kassem *et al*. 2025 (229), and Chapter 2 is based on the preprint El Kassem *et al*. 2024 (230). The text and figures presented in these chapters are adapted in part or in full from the respective publications. Chapter 2 was carried out in collaboration with the Blüthgen Lab at Charité - Universitätsmedizin Berlin, who contributed to both the experimental design, preliminary experiments, and computational analysis.

Chapter 1: Evaluation of Cas13d as a tool for genetic interaction mapping

1.1 Abstract

Genetic networks represent a complex web of interactions between genes and their regulatory elements governing various cellular functions and biological processes. Mapping genetic interactions (GI) is crucial for understanding this complexity by revealing how different genes influence the phenotype of one another, elucidating functional relationships, and identifying key regulatory pathways that control cellular processes. In this study, we investigate the utility of Cas13d, an RNA-targeting CRISPR system, for quantitative GI mapping. First, we compare it to the two most commonly used DNA-targeting nucleases, Cas9 and Cas12a, and find that it produces faster gene perturbations and results in more uniform dual gene perturbed cell populations. By leveraging Cas13d's gRNA array processing ability, we targeted different genes but encountered gRNA-gRNA interference, which we overcame by employing a dual promoter gRNA expression strategy. Finally, we maximized Cas13d knockdown efficiency by combining multiple gRNAs targeting the same gene into a single array, leading to a nine-fold decrease in our library size and larger GI score effect sizes. Together, these strategies enabled the reproducible quantification of GIs between six genes that regulate the chronic myeloid leukemia cell (CML) line K562 response to the BCR::ABL1 kinase inhibitor imatinib. Our findings demonstrate the potential of Cas13d for GI mapping, offering promising insights into drug response pathways with therapeutic relevance.

1.2 Introduction

Understanding GIs is crucial for unraveling the complexity of cellular systems, deciphering disease mechanisms, and identifying novel combinatorial therapeutic strategies. GI mapping has emerged as a powerful method for systematically linking genotype to phenotype and studying the functional relationships between genes that contribute to human disease formation (67,231). A GI occurs when the combined loss-of-function phenotype resulting from systematically perturbing two genes deviates from the expected phenotypes of the single-gene mutants. Beyond elucidating gene functions and therapeutic targets, GI mapping offers critical insights into cellular networks, paving the way for personalized medicine applications.

Despite its potential, the difficulty of systematically perturbing genes in human cells limited GI studies to model organisms such as yeast and *Drosophila* during the early years of

these studies (73,232). The discovery of RNA interference (RNAi) and the adaptation of this technology to systematically modulate gene expression changed this, enabling the study of GIs in mammalian cells. However, RNAi has some limitations due to off-target effects and the relatively low and variable knockdown efficiency of its constructs (192). This changed with the discovery of more precise CRISPR-based gene editing technologies, which rapidly replaced RNAi in GI studies. The Cas9 nuclease from *Streptococcus pyogenes* is the most widely used and studied CRISPR system (70,195,196). However, other systems, such as the Cas12a (Cpf1) nuclease from *Acidaminococcus sp.*, were subsequently discovered and adapted for GI studies as well (197,208). Furthermore, combinations of different CRISPR systems have also been utilized, such as Cas9 and Cas12a (203) and Cas9 with other orthologous CRISPR/Cas9 systems from various microorganisms (198,199).

More recently, RNA-targeting CRISPR systems like Cas13 have been discovered and adopted for precise modulation of gene expression at the RNA level (233). The type VI-D CRISPR-Cas Cas13d system from *Ruminococcus flavefaciens,* also known as CasRx, has rapidly become the most widely used of these systems for precise and efficient RNA targeting gene expression interference (150). In addition to its precision, Cas13d's ability to target non-coding RNAs broadens the scope of GI mapping, offering insights into regulatory elements that influence cellular functions without permanently altering the genome. To date, several studies have demonstrated the utility of Cas13d in pooled genetic screens targeting coding (234) and non-coding transcripts (235), as well as in single-cell combinatorial Perturb-seq screens (152) and massively multiplexed RNA knockdown combinatorial screens in primary human T cells (165). However, its potential for quantitative GI mapping remains unexplored.

For a CRISPR system to be suitable for GI mapping, it must meet several prerequisites. First, the system should generate a uniform population of double-perturbed cells by specifically and homogeneously perturbing two genes within the same cell. The efficiency of double perturbation inherently depends on the efficiency of single perturbations by each gRNA through a multiplicative relationship. Specifically, the efficiency of double perturbation is the product of the efficiencies of single perturbations, meaning that any reduction in the efficiency of either single perturbation significantly lowers the probability of achieving successful double perturbations. DNA-targeting enzymes such as Cas9 and Cas12a perturb gene function by introducing double-strand DNA breaks (DSBs) at the genomic target sites. These DSBs are typically repaired by error-prone cellular DNA damage repair mechanisms like non-homologous end joining (NHEJ) (113) often leading to the formation of INDELs at the target site. This leads to heterogeneous cell populations

with frameshift and in-frame mutations that, in some cases, translate into a functional protein (236). In contrast, Cas13d is an RNA-targeting CRISPR system and therefore overcomes these limitations by acting directly on RNA transcripts. This approach ensures consistent and uniform reduction in target transcript levels across the cell population, leading to a homogeneous perturbed cell population.

Another requirement for quantitative GI mapping is the absence of sequence-specific interference between the gRNAs targeting different genes within the same cell. When the activity of one gRNA depends on the sequence of the second gRNA, it hinders the quantification of GI scores, which are calculated based on the difference between observed double-perturbation phenotypes and those expected from individual perturbations. Unlike Cas9, Cas12a (237) and Cas13d (150) can process multiple concatenated gRNA arrays expressed from a single promoter into single effector gRNAs. While Cas12a gRNA arrays have been used for GI mapping (197,208), the applicability of Cas13d gRNA arrays for this purpose remains unstudied, despite having already been used for multiplexed gene perturbation (152,165).

In this study, we investigate the utility of Cas13d for GI mapping. We hypothesized that Cas13d's RNA-targeting capabilities would enable precise and uniform gene perturbations, overcoming the limitations of DNA-targeting CRISPR systems. We demonstrate that Cas13d induces perturbations faster and more homogeneously than Cas9 or Cas12a, generating highly uniform double-perturbed cell populations. However, we observed that concatenating Cas13d gRNAs into arrays and expressing them from a single promoter can result in sequence-dependent interference, where the efficiency of one gRNA depends on the sequence of the second gRNA in the array, rendering such arrays unsuitable for GI mapping. To address this, we show that expressing individual gRNAs from different promoters eliminates the interference. Additionally, we find that utilizing arrays containing three concatenated gRNAs targeting the same gene enhances knockdown efficiency and proliferation phenotypes while minimizing library size. To this end, we conducted a series of GI screens using single gRNAs or single-gene arrays expressed from separate promoters to systematically map GIs between six genes involved in the response of the CML K562 cell line to the tyrosine kinase inhibitor imatinib. Our results show that single-gene arrays produced larger GI effect sizes compared to single gRNAs, with reproducible and consistent GI scores within each approach and across the two gRNA expression approaches. We established and benchmarked two Cas13d-based strategies for quantifying GIs in therapeutically relevant oncogenic signaling pathways.

1.3 Results

1.3.1 Single gene perturbation properties of Cas9, Cas12a, and Cas13d

DNA-targeting nucleases Cas9 and Cas12a perturb their targets through DSBs that are repaired by the DNA DSB repair pathways introducing frameshift mutations. In contrast, the RNA-targeting ribonuclease Cas13d autonomously degrades its RNA targets, resulting in gene knockdown rather than knockout (Fig. 6). To analyze how these differences in gene perturbation mechanisms affect the performance of single and double perturbations, we compared the perturbation kinetics between the three systems. For this purpose, we introduced the components of the three CRISPR systems into the chronic myeloid leukemia (CML) cell line K562 through lentiviral transduction at a low multiplicity of infection (MOI<0.3), targeting cell surface markers, and measured their expression levels by flow cytometry.

Representative histograms in Figure 8A illustrate how Cas13d reduces CD46 protein levels faster and more uniformly compared to the DNA-targeting nucleases Cas9 and Cas12a. These differences arise from the distinct mechanisms of action of DNA- versus RNA-targeting CRISPR systems. The nearly random nature of the DNA DSB repair outcomes leading to Cas9 and Cas12a-mediated knockout results in a mosaic of bimodal cell populations, featuring both loss-of-function mutations and wild-type levels of CD46. In contrast, Cas13d creates a homogeneous cell population with reduced CD46 levels within three days post gRNA transduction, due to its RNA-degrading mechanism of action (Fig. 8A).

Next, the faster kinetics of gene perturbation with Cas13d compared to Cas9 or Cas12a were demonstrated by targeting the *CD46*, *CD47*, *CD63*, and *CD71* genes with each system and measuring the cell surface protein expression level over time using flow cytometry (Fig. 8B, Supp. Fig. 1). Three days after transduction, minimal to no reduction in target protein levels was observed for either of the two DNA-targeting nucleases, while Cas13d showed an almost complete knockdown. Depending on the target, Cas9 and Cas12a nucleases achieved maximum target protein reduction by day 5 or day 7 post-transduction. For the essential gene *CD71* (K562 DepMap Chronos score: -0.975), Cas9 and Cas12a mediated knockout resulted in a decrease in edited cell populations by day 7 or day 10, respectively. In contrast, Cas13d produced a stable knockdown between 79% and 87% throughout the entire experiment. These results indicate that the complete knockout of CD71 by Cas9 and Cas12a leads to decreased viability of K562 cells. Conversely, a greater than 80% reduction in CD71 levels does not which suggests that

"fine-tuned" Cas13d knockdown could be useful for investigating the function of genes whose complete deletion is cytotoxic.

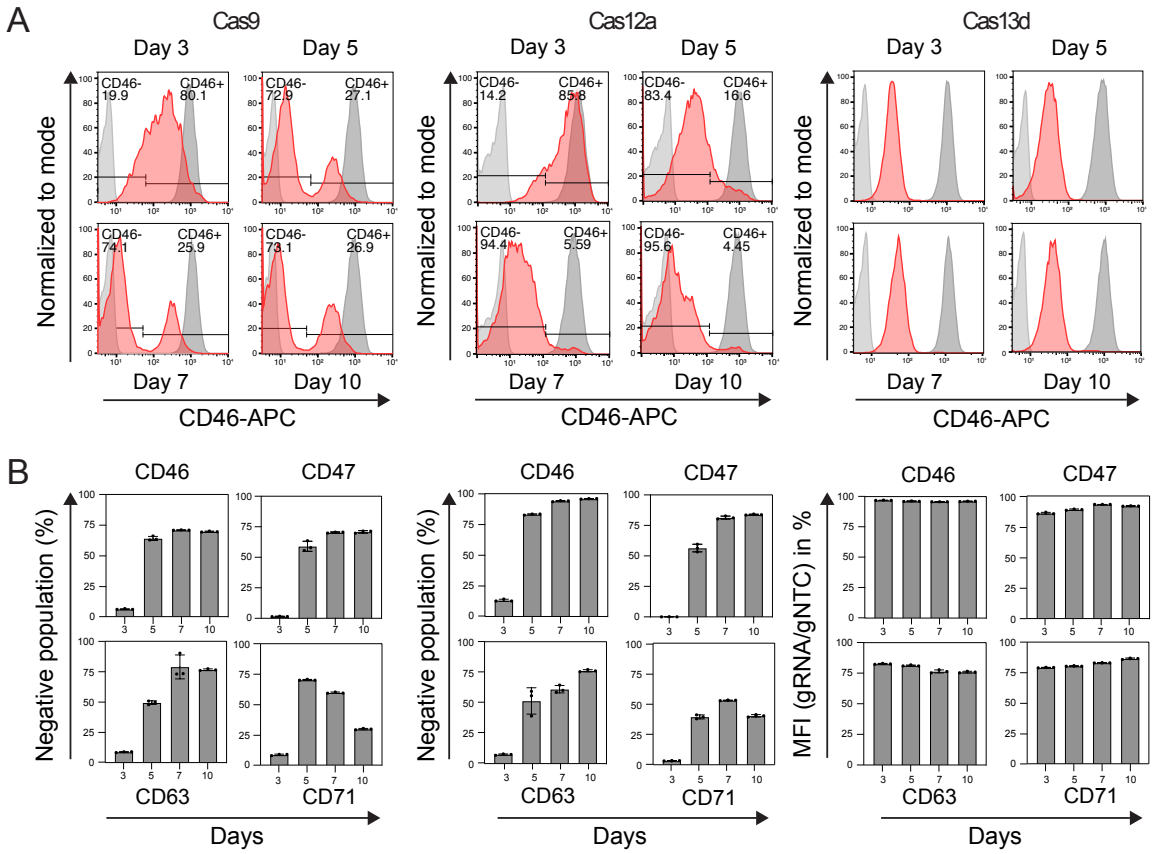


Figure 8 - Comparison of single gene perturbation properties of Cas9, Cas12a, and Cas13d.

***(A)** Histograms of cells perturbed with Cas9, Cas12a, and Cas13d targeting CD46, with protein levels measured via flow cytometry at four different time points post-transduction with the respective CRISPR system. Light grey: Unstained untransduced cells. Red: Perturbed cells. Dark grey: Stained non-target control cells. **(B)** Gene perturbation kinetics of Cas9, Cas12a, and Cas13d targeting CD46, CD47, CD63, and CD71 over 10 days. MFI = Mean fluorescence intensity. Values represent the mean of biological replicates; error bars show SD (n=3).*

1.3.2 Double gene perturbation properties of Cas9, Cas12a and Cas13d

The distinct mechanisms of target gene perturbation of the three CRISPR systems become particularly significant when more than one gene is perturbed per cell, as required for GI mapping. By simultaneously targeting two genes within the same cells, Cas13d was able to generate the most uniform double-perturbed cell populations compared to both DNA-targeting CRISPR systems, making it a promising tool for GI mapping. Figure 9A shows the distribution of cells after single or double gene perturbation from all three tested CRISPR systems of CD46 and CD47. In the case of double perturbation with Cas9 and Cas12a, a mix of unperturbed, single-perturbed, and double-perturbed cell populations remained. In contrast, for Cas13d, the entire cell population uniformly shifted towards

double negative (Fig. 9A, Supp. Fig. 2). The percentages of double-perturbed cell populations were 51.2% +/-5 for Cas9, 81.6% +/-0.7 for Cas12a, and 95.4% +/-0.9 for Cas13d (Fig. 9B).

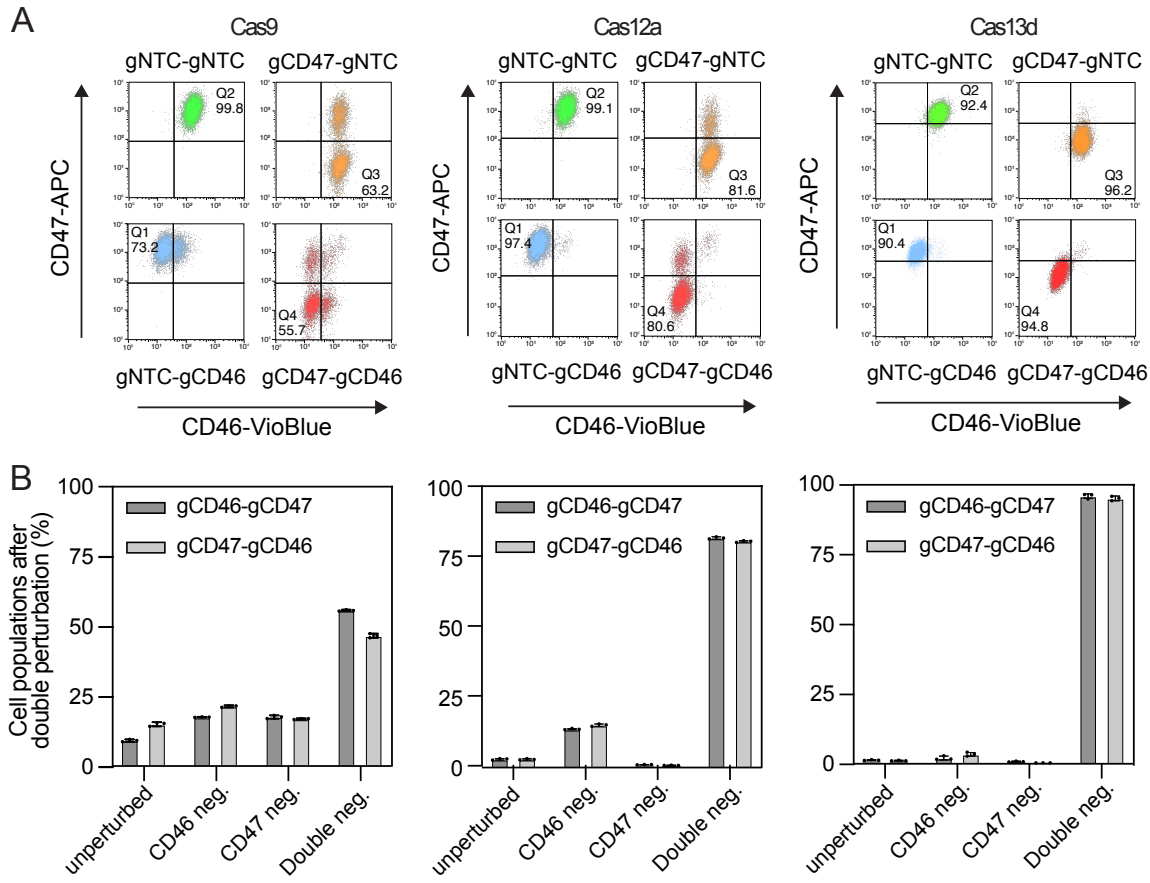


Figure 9 - Comparison of double gene perturbation properties of Cas9, Cas12a, and Cas13d.

***(A)** Distribution of cell populations expressing combinations of CD46, CD47, and NTC gRNAs as indicated. Cell surface marker levels were quantified via flow cytometry at 10 days post-transduction with the respective CRISPR system. **(B)** Quantification of unperturbed, single-perturbed, and double-perturbed subpopulations from cells expressing gRNAs against CD46 and CD47 at 10 days post-transduction. Values represent the mean of biological replicates; error bars show SD (n=3).*

1.3.3 Concatenated Cas13d gRNAs show sequence-dependent interference

Multiple gRNAs can be concatenated into arrays that the Cas13d effector protein processes into individual functional single gRNAs, which can target different RNA constructs as needed for GI mapping (150) (Fig. 10A). Therefore, we assessed the utility of Cas13d gRNA arrays for this purpose using dual gene combinatorial proliferation screening. After determining the proliferation scores (tau) for single and double gene perturbation phenotypes, the GI score between two genes is calculated as the deviation of the measured phenotype of the perturbation of both genes in the same cell from the expected phenotype calculated from the measured perturbation phenotype of each gene

individually. To accurately calculate GI scores, two gRNAs must act identically on their target RNA, irrespective of the sequence of the second gRNA expressed in the same array.

To assess the applicability of Cas13d arrays for GI mapping, gRNAs were designed against six genes whose involvement in the imatinib response of the CML cell line K562 was previously established (198). K562 cells carry a translocation between chromosomes 9 and 22, creating the *BCR::ABL* fusion oncogene, a constitutively active tyrosine kinase that drives uncontrolled cell proliferation (238). Imatinib, a targeted BCR::ABL inhibitor, blocks this aberrant signaling pathway and is used to treat CML (239). Three genes whose perturbation sensitized K562 cells to imatinib treatment, namely, *(BCR)-ABL1*, the direct target of imatinib, *GAB2,* and *SOS1 were* selected, along with three genes whose perturbation made the cells less sensitive to imatinib: *PTPN1*, *NF1,* and *SPRED2*. 27 gRNAs against each target gene were designed using the Cas13d gRNA design rules previously described (234) (Supp. Table 1). The gRNAs were symmetrically cloned into a concatenated dual gRNA array library expressed from a single mouse U6 (mU6) promoter (Fig. 10B). The library showed a narrow distribution, with 99.8% of all elements falling within one order of magnitude (Supp. Fig. 3A). Using this library, we performed a pooled combinatorial CRISPR screen as described in Figure 10C. Each dual gRNA array targeted either a gene in position 1, combined with one of 30 different non-target control gRNAs (gNTC) in position 2 (U6-gRNA-gNTC), or vice versa (U6-gNTC-gRNA).

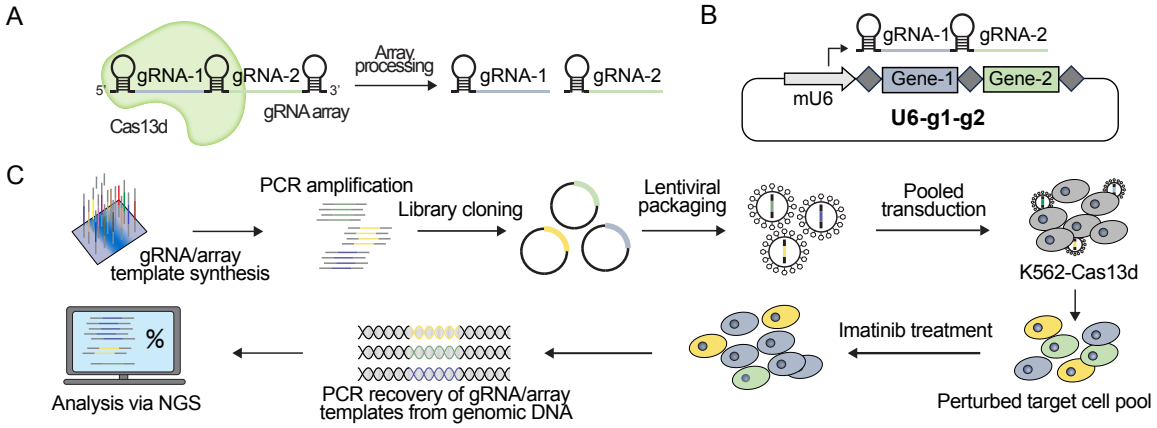


Figure 10 - Overview of the pooled Cas13d screening pipeline in K562.

(A) *Schematic representation of Cas13d processing concatenated gRNA arrays into single gRNAs.* ***(B)*** *Schematic of the dual gRNA concatenation strategy U6-g1-g2. mU6 = mouse U6 promoter, diamonds = direct repeat DR36, squares = target-specific spacer sequence.* ***(C)*** *Overview of the pooled Cas13d screening pipeline. In summary, gRNA template oligonucleotide pools were amplified by PCR and cloned into a lentiviral expression vector. K562 cells stably expressing RfxCas13d were transduced with lentiviral pools at a low MOI to ensure the integration of a maximum of one gRNA expression cassette into the host cell genome. Perturbed target cell pools were treated for 19 days with escalating doses of imatinib, ranging from 100 nM to 300 nM, before gRNA expression cassettes were recovered via PCR from the gDNA. The abundance of gRNA expression cassettes at the beginning (baseline) and in imatinib-treated or untreated cells at the end of the screen was quantified through next-generation sequencing (NGS) of the gRNA expression cassette pools to determine the enrichment or depletion of cells expressing a specific gRNA combination.*

To assess the reproducibility of our screen data, we examined the correlation of replicate normalized read counts and tau scores at different levels of data analysis across untreated and treated conditions. A strong correlation was found between the normalized read counts of both replicates in untreated samples (r = 0.88), while treated samples exhibited slightly lower correlation scores (r = 0.74) (Fig. 11A, 11B). Similarly, high correlations were found between the tau values across replicates before and after filtering in both untreated (r = 0.81 and r = 0.79, respectively) and treated conditions (r = 0.89 and r = 0.9, respectively) (Fig. 11C-F). To evaluate the consistency and robustness of the calculated tau scores, we compared tau values derived from the same gRNAs in both array positions, namely gRNA-gNTC (position 1) and gNTC-gRNA (position 2). The treated samples showed a stronger correlation (r = 0.73) than the untreated ones (r = 0.38); however, the overall heterogeneity in correlations indicated position-dependent performance differences of the same gRNA (Fig. 11G-H, Supp. Data 1).

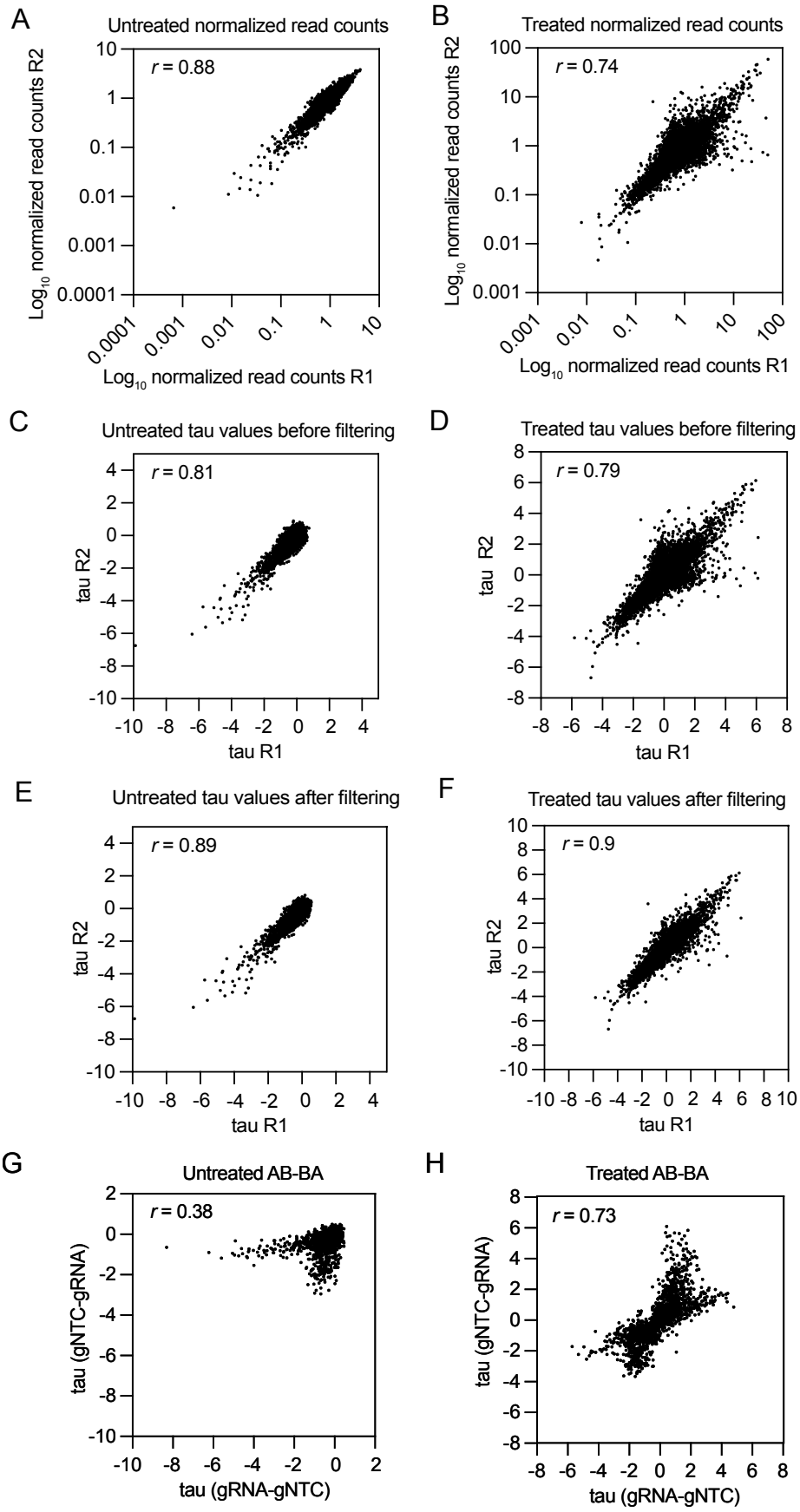


Figure 11 - Correlation plots at different levels of screen data analysis showing high technical reproducibility of the U6-g1-g2 screens.

(A-B) *Correlation between normalized read counts from two technical screen replicates following 19 screen days in untreated (A) and imatinib-treated (B) conditions using the U6-g1-g2 strategy.* ***(C-D)*** *Correlation between tau values from two technical screen replicates following 19 days of screening in untreated (C) and imatinib-treated (D) conditions before filtering for functional gRNAs using the U6-g1-g2 strategy.* ***(E-F)*** *Correlation between tau values from two technical screen replicates following 19 screen days in untreated (E) and imatinib-treated (F) conditions after filtering for functional gRNAs using the U6-g1-g2 strategy.* ***(G-H)*** *Correlation between tau values from gRNA-gNTC and gNTC-gRNA combinations following 19 days of screening in untreated (G) and imatinib-treated (H) conditions using the U6-g1-g2 strategy. Pearson's correlation was used to determine the r values. Screen data was analyzed in collaboration with Jasmine Hillmer.*

To further explore this issue, the three best performing gRNAs against *PTPN1* and *SOS1,* combined with 30 different gNTCs, were selected and ranked based on their performance in position 1 of the gRNA array (Fig. 12, top panel). Large variations in tau values were observed from the same gRNAs, depending on which gNTC sequence they were concatenated with in position 2. For example, the concatenation of all six selected gRNAs with gNTC-28 produced strong positive (gPTPN1) or negative (gSOS1) tau values, while the same gRNAs concatenated with gNTC-21 resulted in only minimal enrichment or depletion of cells. This trend was not observed in the opposite orientation (gNTC-gRNA) (Fig. 12, bottom panel). Nevertheless, we noted lower tau value effect sizes compared to the gRNA-gNTC orientation, and high variability between the three selected gRNAs against the same gene, where the same three gRNAs performed rather consistently in the first orientation, especially those targeting *PTPN1*.

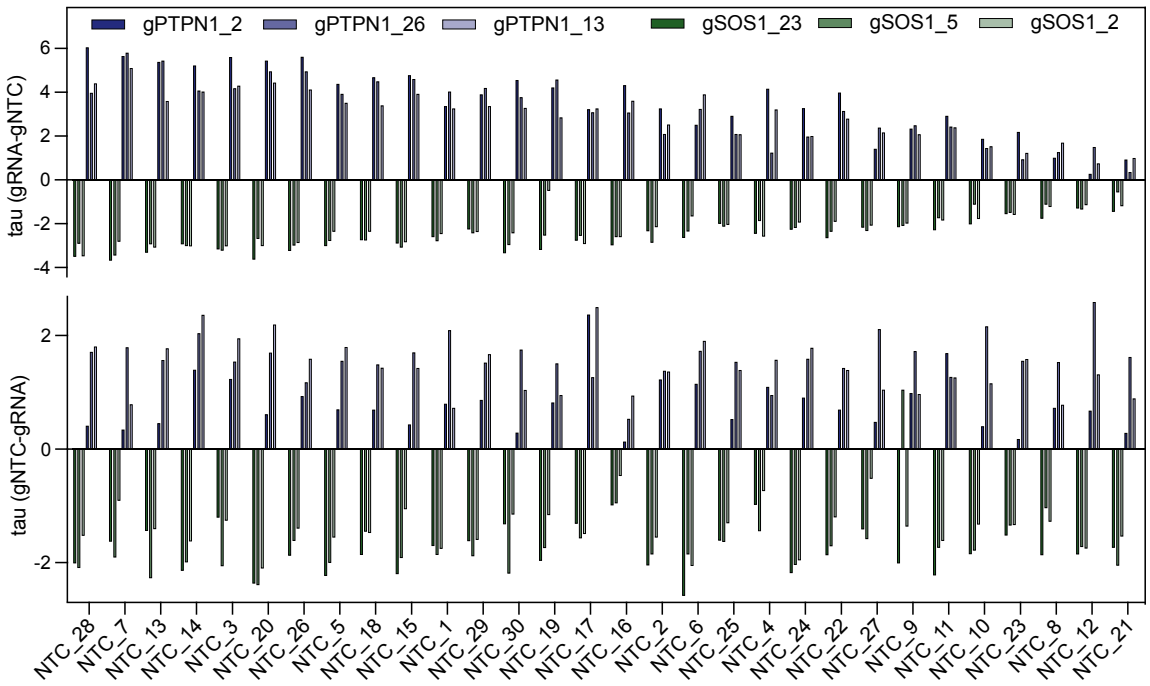


Figure 12 - Differences in gRNA efficiency dependent on position and gNTC identity.

Bar chart showing the tau values for the three best performing gRNAs targeting PTPN1 and SOS1, in combination with 30 gNTCs, presented in both orientations: gRNA-gNTC (top) and gNTC-gRNA (bottom). Data was analyzed in collaboration with Jasmine Hillmer.

To obtain a more systematic overview, tau values from the top six gene-targeting gRNAs per gene, expressed either in position 1 (Fig. 13A) or position 2 (Fig. 13C) in combination with the 30 different gNTCs in the respective other position, were calculated. Similar to the results shown in Figure 12, the effect sizes of tau values from specific gRNAs against a target gene in position 1 varied in a gNTC sequence-specific manner. This gRNA sequence-dependent interference was further demonstrated by the strong positive and negative correlations between arrays with different gene-targeting gRNAs in position 1 (Fig. 13B). On the other hand, no such effect was observed when the gene-targeting gRNA was expressed in position 2 (Fig. 13C). However, smaller effect sizes were observed when the gRNAs were expressed from position 2 compared to position 1. No systematic correlation was observed between gene-targeting gRNAs expressed in position 2, suggesting that gRNA-gRNA interference occurs only when the gene-targeting gRNA is expressed in position 1, but not position 2, probably due to the way *Rfx*Cas13d processes gRNA arrays (Fig. 11D). In summary, these results show that concatenation of Cas13d gRNAs against different target genes is not a suitable method for quantitative GI mapping due to sequence-dependent interference between gRNAs of the same array.

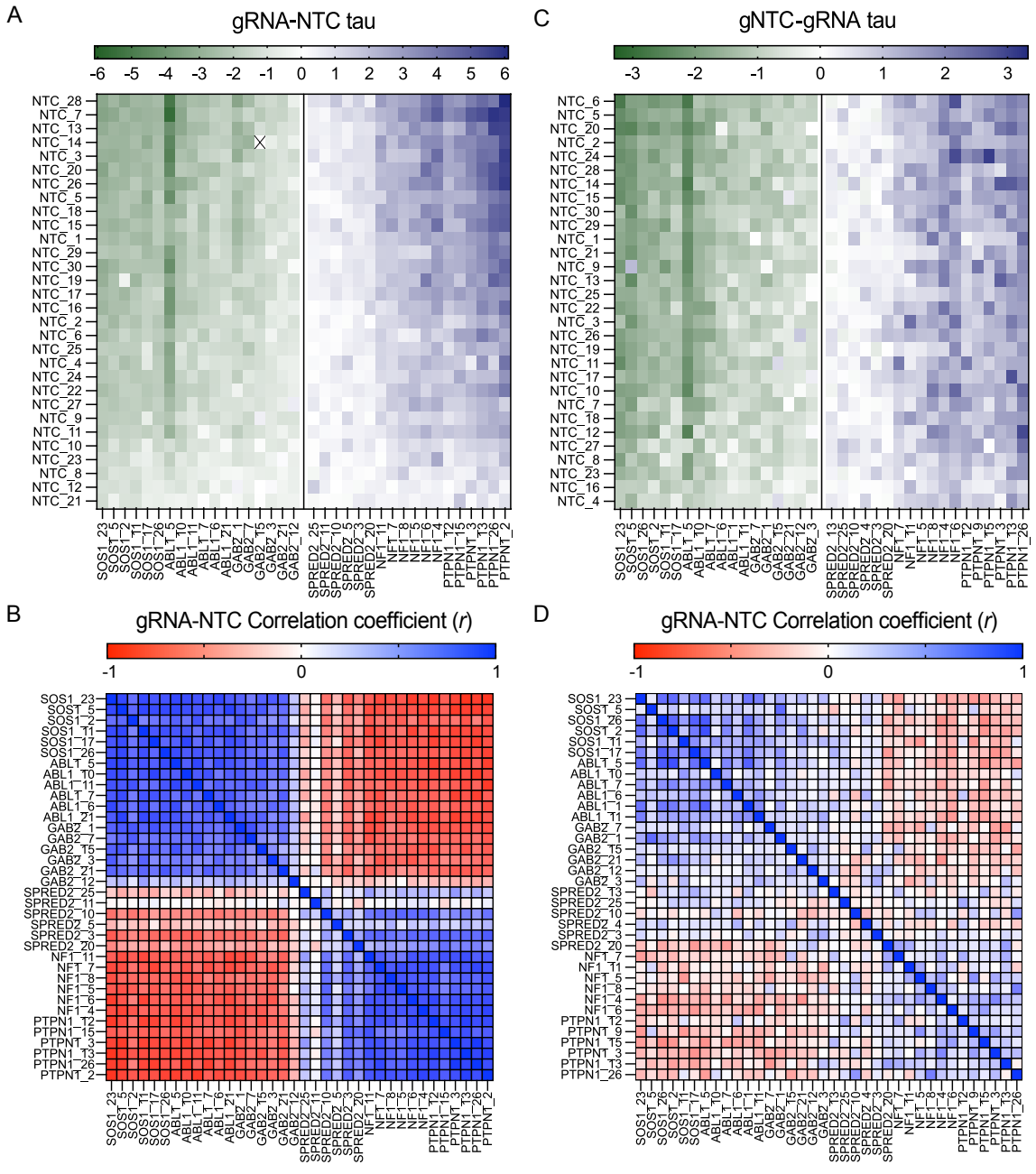


Figure 13 - Concatenated Cas13d gRNAs show sequence-dependent interference.

(A) *Heatmap showing tau values from the six best performing gene-targeting gRNAs in position g1 (x-axis), combined in one array with 30 different non-target control gRNA sequences (gNTC) in position g2 (y-axis) using the U6-g1-g2 strategy.* ***(B)*** *Pearson's correlation (r) between tau values from all columns in (A).* ***(C)*** *Heatmap showing tau values from the six best performing gene-targeting gRNAs (x-axis) in position g2, combined in one array with 30 different gNTCs in position g1 (y-axis) using the U6-g1-g2 strategy.* ***(D)*** *Pearson's correlation (r) between tau values from all columns in (C). Crossed squares (x) indicate not available (n.a.) data points. Screen data was analyzed in collaboration with Jasmine Hillmer.*

1.3.4 Expression of gRNA from separate promoters prevents gRNA-gRNA interference

To address the gRNA-gRNA interference identified with concatenated gRNAs, a modified gRNA expression strategy was implemented, in which each gRNA was expressed from a separate promoter. Different promoters were used to avoid recombination during library cloning. Specifically, a mouse U6 promoter (mU6) drove the expression of gRNA-1, while a human U6 promoter (hU6) was used for gRNA-2 (Fig. 14A). A pooled library was then cloned using this dual-promoter system, targeting the same six genes as before with 27 gRNAs per gene, assigned to either the mU6 (position 1) or hU6 (position 2) promoter (Supp. Table 1). The library showed a relatively narrow distribution, with 69.6% of all elements falling within one order of magnitude (Supp. Fig. 3B). A pooled combinatorial proliferation screen was performed under the conditions described in Figure 10C.

We first checked the quality of the screening data by evaluating the correlation of the screen replicates. High reproducibility was observed between technical replicates in the imatinib-treated and untreated conditions, as demonstrated by a strong correlation in normalized read counts (r = 0.98 after imatinib treatment) and tau values (r = 0.91 after imatinib treatment) (Fig. 14B, Supp. Fig. 4A-E). Additionally, gene-targeting gRNAs expressed in both positions revealed a similarly high correlation of 0.92 (Fig. 14C, Supp. Fig. 4F, Supp. Data 1), indicating that the performance of the gene-targeting gRNAs was not affected by the position from which they were expressed. Similarly, as was done previously with the concatenated gRNA format (U6-g1-g2), effective gRNAs were identified for all six genes. However, in this dual-promoter setup (U6-g1-U6-g2), the functional effects of active gRNAs remained consistent across both positions, regardless of the sequence of the coexpressed gNTC (Fig. 14D and 14F). Moreover, due to the absence of systematic correlation between the different gene-targeting gRNAs and the gNTCs, no sequence-specific performance differences were observed between different gene-targeting gRNAs (Fig. 14E, 14G). This suggested that expressing two gRNAs from separate promoters can overcome the gRNA-gRNA interference observed with concatenated gRNAs. Overall, these results demonstrate the usability of Cas13d for GI mapping.

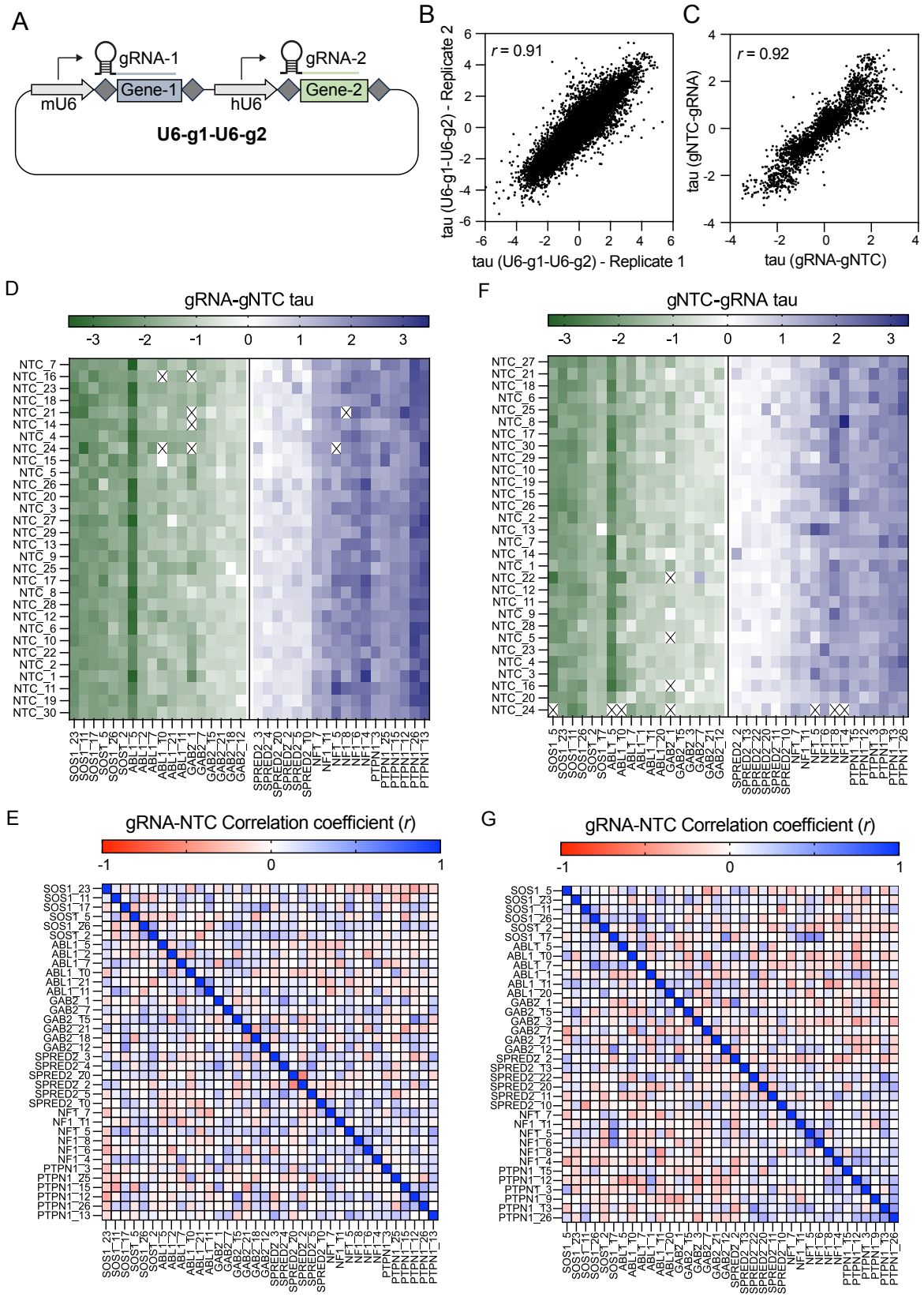


Figure 14 - Concatenated Cas13d gRNAs sequence-dependent interference can be overcome by gRNA expression from separate promoters.

(A) *Schematic of the dual promoter gRNA expression strategy mU6-g1-hU6-g2. mU6 = mouse U6 promoter, hU6 = human U6 promoter, diamonds = direct repeat DR36, squares = target-specific spacer sequence.* ***(B)*** *Correlation between tau values from two independent screen replicates after 19 days of imatinib treatment using the U6-g1-U6-g2 strategy. Pearson's correlation was used to determine the r value.* ***(C)*** *Correlation between tau values from gRNA-gNTC and gNTC-gRNA combinations after 19 days of imatinib treatment using the U6-g1-U6-g2 strategy. Pearson's correlation was used to determine the r value.* ***(D)*** *Heatmap of tau values from gene-targeting gRNAs expressed from the mU6 promoter co-expressed with the indicated gNTCs from hU6.* ***(E)*** *Pearson's correlation (r) between tau values from all columns in (D).* ***(F)*** *Heatmap of tau values from gene-targeting gRNAs expressed from the hU6 promoter co-expressed with the indicated gNTCs from mU6.* ***(G)*** *Pearson's correlation (r) between tau values from all columns in (F). Crossed squares (x) indicate not available (n.a.) data points. Screen data was analyzed in collaboration with Jasmine Hillmer.*

1.3.5 Single-gene arrays generate stronger knockdown and proliferation phenotypes

Although gRNA-gRNA interference prevents the use of Cas13d arrays to target two genes for GI mapping, combining multiple gRNAs that target the same gene may still be a viable approach to enhance knockdown efficiency. To explore this, we synthesized and cloned "single-gene arrays" by concatenating three gRNAs targeting the same gene into one construct (Fig. 15A). We then compared the knockdown efficiency of these arrays to that of the individual gRNAs by targeting cell surface markers CD46, CD47, CD71, and CD63 (Supp. Table 2). Across all targets, the single-gene arrays achieved more robust knockdown than any single gRNA alone, with residual protein expression dropping as low as 0.7% for CD46, 4.9% for CD47, 10% for CD71, and 17% for CD63 (Fig. 15B). To further assess the potential of single-gene arrays for GI mapping, the 27 gRNA sequences used in the previous dual promoter single gRNA library were concatenated into 9 single-gene arrays (Fig. 15C, Supp. Table 3). The dual gene targeting single-gene array library showed a narrow distribution, with 89.5% of all elements falling within one order of magnitude (Supp. Fig. 3C). This strategy resulted in a reduction of library size from 36,864 elements down to 4,096. These findings highlight the effectiveness of Cas13d single-gene arrays in enhancing perturbation phenotypes while reducing the complexity and size of the screening library.

Finally, pooled screens were conducted as shown in Figure 10C using both the single gRNA library (U6-g1-U6-g2) and the single-gene array library (U6-a1-U6-a2), with and without imatinib treatment to compare the performance of both gRNA expression strategies. Like previous screens, the data quality of the screen was assessed through Pearson's correlation analysis of the normalized read counts and tau values of the screen replicates at different stages of data analysis. The high correlation values of tau values in the imatinib-treated technical screen replicates (*r* values between 0.94 and 0.96) demonstrated the reproducibility of the screen (Fig. 15D, Supp. Fig. 6A-F). Additionally,

when comparing gRNAs targeting the same gene but expressed from different positions, a strong correlation of 0.9 was observed (Fig. 15E, Supp. Data 1), indicating that the performance of single-gene arrays was independent of the position from which it was expressed.

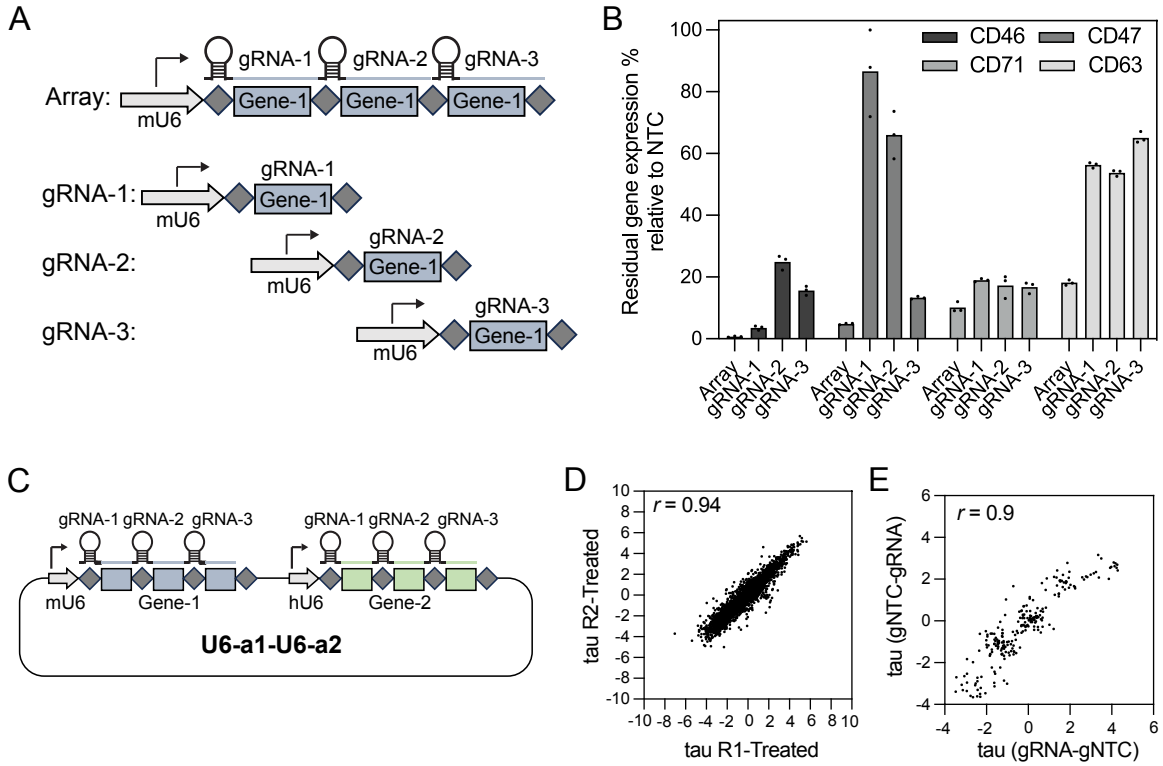


Figure 15 - Single-gene arrays generate stronger knockdown and proliferation phenotypes.

***(A)** Schematic of the single-gene array strategy. **(B)** Flow cytometric quantification of CD46, CD47, CD71, and CD63 residual protein levels in cells expressing either a single-gene array or one of the three single gRNAs concatenated in the single-gene array. Values represent the mean of biological replicates (n=3). **(C)** Schematic of the dual promoter single-gene array expression strategy used for subsequent U6-a1-U6-a2 libraries. mU6 = mouse U6 promoter, hU6 = human U6 promoter, diamonds = direct repeat DR36, squares = target-specific spacer sequence. **(D)** Correlation between tau values from two independent screen replicates, following 19 days of imatinib treatment using the U6-a1-U6-a2 strategy. Pearson's correlation was used to determine the r value. **(E)** Correlation between tau values from gRNA-gNTC and gNTC-gRNA combinations following 19 days of imatinib treatment using the U6-a1-U6-a2 strategy. Pearson's correlation was used to determine the r value. Screen data was analyzed in collaboration with Jasmine Hillmer.*

1.3.6 Cas13d perturbation shows no nonspecific proliferation phenotypes

Recent studies have demonstrated that Cas13d, specifically the RfxCas13d utilized in this work, shows collateral activity upon target binding under certain conditions. This activity leads to nonspecific RNA degradation in eukaryotic cells, which in turn reduces cell proliferation in specific cell types (240). To ensure that the observed growth phenotypes were indeed due to specific Cas13d-mediated gene knockdown, a Cas9 sgRNA library was designed and cloned to target the same six genes as the Cas13d libraries, and Cas9

counter screens were conducted as described previously (Fig. 10C). The Cas9 library consisted of four sgRNAs per gene, accompanied by five NTC and five Safe-cutter controls cloned with a narrow distribution (Supp. Fig. 3D, Supp. Table 4). Proliferation phenotypes from single gene perturbed cells in the untreated and 19 day imatinib treated screens showed a strong correlation between the Cas9, Cas13d single gRNA, and Cas13d single-gene array screens (Fig. 16, Supp. Fig. 5). Additionally, a comparison with Chronos scores from K562 DepMap data (85) also revealed a strong correlation with both the Cas9 and Cas13d results in untreated conditions. Most importantly, no instances of cell death were observed following Cas13d knockdown of a target gene that could not be confirmed by Cas9-mediated knockout and DepMap data. These results indicate the absence of cytotoxicity due to potential collateral activity in this study.

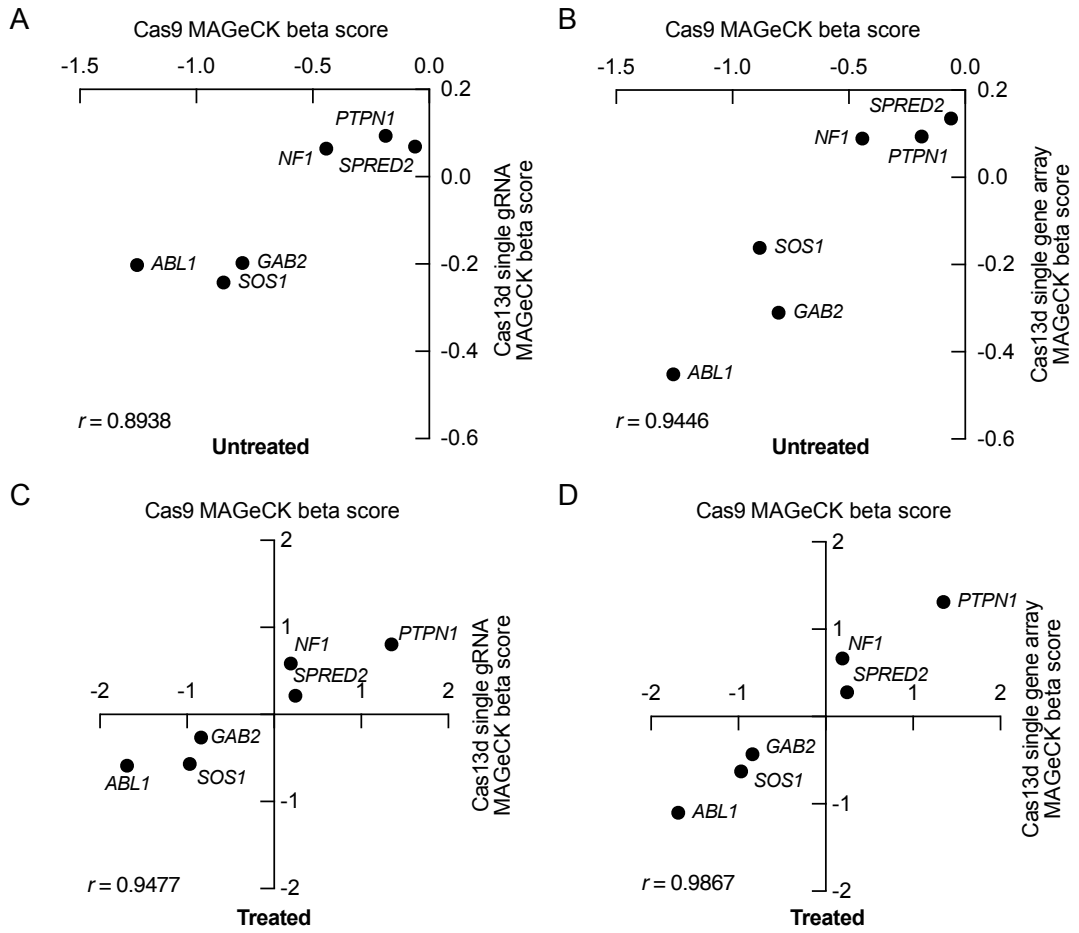


Figure 16 - Cas13d shows no signs of unspecific growth phenotypes.

***(A)** Correlation between MAGeCK MLE beta scores from the untreated Cas9 counter screen and the single gRNA screen (U6-g1-U6-g2). **(B)** Correlation between MAGeCK MLE beta scores from the untreated Cas9 counter screen and the single-gene array screen (U6-a1-U6-a2). **(C)** Correlation between MAGeCK MLE beta scores from the imatinib-treated Cas9 counter screen and the single gRNA screen (U6-g1-U6-g2). **(D)** Correlation between MAGeCK MLE beta scores from the imatinib-treated Cas9 counter screen and the single-gene array screen (U6-a1-U6-a2).*

1.3.7 Cas13d allows reproducible quantification of GIs in therapeutic signaling pathways

Moving to the identification and quantification of GIs in the signaling pathways of K562 cells under both untreated and imatinib-treated conditions. Tau values from all four screens were used to calculate GI scores (Fig. 17A, Supp. Data 1). Consistent with previously observed stronger knockdown effects of single-gene arrays compared to individual gRNAs (Fig. 15B), the tau values and GI scores derived from the single-gene arrays displayed greater effect sizes (Fig. 17B). The U6-g1-U6-g2 dual promoter single gRNA approach revealed only two GIs in the untreated condition in both orientations: *NF1-SOS1* and *NF1-GAB2.* In contrast, three additional GIs were detected with the U6-a1-U6-a2 single-gene array approach, including the well-established buffering interaction between the ABL1 kinase and its antagonist phosphatase, PTPN1 (241,242). Under imatinib treatment, GI scores generally showed larger effect sizes compared to untreated conditions. As noted earlier, single-gene arrays outperformed single gRNAs by producing stronger tau values and GI scores (Fig. 17B, 18A).

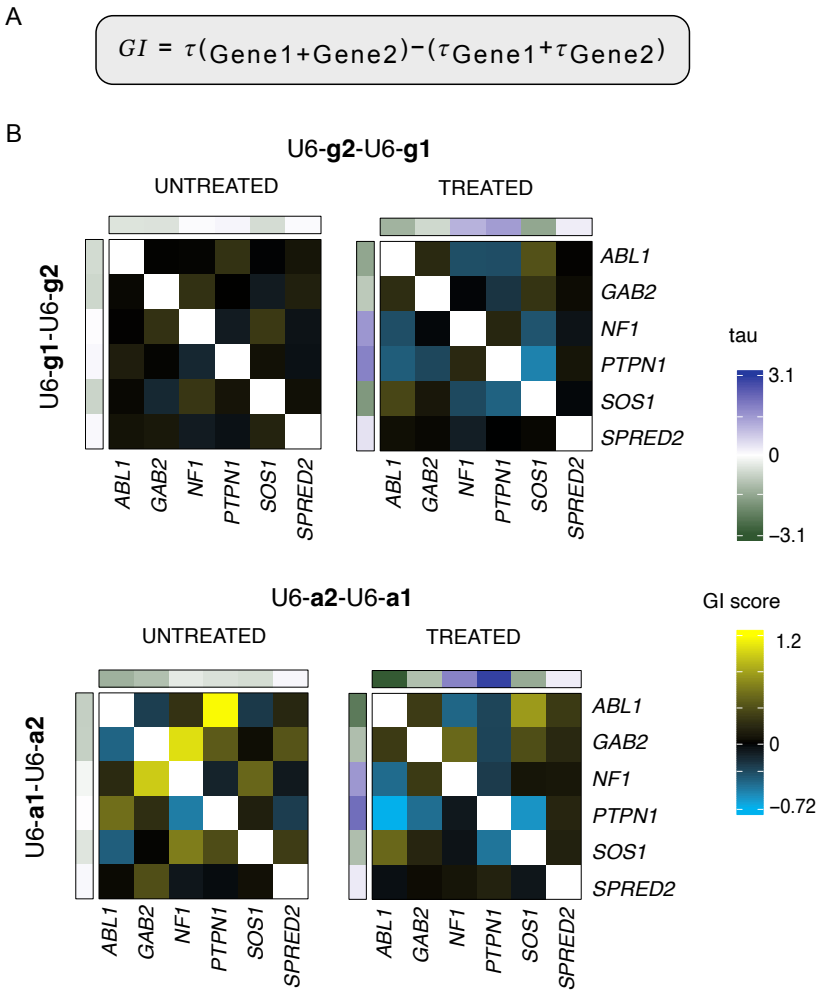


Figure 17 - Single-gene arrays generate reproducible stronger tau and GI scores.

(A) *The formula for calculating GI scores.* ***(B)*** *GI maps showing GI scores from single gRNA (top panel) and single-gene array screens (bottom panel), both with (right panel) and without (left panel) imatinib treatment. Tau values from single gene perturbations in both possible orientations (Gene1-gNTC and gNTC-Gene1) are shown along the edges of each GI map. GI scores derived from gRNA/array orientation 1 are shown at the bottom left of each GI map, while GI scores derived from gRNA/array orientation 2 are shown at the top right. Positive GI scores indicate buffering interactions, while negative GI scores indicate synergistic interactions. Data was analyzed in collaboration with Jasmine Hillmer.*

GI scores showed strong correlations within each gRNA expression method and between the single gRNA and single-gene array approaches, confirming that both reliably identified similar GIs (Fig. 18B, Supp. Fig. 7). A subtle yet consistent negative correlation was observed between GI scores in treated and untreated conditions for both strategies. This indicates that some interactions change their nature in response to imatinib treatment. For instance, the buffering interaction between *ABL1* and *PTPN1* becomes synergistic with imatinib treatment, while the *ABL1-SOS1* interaction shifts from synergistic to buffering. In all four screens, “single gene controls” were calculated as the interaction between the six candidate genes and all non-target control gRNAs. As expected, none of the six investigated genes showed a significant GI with non-target controls (Fig. 18A). Figure 18C shows a *de novo* generated interaction network, consisting of all reproducible GIs calculated from the tau values that were generated by the single-gene array approach in both gene orientations (Gene1-Gene2 and Gene2-Gene1) under untreated and imatinib treated conditions.

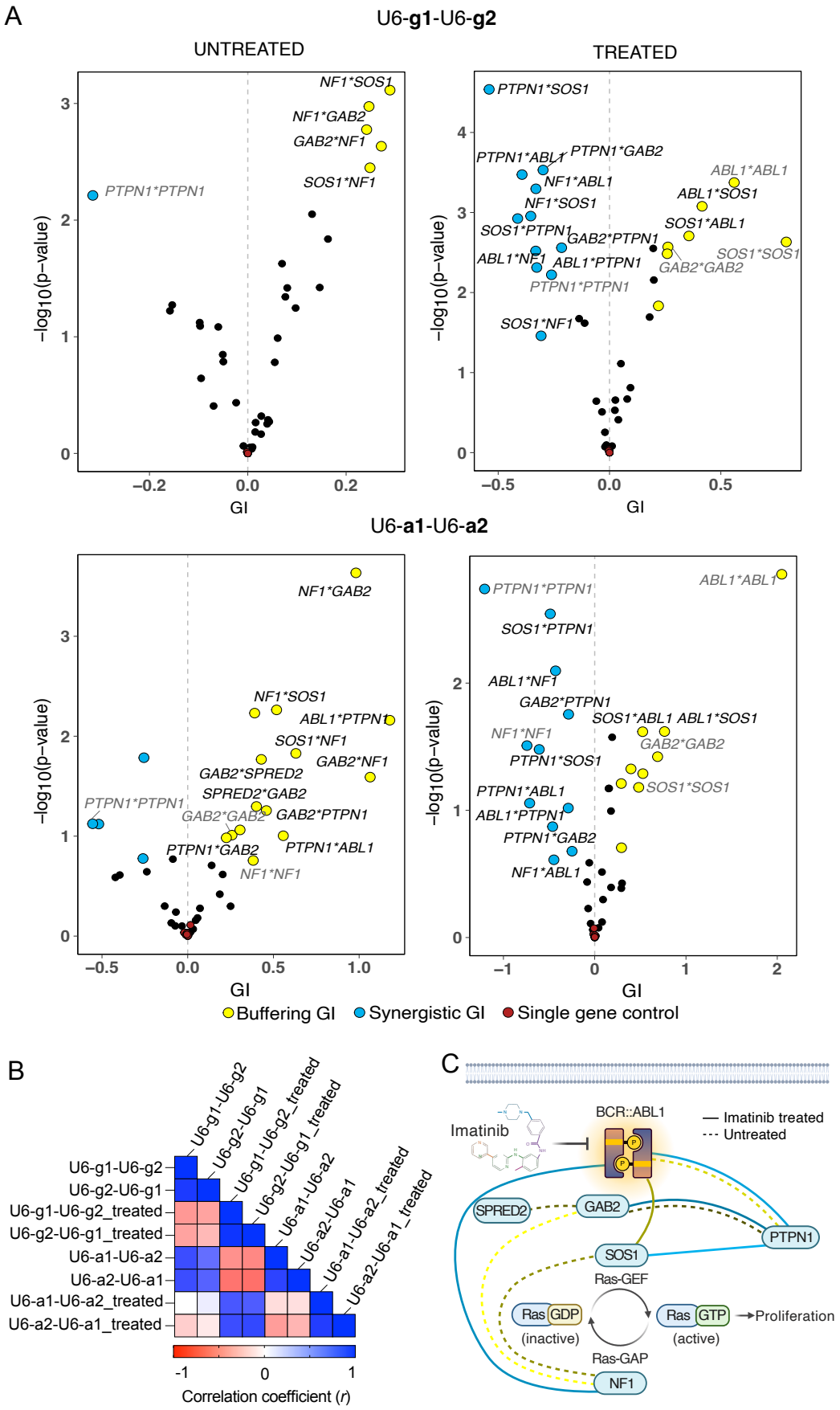


Figure 18 - Cas13d enables the highly reproducible identification of GIs in oncogenic signaling pathways.

(A) *Volcano plots showing GI scores and their associated significance ($-\log_{10}$(p-value)) for all possible gene-gene combinations arising from single gRNA (top) and single-gene array screens (bottom), both with and without imatinib treatment. Blue and yellow data points indicate interactions that meet the threshold of GI >< +/-0.2 and FDR < 0.5. Data points with black labels represent GIs identified in both orientations (gRNA-gNTC and gNTC-gRNA). Gray-labeled data points indicate "same gene GIs" that passed the threshold, where gRNAs in both positions targeted the same gene. Red data points indicate 'single gene controls' where GI scores were calculated as between genes, with one gene replaced by non-target control gRNAs (see methods section 2.7.20 for details).* ***(B)*** *Pearson's correlation (r) between GI scores determined from imatinib-treated and untreated cells via the single gRNA and single-gene array approach, analyzed in both orientations.* ***(C)*** *GI network of reproducible GIs between all six investigated genes, derived from the single-gene array strategy. Only interactions identified by GIs of >< +/-0.2 and FDR < 0.5 in both orientations (gRNA-gNTC and gNTC-gRNA) are shown. Ras-GEF = Ras guanine nucleotide exchange factor. Ras-GAP = Ras-GTPase activating protein. Edges between genes are colored based on the average GI score of screen replicates and gRNA-gNTC orientations. The figure was created using Biorender.com. Data was analyzed in collaboration with Jasmine Hillmer.*

1.4 Discussion

Novel combinatorial screening methods are needed for mapping GIs which is crucial for understanding the functional interdependencies between genes within genetic networks. In the context of cancer, GIs can reveal hidden vulnerabilities, inform combination therapies, and identify synthetic lethal gene pairs. In this chapter, we systematically evaluated of the utility of Cas13d, a type VI-D CRISPR system that targets RNA, for quantifying GIs.

Previous GI mapping approaches relied primarily on the DNA-targeting nucleases Cas9 and Cas12a. Despite advancing GI mapping in mammalian cells, these systems have critical limitations. Both Cas9 and Cas12a can exhibit slow editing kinetics and heterogeneous DNA cleavage and repair dynamics, leading to mosaic populations of edited and unedited cells even several days post-delivery (Fig. 8, 9). It was previously reported that the Cas9 induced non-frameshift indels can lead to over one third of transduced cells with in-frame mutation, or non-perturbed characteristics resulting in true knockouts, heterozygotes, and wild-type cells. This heterogeneity leads to increased noise in functional assays and complicates the analysis of GI data, especially when weak phenotypic effects are being measured (84,115,243).

These limitations are especially problematic when studying essential genes, such as *CD71* (K562 DepMap Chronos score: -0.975). Targeting such genes with DNA nucleases can result in rapid loss of cell viability, which may hide more subtle biological information. In contrast, Cas13d-mediated RNA targeting enables a more controlled and tunable approach which allows the study of partial as well as full loss-of-function phenotypes. In our study, Cas13d achieved 79% reduction of CD71 protein levels within 3 days post-transduction, without signs of CD71 perturbation related cytotoxicity over 10 days, as it was observed after Cas9 as well as Cas12a mediated knockout of CD71 (Fig. 8B). This

observation highlights the advantage of using Cas13d and its ability to modulate gene expression without inducing cell death by providing a more controlled loss of gene expression.

Besides essential genes, Cas13d can also be used to target non-coding RNAs (ncRNAs), such as lncRNAs, which have been shown to be involved in chromatin regulation, splicing, and post-transcriptional gene regulation (244,245). DNA targeting CRISPR systems rely on introducing frameshift mutations to disrupt gene function. Since ncRNAs do not rely on a reading frame, this strategy is ineffective for perturbing non-coding transcripts. Cas13d circumvents this limitation by directly targeting the RNA for degradation, enabling a functional knockdown on the transcript level. Previous studies have demonstrated the effectiveness and specificity of Cas13d in degrading lncRNAs, proving its ability to interrogate the function of both coding and non-coding elements, which allows a systems-level analysis of cellular pathways (150,235).

Another key observation from our study relates to the perturbed cell population dynamics of Cas13d-mediated knockdown compared to traditional DNA-targeting nucleases. We found that in comparison to the RNA-targeting Cas13d, DNA-targeting nucleases generate bimodal populations of full knockout (null mutation) cells and wildtype cells (Fig. 8A). This poses a challenge for studying genes that are relevant for therapeutic intervention because typically, pharmacological inhibitors reduce rather than completely eliminate the activity of their targets. In contrast, Cas13d generated homogenous, uniform knockdown populations that resemble the effects of pharmacological modulation of target activity. The knockdown population uniformity is important in higher-order perturbation screens that are required for GI studies, since the compounding multiplicative effect of the efficiency of single perturbations greatly reduces the double perturbed cell populations. The uniform knockdown mediated by Cas13d led to a higher fraction of double perturbed cell populations in comparison to both DNA-targeting nucleases (Fig. 9B). Overall, this uniformity of perturbed cell populations improves the interpretability of GI data, especially when conducting large-scale combinatorial screens.

Similar logic has been applied for the development of CRISPRi systems, such as those based on dCas9-KRAB, which enable target repression without DNA cleavage (246). However, CRISPRi suffers from strong dependence on chromatin context, positional effects relative to the TSS (247), and can cause substantial off-target effects (248), which reduce its generalizability and scalability. Cas13d avoids these pitfalls by directly targeting mature transcripts, bypassing the complexities of transcriptional regulation and chromatin structure.

In addition to phenotypic consistency, Cas13d offers a major technical advantage by avoiding DNA DSBs, which are the hallmark of Cas9 and Cas12a activity. DNA DSBs can cause cytotoxicity by triggering the p53 response (249,250), and this effect strongly correlates with the number of target loci, raising concerns about the feasibility of using DNA-targeting nucleases in higher-order perturbation screens (251,252). When multiple target loci are edited simultaneously, the cumulative DNA damage can lead to unintended cytotoxicity, confounding the interpretation of observed phenotypes. By targeting RNA rather than DNA, Cas13d overcomes this issue and thus provides a more scalable platform for combinatorial studies.

While Cas13d circumvents many cytotoxicity concerns associated with DSB from DNA-targeting CRISPR systems, several studies have reported varying levels of cytotoxicity in eukaryotic cells caused by Cas13d from *Ruminococcus flavefaciens* (*Rfx*Cas13d) due to collateral RNA cleavage activity. Unlike Cas9 or Cas12a, Cas13 effectors has a nonspecific RNase activity that is allosterically activated upon target RNA binding, leading to the collateral degradation of non-target transcripts. This phenomenon was first seen in bacterial systems and later in select mammalian studies (149,253). Although collateral RNA cleavage was not reported in the original *Rfx*Cas13d publication by Konermann *et al.* 2018 (177), it represents a potential concern for using *Rfx*Cas13d that should be carefully monitored in future studies. Subsequent studies have demonstrated context-dependent cytotoxicity associated with collateral activity based on cell type, target gene expression levels, and the delivery method of the *Rfx*Cas13d system (158–161). However, most studies employing *Rfx*Cas13d did not report any adverse effects from collateral activity (152,165,234,235,254–258). Similarly, in this study, we observed no unexpected growth phenotypes that could be attributed to collateral RNA cleavage (Fig. 16, Supp. Fig. 5), nor did we detect reciprocal degradation between the highly expressed surface proteins CD46 and CD47, as would be expected from collateral RNA trans-cleavage (Fig. 9A). While these results are reassuring, we emphasize that collateral activity remains a relevant concern, especially in large-scale pooled screens or in more sensitive systems such as primary cells or *in vivo* models. To mitigate the effects of collateral activity, newer high-fidelity variants of *Rfx*Cas13d with minimal collateral activity, as well as Cas13d orthologs like *Dj*Cas13d, have demonstrated strong on-target activity with minimal cytotoxicity, even when targeting highly abundant transcripts (160,163).

Although Cas13d offers several inherent advantages, its application to combinatorial GI mapping revealed critical technical challenges that needed resolution to unlock its full potential. Cas13d can process concatenated gRNA arrays from a single transcript through inherent RNase activity, enabling compact and modular expression of multiple gRNAs

from a single promoter. This multiplexing strategy is appealing in combinatorial pooled screening settings, where vector size and cloning complexity impose practical constraints. However, we detected sequence-specific interference between Cas13d gRNAs targeting genes expressed in position 1 with non-target gRNAs expressed in position 2 of the same array (Fig. 11H and 12), making this approach ineffective for quantitative GI mapping. Specifically, when two distinct gRNAs were co-expressed in a tandem array, we observed that the gRNAs in the first position displayed diminished knockdown efficiency depending on the sequence identity of the gNTC in the second position. Notably, this phenomenon occurred in an orientation-dependent and sequence-specific manner (Fig. 11H, 12, 13).

Cas13d gRNA-gRNA sequence-dependent interference effects on efficiency have also been observed in Tieu *et al.* 2024, where Cas13d gRNA knockdown efficiency varied when multiple gRNAs were concatenated into arrays in a sequence-dependent manner. In their study, they used a library of 6400 arrays targeting 576 gene pairs in primary human T cells to identify gene pairs that regulate HA-28ζ CAR T cell proliferation and anti-tumor activity (165). Similar to our Cas13d libraries, their library was designed with gRNA pairs cloned in two orientations. Consistent with our observation, the gRNAs were less efficient when expressed from position 1 compared to position 2 in the array, which can be attributed to gRNA-gRNA sequence-dependent interference (165). Similarly, reduced gRNA efficiency was observed with the dCasRx-RBM25 fusion protein when targeting alternative exons. Li *et al.* 2024 reported that two-spacer arrays showed a moderate decrease in efficiency for exon activation compared to single gRNAs. This effect was further intensified when using five-spacer gRNA arrays (259). Previous studies have attributed this effect to ssRNA secondary structures that render them inaccessible, preventing Cas13d from recognizing the cleavage site 5’ of the DR sequence, thereby severely impacting its array processing efficiency (139,260).

Although concatenating gRNAs against different target genes for GI mapping has been successfully applied with Cas12a (197,208), gRNA-gRNA sequence-dependent interference has been studied more extensively in this CRISPR system due to its more widespread use. These studies can provide insights into the mechanisms behind the Cas13d dual gRNA array efficiency effects observed in this study, despite Cas12a being structurally and functionally different from Cas13d. Several reports have indicated that arrays encoding multiple Cas12a gRNAs often exhibit unpredictable performance for multi-gene perturbation. Small RNA-Seq experiments suggest that array processing by Cas12a is uneven, resulting in gRNAs that may differ in abundance by 10- to 100-fold after processing (209,261,262). Magnusson *et al.* 2021 reported that Cas12a array performance is hypersensitive to the GC content of gRNA spacers, with high-GC gRNAs

impairing the activity of the downstream gRNA (262). Additionally, the presence of complementary sequences in gRNAs that could generate complex secondary RNA structures negatively impacts Cas12a's array processing ability (263,264). Similarly, recent data from CHyMErA (203) screens revealed that polycistronic gRNA processing can substantially impair phenotype induction and reduce the effect size range compared to single-gRNA strategies. Specifically, CHyMErA's hybrid Cas9-Cas12a arrays underperformed both in efficiency and speed of target gene depletion, attributed in part to compromised gRNA activity due to processing interference (265). Further exploration of the underlying reasons behind the Cas13d gRNA sequence-dependent interference and array processing mechanisms may enable the use of concatenated Cas13d gRNAs against different target genes for quantitative GI mapping.

To overcome the Cas13d gRNA-gRNA sequence-dependent interference, we implemented a dual-promoter gRNA expression strategy where each gRNA is expressed under its own U6 promoter in separate transcriptional units. This approach is conceptually analogous to the multiSPAS strategy, which decouples SpCas9 and enAsCas12a gRNA expression to restore optimal gRNA functionality in orthogonal CRISPR systems (265). In our case, this configuration restored efficient knockdown of both targets, enabling accurate quantification of GI phenotypes (Fig. 14) on the expense of slightly increased vector size and library cloning complexity. Nevertheless, for the purpose of high-resolution GI mapping, the benefits of this strategy such as robust knockdown, reproducibility, and minimized gRNA interference, outweigh its drawbacks.

Guide RNA design is essential in all CRISPR applications. Compared to Cas9, that was used in over 1000 whole-genome screens in cancer cell lines, data on Cas13d guide efficacy is lagging behind (266). By analyzing the fitness scores from the gRNA-gNTC combinations in both orientations in our U6-g1-U6-g2 screen, we identified gRNAs that showed absent or minimal fitness effects despite having high predicted efficiency scores in the gRNA selection algorithm from Wessels *et al.* 2020 (234). Therefore, we assessed the efficiency of the gRNA selection algorithm and compared it to the updated version of the algorithm that was published during the course of the project, TIGER (Supp. Table 1) (166,234). The outcome of this comparative analysis revealed that neither model exhibited a reliable prediction of gRNA performance. These prediction models have predominantly been created through the examination of $\log_2$ fold changes derived from essentiality screens. Although essentiality screens offer the advantage of simultaneously testing a large number of gRNAs, they represent a proxy measure for RNA knockdown levels (160,166,234,267). Ideally, model performance should be evaluated by directly comparing predictions to experimental RNA knockdown measurements. This discrepancy highlights

the need for further refinement and optimization of Cas13d gRNA efficiency predictive models.

Next, we explored the potential of concatenating multiple gRNAs targeting the same gene to enhance knockdown efficiency while simultaneously reducing library size. By concatenating three distinct gRNAs against the same target within a single array, we achieved stronger target reduction (Fig. 15B), more pronounced growth phenotypes (Fig. 17B), and reduced the combinatorial library size nine-fold, from 36,864 gRNA-gRNA combinations to just 4,096 array-combinations while enhancing GI effect sizes at the same time (Fig. 18A). The more compact size of the array library allows this approach to be readily adapted to study GIs between hundreds of genes in a single screen. These results show that by leveraging the multiplexing ability of Cas13d, the same-gene arrays can also enable the creation of compact genome-wide libraries by reducing the number of unique library elements while maintaining multiple gRNAs per gene. Such approaches are needed in cases where the cell numbers are limited, such as in primary cells and in scCRISPR-seq approaches. Similar approaches were previously developed for Cas12a and showed that compact Cas12a libraries demonstrated comparable performance to optimized Cas9 libraries (197,212).

One limitation of our dual promoter expression strategy was the recombination of the Cas13d DR sequences in the U6-g1-U6-g2 and the U6-a1-U6-a2 libraries due to the complementarity of the Cas13d DR sequences. Since the libraries were cloned symmetrically with all gRNAs/Arrays present in both positions, recombination during cloning of these libraries was observed when identical gRNA/arrays were present in both positions (Supp. Fig. 3C-D). In these recombined constructs, a single gRNA/array was present under the expression of the mU6 promoter. These events were removed later during analysis, but this underscores the need for further development of efficient alternative Cas13d DR sequences in a similar fashion to the alternative DR sequences developed for Cas12a (197).

Finally, we benchmarked both approaches, the single gRNA (U6-g1-U6-g2) and the single-gene array approach (U6-a1-U6-a2) in a focused screen to study GIs between six genes involved in oncogenic signaling downstream of *BCR::ABL*: ABL1, GAB2, SOS1, NF1, PTPN1, and NF1. The *BCR::ABL* fusion protein is a well-characterized oncogenic driver in CML, where it promotes constitutive tyrosine kinase activity and activates downstream pathways such as Ras-MAPK and PI3K-AKT (40,41). Our results show that the calculated GI scores were highly reproducible, both within and across the two approaches (Fig. 18B). Interestingly, 3 out of the 4 synergistic GIs that could be detected under imatinib treatment were between *PTPN1* and *ABL1*, *SOS1* or *GAB2*, respectively (Fig. 18C), suggesting a

negative upstream regulatory function of PTPN1 on these genes. These findings are consistent with the known role of PTPN1 in dephosphorylating and attenuating ABL1 activity, which has been documented in several prior studies (241,242). While there have been no reports on the direct interaction between PTPN1 and the adaptor protein GAB2 or the Ras-GEF SOS1 yet, the observed synergistic interactions could be explained by the well-established function of both proteins as direct physical interaction partners of BCR::ABL, the target of PTPN1 (268). In addition, we also identified a strong synergistic interaction between *ABL1* and *NF1*, which can be rationalized by their opposing roles in regulating Ras activity. ABL1 positively regulates Ras through phosphorylation of Ras-GEFs and activation of adaptor complexes, while NF1 is a GAP protein that promotes Ras inactivation (44–46,269). We also observed a buffering interaction between *ABL1* and *SOS1* under imatinib treatment, wherein the combined knockdown rendered cells less sensitive to the drug than expected based on individual perturbations, possibly due to their involvement in the same signaling complex (38). This counterintuitive result can be interpreted as a case of negative epistasis within the same pathway, where both genes participate in a shared signaling module. Taken together, these results demonstrate that Cas13d-mediated GI mapping is not only technically robust but also biologically informative, capable of recapitulating known interactions and uncovering new ones.

1.5 Conclusion

This study establishes Cas13d, an RNA targeting CRISPR system, as an efficient tool for quantitative GI mapping in mammalian cells. By comparing Cas13d to other traditional DNA-targeting CRISPR systems, such as Cas9 and Cas12a, we show that Cas13d allows rapid RNA knockdown and results in more uniform perturbation across cell populations. In addition, Cas13d acts post-transcriptionally which enables its use in contexts where DNA DSB-based perturbation is undesirable or infeasible, such as studies involving essential genes where full gene knockouts are not tolerated, as well as ncRNAs where the function is independent of the presence of reading frames. These findings position Cas13d as a complementary tool to other DNA-targeting CRISPR platforms for gene function and GI studies. Our dual-promoter strategy for expressing multi-gene-targeting gRNAs overcomes gRNA-gRNA sequence-dependent interference observed with concatenated gRNAs expressed under a single promoter as multi-gRNA arrays. Furthermore, we demonstrate that concatenating gRNAs targeting the same gene into same-gene arrays increases the knockdown phenotype effect sizes and allows the detection of GIs within oncogenic signaling pathways. In conclusion, we show that Cas13d has a strong potential for GI mapping to advance our understanding of drug response pathways.

Chapter 2: Perturb seq reveals TCF7 as a nexus of interaction between MAPK- and Wnt-driven gene expression

2.1 Abstract

The MAPK pathway is a crucial cellular signaling cascade whose dysregulation is implicated in numerous cancers. While its upstream regulators are well characterized, how its activation translates into different transcriptional responses remains poorly understood. This study addresses this gap by using targeted Perturb-seq against 22 transcription factors in an inducible model system for RAF-MAPK signaling. Leveraging the power of single-cell CRISPR screening, which enables high-resolution mapping of complex phenotypes through single-cell transcriptomic profiling, we capture the diverse transcriptional programs induced by gene perturbations. A topology-based modeling approach is applied to the obtained data to construct a transcription factor directional interaction network which identified a positive feedback loop between EGR1, a downstream regulator of the RAF-MAPK response, and TCF7, a transcription factor best known for its involvement in Wnt-signaling. This work demonstrates how the combination of single-cell CRISPR screening and gene network analysis can provide novel insights into gene cross-talk and the transcriptional architecture downstream of RAF-MAPK activation.

2.2 Introduction

Cell surface signals traverse a complex network of signaling proteins that lead to the activation of intricate gene expression programs. Investigating the relationship between these signaling proteins and the gene expression programs they activate is challenging due to the potential activation of different signaling pathways by a single receptor and the ability of different receptors to activate the same pathway. Moreover, the transcriptional response to a signal is not fixed but rather is shaped by the cellular context, including the presence of other signaling inputs, feedback regulators, and epigenetic states (270,271). Activated pathways can intersect, antagonize each other, or converge on shared downstream effectors through the presence of feedback loops, which add further complexity (272).

One of the hallmarks of cancer formation is mutations of tumor activator and tumor suppressor genes that have a critical role in intracellular signal transmission. For example, G12D activating KRAS mutations and APC loss-of-function mutations lead to

dysregulation of the MAPK (273) and Wnt signaling pathways (274). Both these mutations are seen in 21% and 14% of all cancer cases, respectively, and can occur in isolation as well as in combination in a large number of cancers (275–277). Interactions between the MAPK and Wnt pathways have been demonstrated at multiple levels, including cross-talk at the receptor activation level, interactions at the cytosolic kinase level, and transcriptional co-regulation of target genes (278). One established interaction is MAPK-dependent β-catenin phosphorylation, a main Wnt signaling effector, which may influence its stability and nuclear localization. Similarly, EGR1, a MAPK-inducible immediate-early gene, is involved in β-catenin/TCF-dependent transcriptional regulation (279). Dissecting the mechanisms of such cross-talk and studying how the extracellular signals are converted into gene expression programs is still one of the challenges of systems biology.

The MAPK cascade is a conserved and central regulator of essential cellular functions such as proliferation, differentiation, and survival (15,271). It is activated by various extracellular cues such as growth factors, cytokines, and environmental stress, and it relays signals from cell surface receptors to intracellular mediators, and finally to the nucleus where targeted transcriptional programs are activated. MAPK signaling has been extensively studied and therefore it serves as a model to investigate signal-induced transcriptional regulation (280). While individual MAPK isoforms are structurally conserved, they have evolved functionally and acquired distinctive activation profiles and substrate specificities that ensure cellular homeostasis as well as unique responses to various upstream stimuli (15,16).

The MAPK pathway is divided into the canonical and the non-canonical pathways. Signal transmission via the MAPK pathway involves a tiered phosphorylation cascade in which MAPKKKs phosphorylate MAPKKs, which in turn activate MAPKs through dual phosphorylation on conserved motifs. In this study, we focus on the RAF-MEK-ERK pathway which is part of the canonical MAPK pathway. After Ras activation by growth factor-induced and mitogen-induced signals at the cell plasma membrane, GTP-loaded Ras directly interacts with RAF kinases, inducing catalytic activity via RAF kinase domain dimerization (25,281). Serving as a MAPKKK, RAF subsequently initiates the activation of the dual-specificity kinase MEK (MAPKK) by phosphorylating its activation segment. This signal is then propagated by MEK, which phosphorylates and activates ERK (27). Once activated, ERK translocates to the nucleus where it phosphorylates a diverse set of transcription factors (TFs) triggering specific transcriptional programs (28).

Time-resolved transcriptome analysis following ERK activation has provided insights into the temporal dynamics and hierarchical organization of ERK-target genes. The studies have shown that cell outputs are subject to the duration of ERK signals, with transient

activation prompting cell proliferation, and sustained signaling leading to differentiation or apoptosis (282,283). The duration of ERK activity itself is also regulated by the types of ligand, receptor abundance, and ERK-dependent negative feedback loops (284).

In addition, transcript half-life also influences the kinetics of this initial cellular response (285). ERK signaling is generally associated with fast and transient induction of primary response genes (PRGs) or immediate early genes (IEGs), which are independent of *de novo* protein synthesis and share regulatory motifs within their promoters (286,287). Many IEGs encode TFs that subsequently initiate activation of secondary response genes (SRGs) after a temporal delay (288–290). This delay, modulated by mRNA half-lives, enables the cell to differentiate between transient and sustained signals effectively (291,292). Delayed primary genes frequently encode negative feedback regulators, attenuating growth factor signaling at specific nodes within the network (293–295).

Despite extensive characterization of ERK-induced transcriptomic responses and associated regulatory proteins (296), significant gaps remain in understanding the detailed structure of the transcriptional network. Specifically, the regulatory interplay between the TFs and how they regulate secondary response genes remains poorly studied. A better understanding of the transcriptional network topology and dynamics is crucial to decoding how MAPK pathways orchestrate distinct and often opposing phenotypes. Reverse engineering approaches represent a powerful strategy for dissecting these complex transcriptional networks downstream of MAPK signaling to understand how physiological and pathological MAPK signals are integrated into a cellular response. Early studies using RNAi based perturbations and semi-quantitative modeling identified a small seven node TF network downstream of Ras/MAPK signaling that modulates cell growth and transformation (290). Advances in genome editing technologies and coupling of CRISPR-based genetic perturbations with single-cell RNA sequencing (scCRISPR-seq), such as in Perturb-seq (214,215) and CROP-seq (223), have overcome limitations that restricted broader and more profound analyses (79). Pooled genetic screens rely on selectable phenotypes like cell growth, drug resistance, or reporter gene expression with fluorescence-activated cell sorting that are intrinsically limited in their ability to capture complex biological processes, particularly those involving transcriptional network topologies (175). scCRISPR-seq integrates single-cell transcriptomics with targeted genetic perturbations at scale to enable the systematic and simultaneous dissection of complex transcriptional interactions. These technologies now offer unprecedented capacity to map network topology, decode signaling cascades, and achieve deeper mechanistic insight into how physiological and pathological MAPK signaling are coordinated into cell response.

In this study, we used a targeted Perturb-seq (TAP-seq) strategy (228) to evaluate the transcriptional changes resulting from the knockout of 22 TFs activated downstream of RAF1-mediated MAPK signaling. Using this data, we constructed a model of the genetic interaction network and demonstrated the complex transcriptional interdependence that describes how these transcriptional elements regulate each other's expression. In this network, EGR1 emerged as the most upstream regulator of the transcriptional RAF-MAPK response, and together with FOS, it frequently co-regulates overlapping target gene sets in an orthogonal manner. Furthermore, we also identified a positive feedback loop between EGR1 and TCF7, thereby establishing a nexus that links the MAPK and Wnt signaling pathways. Our results demonstrate the effectiveness of Perturb-seq in reconstructing topological networks and identifying dynamic interactions within cellular transcriptional networks.

The work presented in this chapter was conducted in collaboration with the Blüthgen Lab at Charité - Universitätsmedizin Berlin, who contributed to the experimental design, preliminary experiments, and computational analysis.

2.3 Results

2.3.1 Identification of transcription factors up-regulated by RAF1-induction

To characterize the transcriptional networks activated by RAF-MAPK signaling, a HEK293ΔRAF1:ER cell line engineered with a tamoxifen-inducible CR3 kinase domain of RAF1 was used (Fig. 19). This cell line, previously established by Samuels *et al.* 1993 (297), allows precise control of RAF1 activity. Unlike cell culture systems stimulated by growth factors, this model of RAF-MAPK signaling is activated independently of the upstream G-protein receptor activation of Ras, thereby reducing the potential for pathway divergence and the initiation of feedback mechanisms.

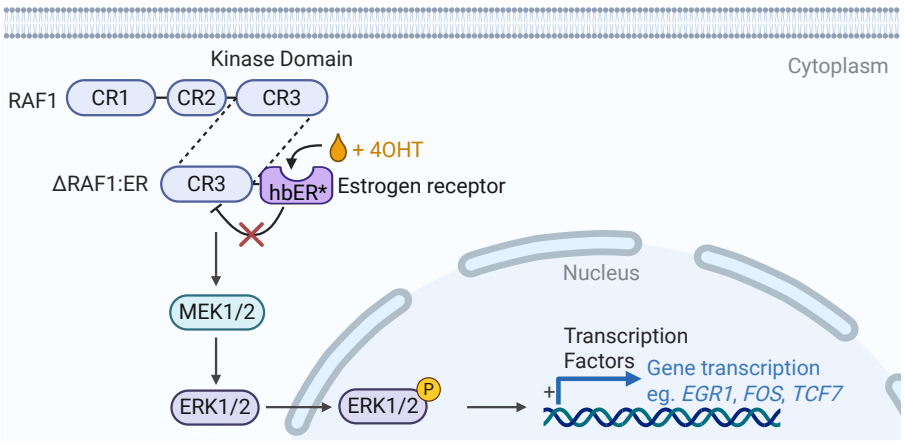


Figure 19 - Schematic structure of the 4OHT-inducible RAF1-CR3 kinase domain in HEK293ΔRAF1:ER cells.

The RAF1 CR3 domain is fused to an estrogen receptor hormone-binding domain (hbER) to create the ΔRAF1:ER fusion protein. Adding 4-hydroxytamoxifen (4OHT) relieves the inhibitory activity of the hbER* domain, allowing the ΔRAF1:ER to phosphorylate and activate MEK1/2. Activated MEK1/2 subsequently phosphorylates ERK1/2, resulting in ERK1/2 activation. Phosphorylated ERK1/2 translocates into the nucleus, promoting the activation of TFs and the transcription of MAPK pathway target genes. The figure was created using Biorender.com.*

RAF1 activation was induced by treatment with 4-hydroxytamoxifen (4OHT) for periods ranging from 0.5 h to 8 h, and the resulting changes in the cellular transcriptome were analyzed using bulk RNA sequencing (RNA-Seq). Out of 1,142 significantly upregulated genes, we selected 22 TFs that were induced at different time points following RAF1 induction for further analysis. The basal expression levels of the selected TFs varied over several orders of magnitude, and the degree of their upregulation after RAF1 activation was independent of their basal expression levels (Fig. 20A).

Pulse activation of RAF1 signaling enabled the classification of the selected TFs into four distinct response classes previously described by Uhlitz *et al.* 2017 (285) (Fig. 20B, Table 1). The first group consists of primary response genes, including immediate-early genes (IEG) and immediate-late genes (ILG), such as members of the *EGR* gene family, *FOS*, *FOSB*, and *JUNB*. These genes are rapidly induced upon signal activation and display short mRNA half-lives, resulting in a quick decline in expression following the removal of 4OHT. The second group belongs to the immediate-late genes (ILGs), which include *FOSL1* and *FOSL2.* These genes are also rapidly induced within the first 1-2 h of RAF1 induction but show longer mRNA stability and remain elevated even after the pulse has concluded. The third group consists of secondary response genes (SRG), representing delayed-response TFs that require longer induction periods for activation because they depend on the translation of IEGs and DEGs. For example, TCF7 expression is detectable only after 4 h and reaches maximal levels after 8 h of stimulation (285). The distinct induction kinetics for each of the 22 selected candidate TFs are shown in Fig. 20B. Using this data, we calculated the half-maximal induction times of the TFs, which ranged between 17 min and 5 h (Table 1). These induction kinetics indicate the temporal role of each TF within the RAF-MAPK response. However, they do not by themselves clarify whether or how these TFs contribute functionally to downstream transcriptional regulation of the RAF-MAPK pathway. To address this, we proceeded with a series of Perturb-seq loss-of-function screens.

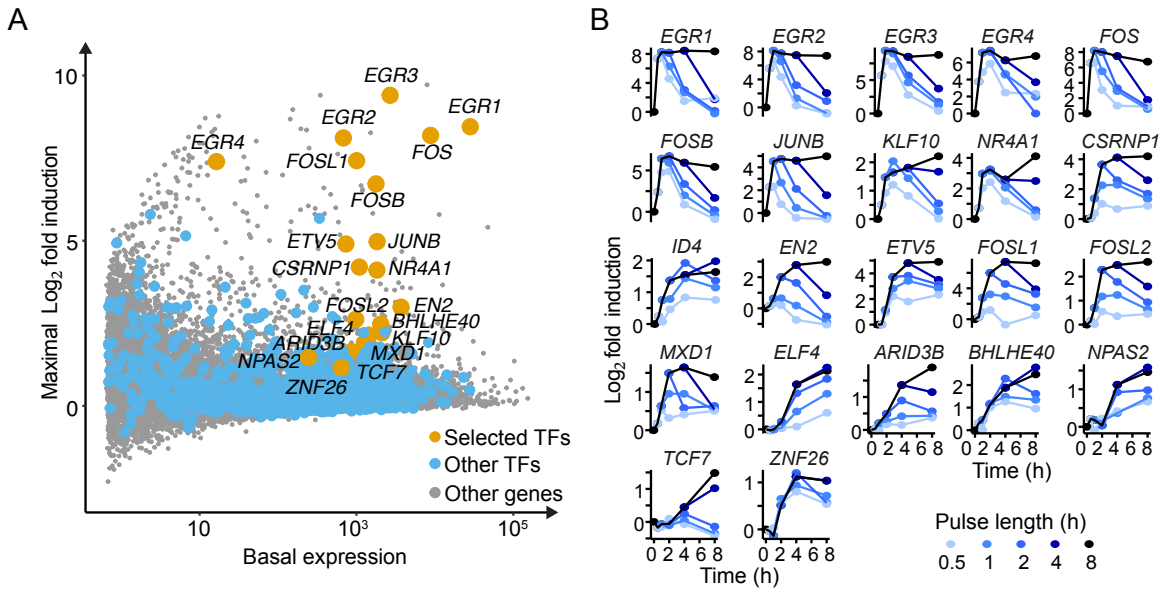


Figure 20 - Transcriptional profiling of HEK293ΔRAF1:ER cells following RAF1 induction.

(A) *The maximum log₂ fold change in expression of the selected candidate TFs is plotted against their expression levels in non-induced HEK293ΔRAF1:ER cells from bulk RNA-Seq analysis.* ***(B)*** *The time-resolved log₂ fold expression changes of candidate TFs are shown after the indicated pulse lengths of 4OHT-mediated RAF1 induction. Data was generated by Florian Uhlitz and analyzed by David Steinbrecht.*

Table 1 - Selected candidate genes and their classification based on their response time to RAF1 induction (285).

IEG = Immediate-early genes, ILG = Immediate-late genes, DEG = Delayed-early genes, SRG = Secondary response genes.

Gene symbol	Name	Class (Uhlitz *et al.* 2017)	Half-maximal induction (h:min)
EGR1	Early Growth Response 1	IEG	0:17
FOS	Fos Proto-Oncogene, AP-1 Transcription Factor Subunit	IEG	0:19
EGR2	Early Growth Response 2	IEG	0:22
EGR3	Early Growth Response 3	IEG	0:25
EGR4	Early Growth Response 4	ILG	0:29
JUNB	JunB Proto-Oncogene	IEG	0:33
FOSB	FosB Proto-Oncogene	IEG	0:37
NR4A1	Nuclear Receptor Subfamily 4 Group A Member 1	ILG	0:47
KLF10	KLF Transcription Factor 10	DEG	0:49
ID4	Inhibitor Of DNA Binding 4	DEG	1:05
MXD1	MAX Dimerization Protein 1	DEG	1:13
FOSL1	FOS Like 1	ILG	1:17
CSRNP1	Cysteine And Serine Rich Nuclear Protein 1	SRG	1:19
FOSL2	FOS Like 2	ILG	1:23
ETV5	ETS Variant Transcription Factor 5	IEG	1:31
EN2	Engrailed Homeobox 2	SRG	1:37
ZNF26	Zinc Finger Protein 26	DEG	2:09
BHLHE40	Basic Helix-Loop-Helix Family Member E40	DEG	2:15
ARID3B	AT-Rich Interaction Domain 3B	SRG	3:03
ELF4	E74 Like ETS Transcription Factor 4	SRG	3:10
NPAS2	Neuronal PAS Domain Protein 2	SRG	3:16
TCF7	Transcription Factor 7	SRG	5:08

2.3.2 *Perturb-seq screens for the functional characterization of RAF1 upregulated TFs*

To investigate the transcriptional targets of the 22 selected RAF-induced TFs and identify key regulators of the RAF-MEK-ERK signaling network, we performed pooled CRISPR Cas9 screens with scRNA-seq readout via direct capture Perturb-seq (225), alongside pooled CRISPR screens with proliferation read-out. A pooled CRISPR sgRNA library was designed and cloned targeting each of the 22 candidate TFs with 4 sgRNAs per gene. Additionally, the library included 10 non-target control sgRNAs (sgNTC), 10 safe-cutter sgRNAs that cleave gene-free regions to control for Cas9-induced DNA DSB effects, and 4 sgRNAs targeting the RAF1-CR3 kinase domain as a positive control. After lentiviral packaging of the sgRNA library, the HEK293ΔRAF1:ER cells were transduced at low multiplicity of infection (MOI = 0.2) and cultured for 10 days to achieve efficient knockout of the targeted TFs (Supp. Fig. 8, 9). Following the editing period, the ΔRAF1:ER transgene was activated by treatment with 4OHT. Cells were treated for 6 h, 12 h, and 18 h to capture the temporal dynamics of transcriptional responses within the RAF-MEK-ERK pathway. Single-cell transcriptomic profiles were then generated using the 10x Genomics 3' scRNA-seq pipeline with "Feature Barcode technology" to simultaneously detect the transcriptome and the sgRNA expressed in each single cell. All experiments were conducted in duplicate to enhance the robustness of the findings.

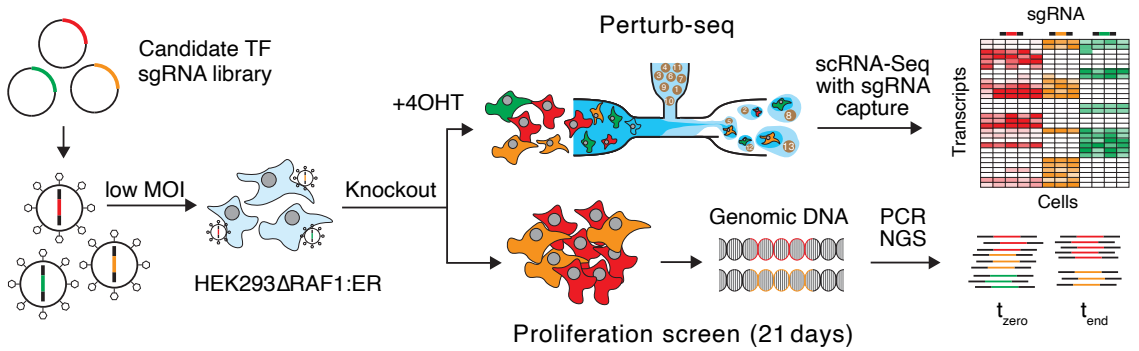


Figure 21 - Schematic of Perturb-seq screens (top panel) and proliferation CRISPR screens (bottom panel).

A pooled sgRNA library targeting candidate TFs is packaged into viral particles and used to infect HEK293ΔRAF1:ER cells at low MOI. Following induction with 4OHT, cells are divided into two branches: ***Top panel:*** *scCRISPR-seq screens to identify the transcriptional targets of the 22 candidate TFs at 6 h, 12 h, and 18 h post RAF-induction. For this purpose, perturbed cell pools were run through the 10x Genomics scRNA-seq pipeline with simultaneous sgRNA capture. Consequently, the identity of the sgRNA expressed in a given cell could be determined along with the cell's transcriptome.* ***Bottom panel:*** *Negative selection screens were performed to identify proliferation phenotypes caused by the knockout of the 22 candidate TFs.*

2.3.3 Quality control of Perturb-seq screen data

We confirmed the high quality of our data using various metrics. The cloned sgRNA library showed a narrow distribution, with 95.5% of the normalized sgRNA sequence read counts falling within one order of magnitude (Fig. 22A). This narrow distribution is crucial for Perturb-seq screens, ensuring comparable numbers of analyzable cells per sgRNA throughout the experiment. The low MOI during transduction resulted in predominantly single sgRNA integrations per cell. Cells exhibiting either no sgRNA or more than one sgRNA were detected, likely due to multiple cells being captured together with a single 10x bead in the same emulsion droplet as a result of intentionally overloading the 10x Genomics Chips with cells. These cells were excluded from further analysis, resulting in approximately 10,000 analyzable cells per sample with exactly one sgRNA (Fig. 22B). Subsequent sequencing of the CRISPR libraries after the scCRISPR-seq experiments yielded high saturation levels, with median unique molecular identifier (UMI) counts per sgRNA ranging from 26 to 84, depending on the sample and time point (Supp. Fig. 10A). Due to the narrow distribution of the sgRNA library, the number of analyzable cells per sgRNA was also very uniform across the six Perturb-seq CRISPR library samples, with median values ranging from 75 to 89 analyzable cells per sgRNA. sgRNAs with fewer than 30 cells per sgRNA were excluded from further analysis (Fig. 22C). These results demonstrate the high technical quality and depth of coverage achieved in our experimental design, forming a solid foundation for all subsequent analyses.

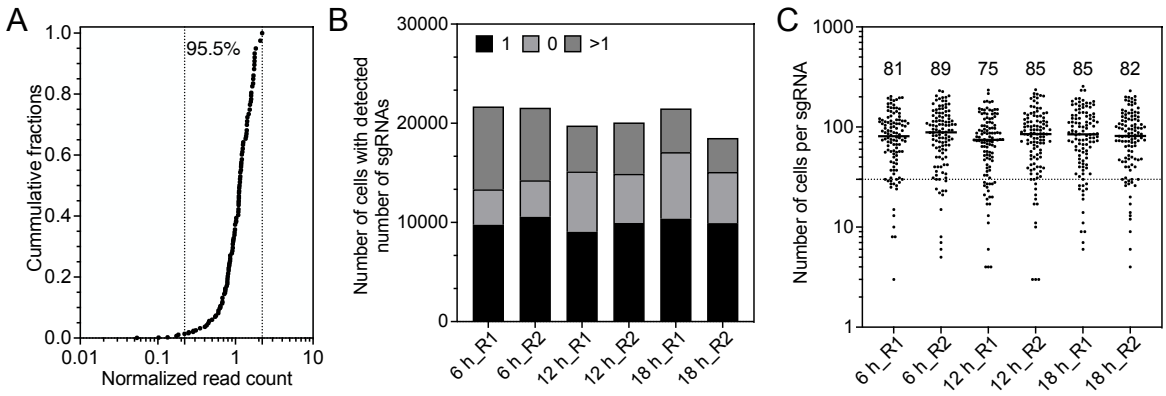


Figure 22 - Perturb-seq screens QC in HEK293ΔRAF1:ER cells.

***(A)** Distribution of the pooled sgRNA library used for all Perturb-seq and proliferation screens. **(B)** Total number of recovered cells and number of cells with 0, 1, and >1 sgRNAs detected in the respective Perturb-seq samples. **(C)** Median number and distribution of cells per sgRNA detected in the respective Perturb-seq samples. Dashed line at 30 cells per sgRNA indicates the cut-off cell number per sgRNA used for further analysis. sgRNA library was designed by Michael Böttcher. Perturb-seq screens data was analyzed by Nils Blüthgen.*

After integrating the scCRISPR screen data at 6 h, 12 h, and 18 h post RAF1 induction and visualizing it using UMAP projections, we observed robust clustering of biological replicates at each induction time point, highlighting the high reproducibility of our dataset. Cells from the 12 h and 18 h induction time points formed distinct clusters, separate from those of the 6 h samples, suggesting that major transcriptional changes occur between 6 h and 12 h following RAF1 induction (Fig. 23A). This observation emphasizes the temporal dynamics of the cellular response. To evaluate potential confounding effects of the cell cycle on our data, we examined the distribution of cells across different cell cycle phases within the UMAPs. We found a uniform representation of G1, S, and G2M phase cells within each cluster (Fig. 23B), indicating that cell cycle heterogeneity did not contribute to the observed transcriptional differences. This is particularly noteworthy as the cell cycle is a common source of biological noise.

To assess the transcriptional consequences of RAF1 induction, we compared *RAF1* transgene knockout cells with safe-cutter control cells. The *RAF1* knockout cells formed a distinct cluster, separated from the safe-cutter controls (Fig. 23C). This separation confirmed the expected absence of transcriptional response to RAF1 activation in the knockout cells and validated that the transcriptional changes observed in the safe-cutter controls were specific to RAF1 induction by 4OHT. Furthermore, pseudo-bulk analysis of the differential expression of known RAF-MAPK target genes revealed a normally distributed peak around zero for the safe-cutter controls. Simultaneously, *RAF1*-knockout cells showed a -0.5 $\log_2$ fold change shift towards negative (Fig. 23D). The absence of the shift in the safe-cutter controls compared to sgNTC indicates a lack of global transcriptional changes in response to Cas9-induced DSB in our scCRISPR-seq results. Additionally, the shift in the *RAF1* knockout cells indicates the inability of these cells to activate RAF-MAPK pathway genes in response to 4OHT treatment, reinforcing the functional impact of the knockout. Examination of individual sgRNAs targeting *RAF1* confirmed the effectiveness of three of the four guides (the fourth *RAF1* sgRNA was represented in less than 30 cells and hence was omitted from analysis), each showing between 12% and 37% of significantly deregulated genes among the RAF1 induced targets identified in bulk RNA-Seq (Supp. Fig. 10B). Overall, these findings demonstrate that the Perturb-seq screens performed as intended, with both positive and negative controls behaving as expected, thereby establishing a solid foundation for downstream analyses.

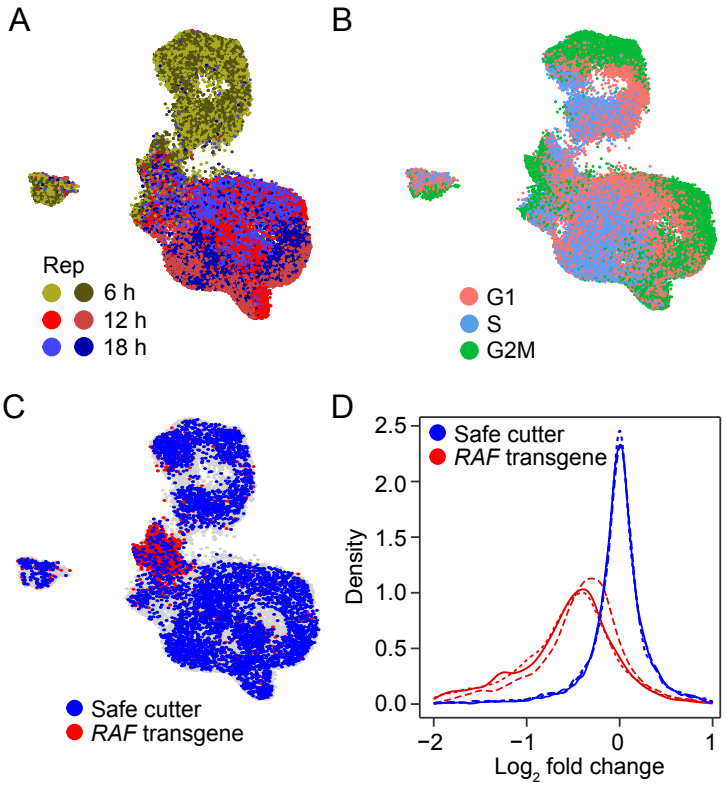


Figure 23 - UMAP clustering and RAF knockout validation confirm high-quality Perturb-seq data.

***(A)** UMAP integration of Perturb-seq samples from three different time points with two replicates each. **(B)** Distribution of cell cycle phases G1, S, and G2M on the integrated UMAPs. **(C)** Distribution of the safe cutter sgRNA cells and the RAF1-knockout cells on the UMAP clusters. **(D)** Histograms of the log_2 fold change of differentially expressed genes between RAF1-knockout (red) and the safe cutter control cells (blue), relative to the non-target control cells. Dashed line: 6 h time point, Dotted line: 12 h time point, Solid Line: 18 h time point. Perturb-seq screens data was analyzed by Nils Blüthgen.*

2.3.4 Proliferation CRISPR screens of candidate TFs in HEK293ΔRAF1:ER cells

To investigate whether the knockout of candidate TFs has strong effects on cell growth, we performed pooled CRISPR screens with a proliferation readout, both with and without RAF1 induction (Fig. 21). The proliferation screens revealed that most TFs show no significant depletion or enrichment compared to sgNTC controls. Perturbation of *FOS* and *EGR2* had a mild yet significant negative impact on the proliferation of both RAF1-induced and uninduced cells (Fig. 24), confirming their established roles in cell proliferation. In contrast, the knockout of the *RAF1* transgene itself (positive control) resulted in strong enrichment of cells upon RAF1 induction but not under uninduced conditions (Fig. 24). This finding is consistent with the observation that prolonged RAF-MAPK activation in HEK293ΔRAF1:ER cells leads to Caspase-8-mediated apoptosis induction (282) and suggests that cells in which the *RAF1* transgene was successfully knocked out gain a proliferative advantage under RAF1-activating conditions. Most importantly, the absence

of strong depletion of sgRNAs targeting candidate TFs also ensured that Perturb-seq experiments could be conducted without losing sgRNA library representation.

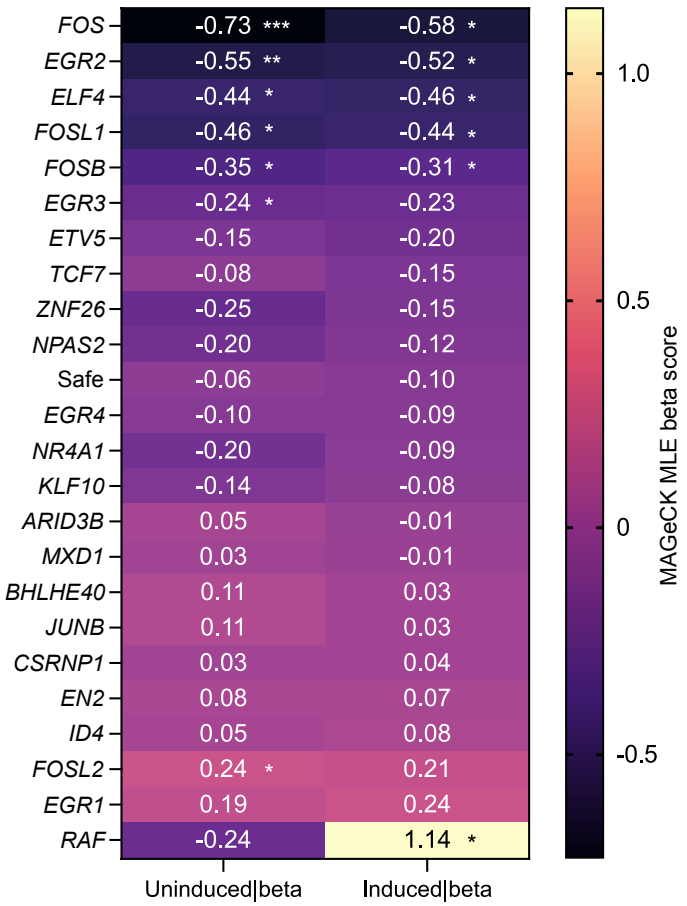


Figure 24 - CRISPR-Cas9 proliferation screen in HEK293ΔRAF1:ER cells.

*MAGeCK MLE beta scores and corresponding significance are shown. * Wald FDR <0.05, ** Wald FDR <0.01, *** Wald FDR <0.001.*

2.3.5 Number of target genes varies greatly between candidate TFs

Perturb-seq experiments are constrained by high costs, inadequate detection of lowly expressed genes, and complex data analysis. Targeted Perturb-seq (TAP-seq) addresses these challenges by sequencing PCR-amplified specific genes instead of the entire transcriptome, enhancing scalability, cost-effectiveness, and sensitivity (228). We employed a modified TAP-seq method using PCR-amplified cDNA as a template. With this approach, we successfully amplified 140 genes selected based on their activation and expression following RAF1 induction, after which we combined the targeted and untargeted Perturb-seq sequencing data to optimally analyze the 140 selected transcripts at high sequencing depth while simultaneously covering the entire transcriptome from the standard gene expression library (Supp. Data 5).

Examining the number of deregulated genes in Perturb-seq screen results at 6 h, 12 h, and 18 h after RAF1 induction, we found that the vast majority of the 10 safe-cutter negative control sgRNAs produced no significantly deregulated target genes across all three time points when compared to sgNTC controls (Fig. 25A). These results aligned with the absence of a shift for the safe-cutter controls in the pseudo-bulk analysis of differential expression in Figure 23C-D. In contrast, sgRNAs targeting the positive control *RAF1* consistently yielded between 500 and 1,000 significantly deregulated target genes at the three post-induction time points. For the 22 candidate TFs, the degree of transcriptional deregulation following their perturbation varied greatly among the individual TFs. For instance, while perturbation of EGR1 resulted in the deregulation of hundreds of target genes, the perturbation of more than half of the candidate TFs did not lead to significant deregulation of target genes above the threshold defined by the safe-cutter controls. These results underscore the differential impact of individual TFs on the RAF1-driven transcriptional program.

Figure 25B shows a heatmap of all candidate TFs whose perturbation resulted in significant deregulation of target genes, including the *RAF1* positive control. As expected, the perturbation of *RAF1* resulted in a strong downregulation of known RAF1 target genes following 4OHT treatment compared to 4OHT-induced sgNTC cells, indicating the knockout cells' inability to activate RAF1 signaling. The data furthermore reveal that a substantial fraction of RAF1 target genes were already deregulated after 6 h of induction in *RAF1* perturbed cells, while additional genes showed significant deregulation only at 12 h, consistent with the temporal induction kinetics described in Figure 20. No substantial differences were observed between the 12 h and 18 h time points, suggesting that the complete transcriptional response to RAF1 activation occurs by 12 h (280).

Among the 22 candidate TFs analyzed, EGR1 and FOS emerged as the primary regulators of the RAF-MAPK transcriptional program, showing the largest number of significantly deregulated target genes (Fig. 25A). Unlike *RAF1*, the perturbation of *EGR1* and *FOS* influenced both the upregulation and downregulation of distinct gene modules, suggesting that RAF1 activation both inhibits and induces these subsets of EGR1 and FOS target genes. These differentially expressed target gene modules only partially overlap with differentially expressed genes following *RAF1* knockout. Interestingly, these two TFs appeared to operate in an orthogonal and sometimes opposing manner on overlapping sets of target genes, where one factor induced the expression of a gene module while the other repressed it (Fig. 25B). This observation indicates that EGR1 and FOS govern divergent and potentially compensatory transcriptional programs in response to RAF1 activation.

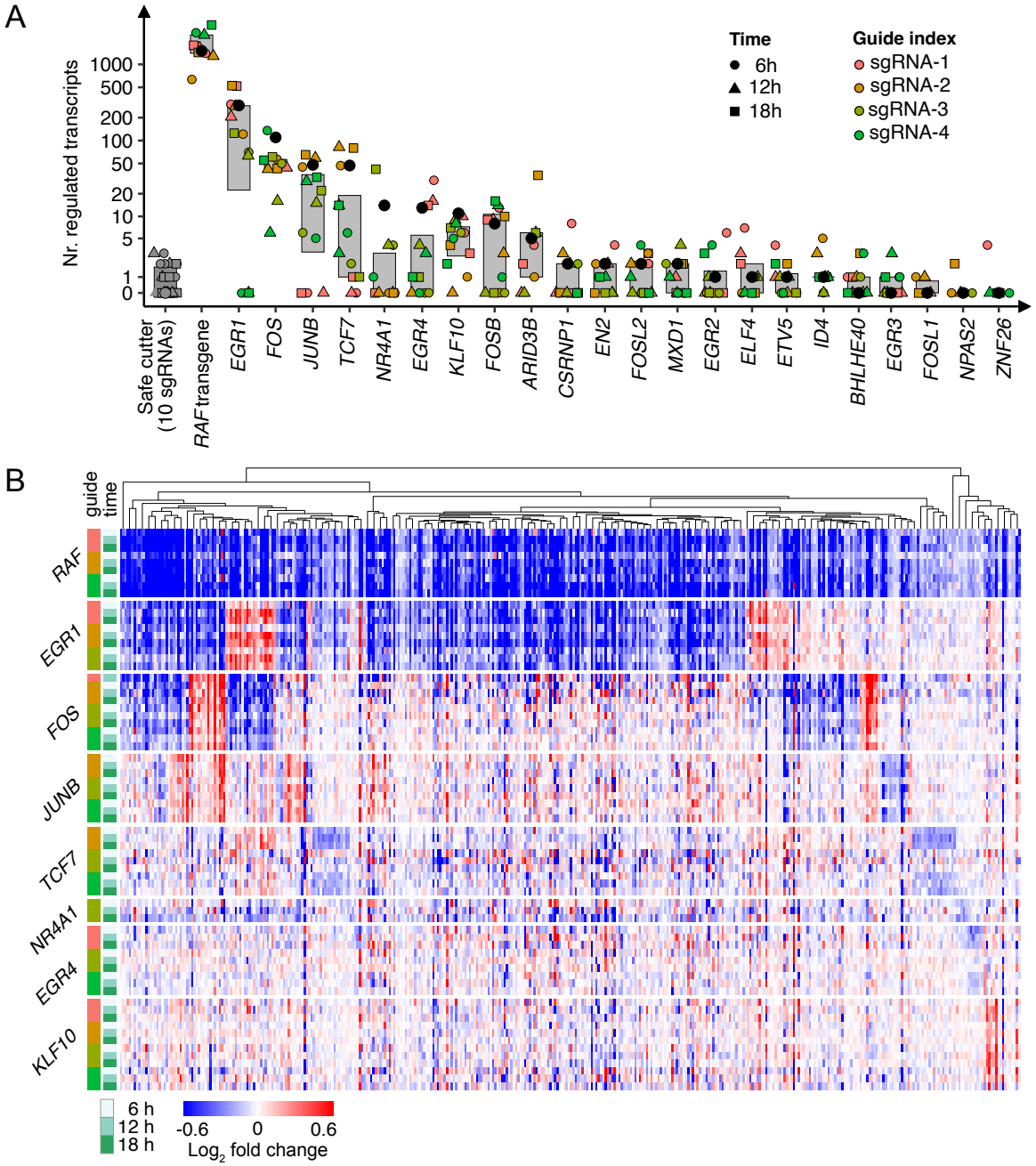


Figure 25 - Summary of Perturb-seq screen results.

(A) *Number of significantly deregulated transcripts (adjusted p-value <0.05) following perturbation of the indicated candidate TF separated by sgRNAs and time points.* ***(B)*** *Heatmap of the log_2 fold change of significantly deregulated genes in response to all candidate TFs whose perturbation resulted in significant deregulation of target genes, including the RAF transgene positive control. The results from perturbed target genes separated by sgRNAs and time points are shown. Negative values indicate lower and positive values higher target transcript levels in the perturbed cells relative to cells expressing sgNTCs. Perturb-seq screens data was analyzed by Nils Blüthgen. TAP-seq PCR amplification of 140 target transcripts was performed by Bertram Klinger and Anja Sieber.*

2.3.6 De novo construction of a TF core network identifies an EGR1-TCF7 positive feedback loop within the RAF-MAPK transcriptional response

To analyze the transcriptional impact of each candidate TF at the gene level, we aggregated data across all time points and sgRNAs targeting the same gene to calculate gene expression fold changes. Fisher's method was then used to calculate significantly deregulated targets by combining the p-values from individual sgRNAs and time points. This allowed us to create a "core heatmap" focusing only on the candidate TFs and how they regulate each other (Fig. 26A). We began by examining the perturbation effects of each TF on its own expression. We observed a consistent trend of target gene downregulation, likely attributable to the introduction of frameshift mutations by Cas9-mediated INDEL formation in early exonic regions, which triggers nonsense-mediated mRNA decay (298). In contrast, *EGR1*, *FOS*, and *CSRNP1* displayed an unexpected upregulation of their transcripts following their knockout, which may reflect compensatory auto-regulatory feedback mechanisms, where the loss of functional protein leads to increased transcription or stabilization of the edited, stop-gain mutation-containing mRNA. Together, these results highlight the importance of careful interpretation when evaluating perturbation efficiency, considering both the downregulation of the targeted transcript and the secondary effects on downstream gene expression.

Next, we examined how the candidate TFs regulate the expression of the other TFs in our list. As seen previously in Figure 25, EGR1 and FOS emerged as central regulators within the TF core heatmap, controlling the expression of several other candidate TFs in an orthogonal manner. EGR1 primarily acts as an activator, while FOS functions as an inhibitor within the network. EGR1 activated the same targets that FOS inhibited, including JUNB, KLF10, ID4, and NR4A1 (Fig. 26A). Notable exceptions to this pattern include FOSL1, which EGR1 inhibited but FOS activated, and EN2, the only candidate TF co-activated by both EGR1 and FOS. Additionally, FOS directly inhibited EGR1 expression, reinforcing its inhibitory role in the network. Collectively, these findings highlight the previously established central and contrasting roles of *EGR1* and *FOS* and emphasize their central role in orchestrating the transcriptional response downstream of RAF1.

Finally, we created a core transcriptional network using the TF core heatmap. Each TF is represented as a node, connected by directional edges to its significantly affected target genes. Upon examining this network, we identified multiple coherent feed-forward loops. For instance, FOS inhibits JUNB expression both directly and indirectly by inhibiting the upstream EGR1. However, a limitation of the perturbation-based approach is that it does not allow us to differentiate between direct and indirect regulatory relationships, as both configurations yield identical response patterns in the resulting heatmap (see example in

Fig. 26B). To resolve this ambiguity, we simplified the model by excluding edges that represent potential feed-forward loops, retaining only the minimal set of regulatory connections necessary to explain the observed data (Fig. 26C).

Among the interactions revealed by this core network, we identified a positive feedback loop between *EGR1* and *TCF7*, in which both genes activate one another. Analysis of the bulk RNA-Seq data shown in Figure 20B reveals that these two genes belong to distinct temporal classes of RAF1-responsive genes: *EGR1* is classified as an immediate-early gene with a half-maximal induction time of 17 min, whereas *TCF7* is part of the secondary response genes, reaching half-maximal induction only after 5 h and 8 min. In the following section, we investigate this loop in greater detail to determine its regulatory dynamics and potential biological significance within the context of MAPK and Wnt signaling.

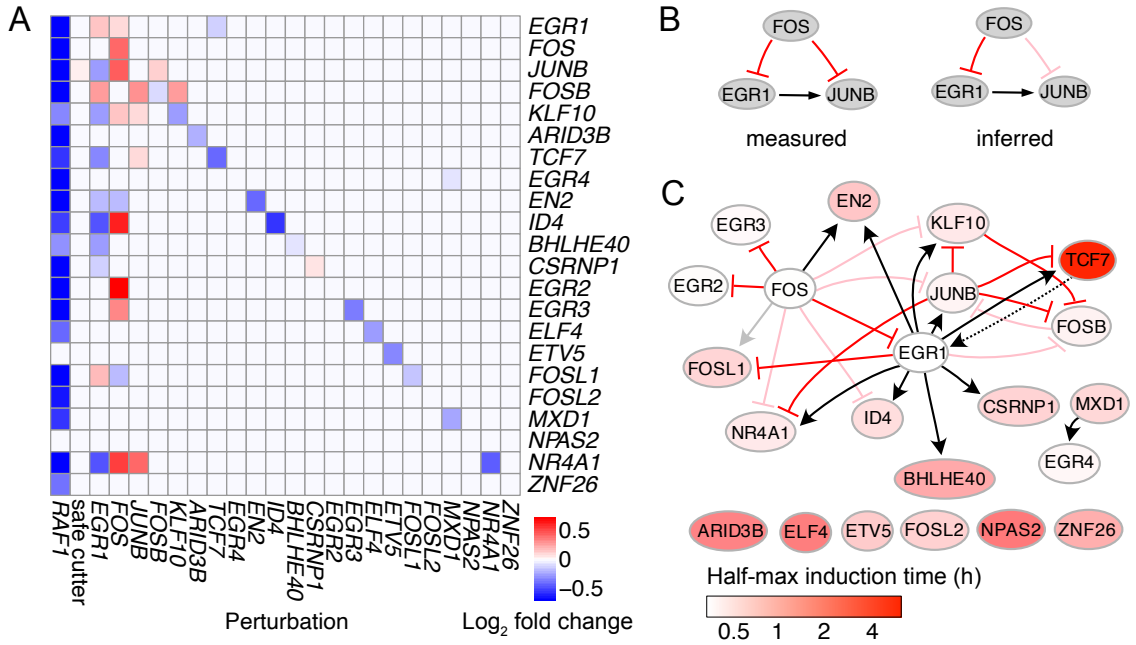


Figure 26 - Model of transcriptional interactions between the 22 candidate TFs.

***(A)** Heatmap of candidate TF expression changes following the perturbation of all 22 TFs and RAF1, showing the log_2 fold change upon perturbation for the significantly differentially expressed TFs (adjusted p-value <0.05). **(B)** Example of the removal of edges from coherent feed-forward loops: The measured perturbation data in A is compatible with a feed-forward loop, where FOS directly and indirectly inhibits JUNB, but it is also compatible with a cascade, in which FOS inhibits JUNB via EGR1 only. To create the most parsimonious network, we removed the feed-forward loop from FOS to JUNB in the inferred network. **(C)** De novo model of the TF core network showing directional interactions between all perturbed TFs with inferred interaction type: Black = Activating, Red = Inhibiting. Transparent edges represent the removed feed-forward loops. TFs are color-coded by their half-maximal induction time. Perturb-seq screens data was analyzed by Nils Blüthgen.*

2.3.7 RAF-MEK-ERK and Wnt signaling pathways cross-talk mediated by TCF7

One of the critical interactions in cancer involves the MAPK pathway and the Wnt/β-catenin signaling pathway. Figure 27A shows a simplified schematic of both signaling pathways, including the *EGR1-TCF7* positive feedback loop identified by the Perturb-seq screen. EGR1, a well-characterized effector of the MAPK pathway, is rapidly induced in HEK293ΔRAF1:ER cells upon treatment with 4OHT. On the other hand, TCF7 is more prominently recognized for its central role in mediating transcriptional responses to Wnt pathway activation, despite also being upregulated following RAF1 induction (Fig. 20). In HEK293 cells, Wnt signaling can be induced by treatment with the GSK3β inhibitor CHIR99021, which blocks the ubiquitination and subsequent proteasomal degradation of β-catenin, leading to the activation of Wnt signaling (Fig. 27A).

To validate the *EGR1-TCF7* positive feedback loop observed in the Perturb-seq screen, we performed RT-qPCR to measure *EGR1* and *TCF7* transcript levels following their respective knockouts. Knockout of *TCF7* resulted in a modest but significant reduction in *EGR1* fold change expression after treatment with 4OHT from 152-fold (+/- 5.1) to 135-fold (+/- 8.2), indicating that TCF7 functions as an activator of EGR1 (Fig. 27B). Conversely, *EGR1* knockout resulted in a decrease in *TCF7* fold change expression after treatment with 4OHT from 5.3-fold (+/- 0.6) to 1.68-fold (+/- 0.2), confirming that EGR1 promotes TCF7 expression as well (Fig. 27C). Together, these findings support the existence of the *EGR1-TCF7* positive feedback loop initially inferred from the Perturb-seq data.

To confirm the selective activation of MAPK and Wnt signaling pathways in response to the treatment of cells with either 4OHT or CHIR99021, we measured the expression of *EGR1*, *TCF7*, and *AXIN2*, an established Wnt signaling marker, following treatment with either compound. Treatment with 4OHT resulted in a 188-fold (+/- 8.7) increase in *EGR1* and a 5.8-fold (+/- 0.2) increase in *TCF7* mRNA expression levels, while *AXIN2* mRNA levels showed only a slight increase of 1.3-fold (+/- 0.3), consistent with specific MAPK pathway activation. In contrast, CHIR99021 treatment led to an 8-fold (+/- 0.8) and a 3-fold (+/- 0.2) increase in *AXIN2* and *TCF7* expression, respectively, which indicates efficient Wnt signaling activation. Interestingly, treatment with CHIR99021 also caused a 5.2-fold (+/- 0.6) increase in *EGR1* expression (Fig. 27D)

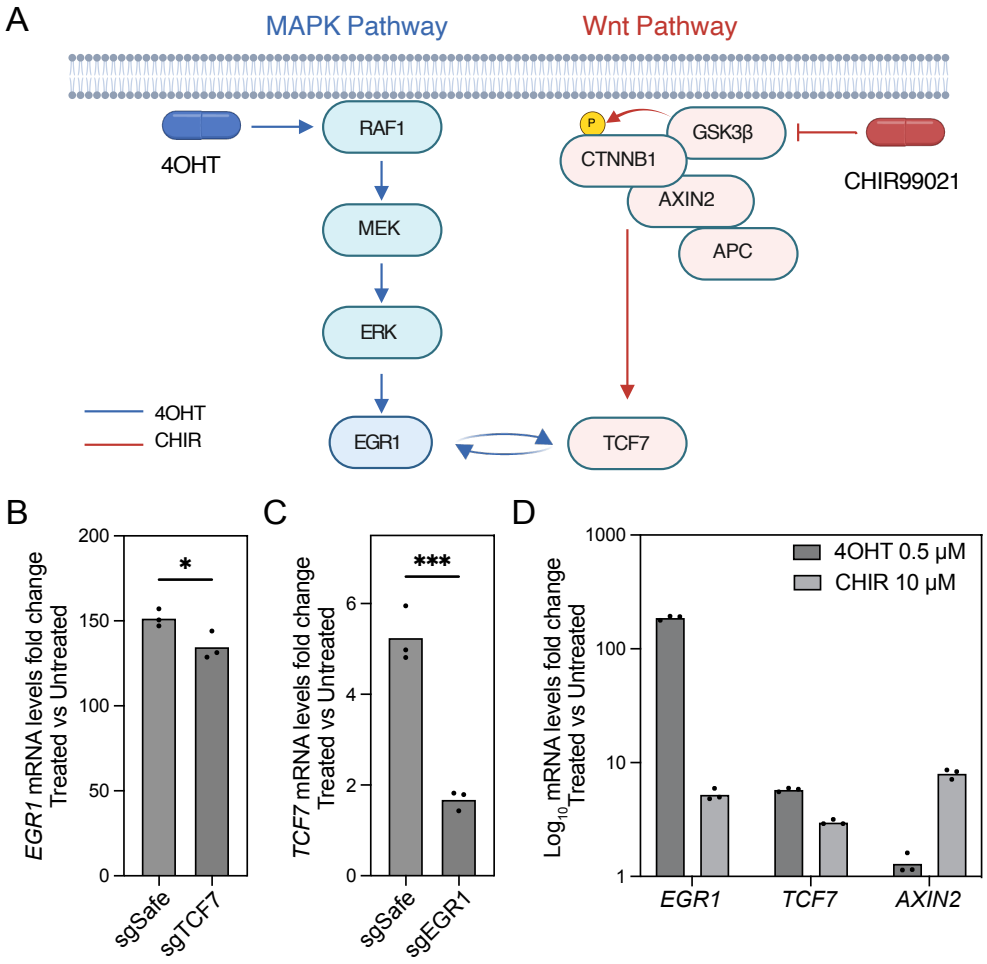


Figure 27 - Characterization of MAPK and Wnt signaling pathways cross-talk through EGR1-TCF7 positive feedback loop.

***(A)** Schematic representation of the MAPK-Wnt signaling cross-talk via the EGR1-TCF7 positive feedback loop. The RAF-MEK-ERK pathway is activated by 4OHT treatment. Wnt signaling is initiated by CHIR99021, a GSK3β inhibitor. **(B)** Validation of the TCF7-EGR1 arm of the positive feedback loop identified in the Perturb-seq screens through RT-qPCR measurement of the fold change in EGR1 mRNA levels following the perturbation of TCF7 after treating HEK293ΔRAF1:ER cells with 0.5 μM 4OHT for 12 h. Values represent the mean of biological replicates (n=3). **(C)** Validation of the EGR1-TCF7 arm of the positive feedback loop identified in the Perturb-seq screens through RT-qPCR measurement of the fold change in TCF7 mRNA levels following the perturbation of EGR1 after treating HEK293ΔRAF1:ER cells with 0.5 μM 4OHT for 12 h. Values represent the mean of biological replicates (n=3). **(D)** RT-qPCR measurement of fold change in EGR1, TCF7, and AXIN2 mRNA levels after treating HEK293ΔRAF1:ER cells with 0.5 μM 4OHT or 10 μM of CHIR99021. Values represent the mean of biological replicates (n=3). * p-value <0.05, ** p-value <0.01, *** p-value <0.001.*

To further explore the cross-talk between the two pathways, we performed bulk RNA-Seq after CHIR99021 treatment of the HEK293ΔRAF1:ER, both with and without *TCF7* knockout, to identify the genes regulated by TCF7 in the context of Wnt pathway activation. Our results showed that activation of Wnt signaling led to the *TCF7*-dependent upregulation of *DKK1*, *CALCB*, *HES6*, *FAH*, *GPRIN1*, and *BICDL1* (Fig. 28A). Among these, we chose *DKK1* for further investigation using RT-qPCR under different treatment

conditions: 4OHT alone, CHIR99021 alone, and a combination of both. Our results showed that 4OHT treatment had no effect on *DKK1* expression levels, while CHIR99021 treatment caused a 2-fold increase in its expression. However, simultaneous activation of MAPK and Wnt signaling led to an additional 0.9-fold increase in *DKK1* expression, resulting in a total increase of 2.9-fold (Fig. 28B). This indicated that MAPK signaling enhances Wnt signaling gene expression and further validated the cross-talk between both pathways. *TCF7* knockout, on the other hand, resulted in a decrease in *DKK1* levels with CHIR99021 and the combined treatment. Nonetheless, some expression was still detected, indicating that additional regulators beyond TCF7 may contribute to *DKK1* transcription (Fig. 28B). Taken together, our results suggest that *TCF7* mediates the cross-talk between the MAPK and Wnt signaling pathways by enhancing Wnt signaling gene expression when both pathways are activated.

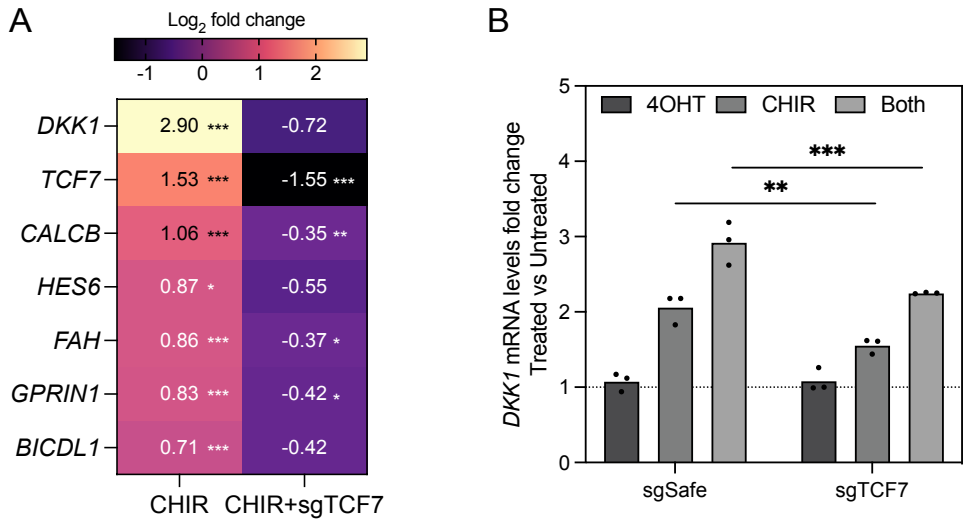


Figure 28 - TCF7 mediates the cross-talk between the MAPK and Wnt signaling pathways.

(A) *Heatmap showing the transcriptional changes detected via bulk RNA-Seq from cells treated with 10 µM of CHIR99021, both with and without TCF7 knockout. Log₂ fold change values are calculated with reference to untreated cells. Values represent the mean of biological replicates (n=3).* ***(B)*** *RT-qPCR measurement of DKK1 mRNA levels fold change after treating cells with 0.5 µM 4OHT for 12 h, 10 µM CHIR99021 for 6 h, or both treatments combined. Values represent the mean of biological replicates (n=3). * Adj. p-value <0.05, ** Adj. p-value <0.01, *** Adj. p-value <0.001. Bulk RNA-Seq data was analyzed by Nils Blüthgen.*

2.4 Discussion

The MAPK signal pathway is a central cascade in cell communication with mutations in its components associated with numerous human diseases such as cancer (30). Extensive research has been dedicated to characterizing the upstream regulation of this pathway, which resulted in compiling a comprehensive map of its elements and their interactions (271). These studies have identified the mechanisms through which MAPK signaling

generates distinct temporal dynamics, which in turn control cell outcomes (299–302). Time-resolved transcriptome analysis following ERK activation enabled the systematic study of the order and kinetics of ERK-target genes response. These experiments demonstrated that the mRNA half-life of transcripts filters short pulses and is a central element of the initial cellular response (285). Despite the detailed research of ERK-induced transcriptome kinetics and associated proteins (296), the structure of the transcriptional network and the mechanisms through which these dynamic signaling patterns are translated into defined transcriptional programs remain unclear (303). More precisely, which TFs govern immediate early genes and how these induced TFs, in turn, regulate the secondary response genes remains understudied. Gaining insights into this network's topology and dynamics is important for understanding how this pathway manages divergent, and at times contradictory, outcomes and how it processes and interprets genetic information. One of the promising methods for bridging this gap is transcriptional network modeling from perturbation data (290). scCRISPR-seq technologies now enable systematic profiling of gene expression changes in response to genetic perturbations at scale, which enables *de novo* construction and tuning of transcriptional network models associated with the RAF-MAPK signaling pathway.

In this study, we used a targeted Perturb-seq (TAP-seq) strategy (228) to evaluate the transcriptional consequences of disrupting 22 TFs activated downstream of RAF1-mediated MAPK signaling in a HEK293ΔRAF1:ER model cell line engineered with a tamoxifen-inducible RAF1-CR3 kinase domain. The selection of the 22 TFs for perturbation was guided by prior knowledge of their inducibility upon ERK activation (285) as well as bulk RNA-Seq measurement of their expression kinetics after pulse activation of RAF1 signaling (Fig. 20). Calculation of the half maximal induction time enabled the classification of the selected TFs into four distinct response classes: IEG, ILG, DEG, and SRG (Table 1). Many of these factors, including *EGR1*, *FOS*, and *JUN*, are well-characterized downstream effectors of MAPK signaling, rapidly induced in response to pathway activation and implicated in controlling key cellular decisions such as proliferation, differentiation, and apoptosis (304–307). Others, such as *TCF7* and *BHLHE40*, represent less-characterized but transcriptionally responsive factors that may participate in feedback or cross-pathway regulation. By including both canonical and less-studied but ERK-responsive TFs that belong to different response classes of ERK signaling, the experimental design captures a broad spectrum of the transcriptional landscape downstream of RAF-MAPK signaling.

To analyze the generated Perturb-seq data, we applied a topology-based modeling strategy to construct a "core network", outlining directional interactions among the 22

perturbed candidate TFs (Fig. 26C). We identified several coherent feed-forward loops where a TF influences a target gene both directly and indirectly through another TF. Since these feedforward loops could not be resolved with the available data, they were removed from the network to reach the most parsimonious network structure. This method of network simplification was previously used in transcriptional network inference tools such as ARACNE, which discards the least supported edge within a triangular regulatory relationship to reduce false positives (308). Since coherent feed-forward loops are well-documented in biological systems and often fulfill important regulatory roles (309), future studies aiming to study specific biological functions of such motifs could benefit from orthogonal CRISPR strategies in which gene knockouts are combined with simultaneous gene activation within the same cell to dissect directional regulatory interactions (198). Notably, the immediate early genes *EGR1* and *FOS* emerged as the most interactive upstream regulators, with orthogonal regulatory influence on the expression of other TFs and downstream targets within the network.

EGR1 and *FOS* are part of the IEGs which are among the earliest responders to *ERK* activation. These genes are directly induced by ERK-regulated SRF-TCF complexes and play a significant role in inducing cell proliferation (310). *FOS* both serves as an initial response gene and as a crucial ERK signal duration sensor (311,312). Under continuous ERK signaling, FOS protein becomes stabilized, increasing its ability to promote the transcription of specific secondary response genes. Transient ERK activity, on the other hand, results in a decline of FOS protein levels before substantial accumulation occurs, limiting the transcriptional response (313). This temporal regulation is also adjusted by downstream phosphorylation cascades and autoregulatory feedback loops that further amplify the accumulation and activation of IEG TFs, enabling the cell to effectively distinguish between sustained and recurrent signals.

Several delayed early and secondary response genes have negative feedback regulatory functions in signaling networks (294,295). This suggests that disruption of EGR1 or FOS function could impair negative feedback mechanisms, potentially resulting in increased MAPK signaling. This effect may also be observed in our core network, where we see that FOS and EGR1 have an inhibitory influence on multiple immediate early TFs and immediate early genes. However, to fully study the complex feedback regulation between transcriptional networks and signal transduction, perturbation experiments that not only capture transcriptomic responses but also monitor dynamic changes in signaling activity directly through specific MAPK signaling reporters are required (314,315).

In addition, the core network revealed a positive feedback loop between *EGR1* and *TCF7*, linking MAPK and Wnt signaling (Fig. 26C). Signaling pathways often interact with each other through a process known as pathway cross-talk, which allows cancer cells to evade normal regulatory mechanisms. In the context of cancer, the MAPK pathway and the Wnt/β-catenin signaling pathway interact extensively (278). While EGR1 is a downstream TF of the RAF-MEK-ERK MAPK signaling pathway, TCF7 is a key transcriptional regulator within the Wnt signaling pathway (Fig. 27). Our results indicate that the interplay between these two TFs provides a mechanism by which MAPK signaling can reinforce Wnt pathway activity, potentially enhancing oncogenic transcriptional programs (Fig. 28). Such interaction has been previously identified where a hidden positive feedback loop caused by cross-talk between the Wnt and ERK pathways can contribute to sustained activation of both pathways in cancer (316). In addition, it has been shown that *EGR1* and *TCF7* cooperate in a positive feedback loop that amplifies *LCN2* expression and promotes aggressive cancer behaviors in esophageal squamous cell carcinoma cells (317). In other cancers, such as colorectal cancer or melanoma, where aberrant activation of both the MAPK and Wnt pathways is commonly observed, targeting the *EGR1-TCF7* feedback loop may offer a novel therapeutic strategy to interrupt this pathological cross-talk, reducing the growth potential of tumors that rely on sustained MAPK and Wnt pathway signaling (318–321). The identification of the *EGR1-TCF7* positive feedback loop provides an additional example of this phenomenon, where these TFs that are key regulators in different signaling pathways mutually reinforce each other's expression.

While this study provides significant insights into the transcriptional networks of RAF-MAPK signaling using the direct capture Perturb-seq approach, there are several methodological limitations that warrant discussion. First, although we perturb 22 key TFs activated by RAF-MAPK signaling, transcriptional networks are inherently complex and therefore, many additional TFs not considered in our study may also play important regulatory roles in mediating MAPK pathway responses. In addition, we did not detect significant deregulation of target genes for more than half of our candidate TFs. Since scRNA-seq has limitations of low mRNA capture efficiency and high dropouts (217), this can be partly explained by the relatively low number of genes that are detected in Perturb-seq experiments. In our experiments, we detected between 2,261 and 3,599 genes per cell, therefore, target genes of TFs that are lowly expressed could not be detected (Supp. Fig. 10C). Another likely explanation is the presence of compensatory mechanisms that buffer the loss of individual TFs, such as functional redundancy within TF gene families or the activity of gene paralogues with overlapping regulatory roles. These compensatory effects may mask the transcriptional consequences of single TF perturbations (50). Future

studies could overcome this limitation by expanding the gene panel used in the targeted Perturb-seq approach, employing CRISPRa methods to increase the expression levels of candidate TFs along with their downstream target genes, or implementing higher-order combinatorial Perturb-seq screens to simultaneously perturb multiple TFs and overcome compensatory regulatory mechanisms (138,152,225,322).

Another limitation of this study is the use of a single synthetic model cell line, which may not fully capture the cell type specific variations in MAPK signaling and transcriptional regulation. The regulatory networks downstream of RAF-MAPK signaling can vary significantly between different cell types, including different cancer models (323,324). Therefore, the network we reconstructed may not fully reflect the dynamics in other cellular contexts. Future studies should extend these analyses to a wider array of cell types, including cancer cell lines or primary cells, to better understand how MAPK signaling regulates transcription across different biological contexts.

Finally, a key component of the direct capture Perturb-seq method is the introduction of specific capture sequences within the Cas9 sgRNA tracr sequence, which enables 3' based annealing and sequencing of the sgRNA sequences for the identification of perturbations at the single-cell level (225). A critical consideration in the design of these capture sequences is to ensure that their addition does not interfere with the sgRNA's ability to guide the Cas9 nuclease to the intended DNA target or with its on-target efficiency. Initial studies with CRISPRi, have demonstrated that the modifications to the Cas9 tracr sequence did not affect the efficacy of the sgRNA in mediating gene perturbation (225,227). However, a recent study comparing three different scCRISPRseq approaches revealed that CS1-modified sgRNAs are less active than unmodified sgRNAs (325). High on-target efficiency is important in Perturb-seq experiments since efficient gRNAs ensure that a sufficient proportion of cells in the population carry the intended gene perturbation, which is necessary to observe phenotypic changes in the subsequent transcriptomic analyses. Therefore, using highly efficient gRNAs improves the signal-to-noise ratio in the downstream data by maximizing the impact of the target perturbation on gene expression. Future studies can benefit from the use of single-cell 5' CRISPR screening approaches. In this method, a guide-specific RT oligo is used that anneals to the standard constant region present in most sgRNA designs during the reverse transcription process. With this strategy, gRNA libraries with unmodified Cas9 tracr sequences can be used without the need for the addition of capture sequences. As such, this method shows superior gRNA perturbation efficiency compared to the 3' CRISPR screening method (325). These limitations highlight areas where future research can improve upon our findings and address unresolved questions within the field.

2.5 Conclusion

The MAPK signaling pathway is implicated in the regulation of fundamental cellular processes, including proliferation, differentiation, survival, and apoptosis, and its dysregulation is commonly associated with various human diseases such as cancer. While upstream regulators of this pathway are well studied, how its activation translates into different transcriptional responses remains poorly understood. Perturb-seq, which combines single-cell RNA sequencing with CRISPR-based genetic perturbations, provides a powerful approach for dissecting complex gene interactions within signaling networks and allows mapping of transcriptional changes at a single-cell level following gene perturbations. In this study, we applied Perturb-seq to study the interactions between 22 TFs downstream of RAF-MAPK signaling in an inducible model cell line. Using a topology-based modeling approach, we constructed a TF interaction network where we identified a positive feedback loop between *EGR1* and *TCF7*, establishing an essential nexus linking MAPK and Wnt signaling pathways. Further investigation of this cross-talk showed that *TCF7* mediates the cross-talk between the MAPK and Wnt signaling pathways by enhancing Wnt signaling gene expression when both pathways are activated. In conclusion, the approach outlined in this chapter provides a foundational framework for the systematic investigation of transcriptional network dynamics. Its scalability and robustness make it well-suited for broader applications beyond RAF-MAPK signaling.

Conclusion and Remarks

Understanding how genes regulate each other and interact within complex networks is important for resolving the complexity of cellular systems. Gene regulation and genetic interactions control critical cellular processes such as proliferation, differentiation, and apoptosis, often through tightly controlled signaling pathways. These pathways translate external signals into specific transcriptional responses that determine cell fate. By systematically perturbing genes, scientists can study transcriptional programs and map genetic interactions, which allows them to study the functional architecture of cell signaling pathways, and identify vulnerabilities in disease related pathways offering new targeted therapy options. This thesis presents two complementary CRISPR-based approaches to understanding gene function, regulation, and interactions.

In Chapter 1, we investigated the utility of Cas13d as a robust RNA targeting platform for quantitative genetic interaction mapping. Contrary to traditional DNA targeting systems such as Cas9 and Cas12a, Cas13d allows precise and rapid post-transcriptional perturbation of gene expression, generating homogenous knockdown cell populations. By systematically optimizing gRNA expression using two promoter systems and using single-gene arrays, we addressed gRNA-gRNA sequence-dependent interference, increased effect sizes of perturbation phenotypes and GI scores, and reduced the library size by nine-folds. We finally use pooled combinatorial screening to demonstrate Cas13d's ability to reveal both known and novel genetic interactions between six genes involved in the chronic myeloid leukemia model cell line K562's response to tyrosine kinase inhibitor imatinib. The methodological improvements presented here lay the groundwork for future large-scale, compact, Cas13d-based combinatorial screening in different disease and cellular contexts.

In Chapter 2, we apply scCRISPR-seq screening to study the downstream transcriptional regulatory networks of 22 transcription factors in an inducible RAF1 model cell line. Using the direct capture Perturb-seq data, we reconstruct a *de novo* transcriptional regulatory network between our candidate transcription factors, where *EGR1* and *FOS* emerge as central regulators in the RAF-MAPK pathway. In addition, we uncover a positive feedback loop within this network between *EGR1* and *TCF7* that bridges MAPK and Wnt signaling and allows the cells to control the signaling intensities of these two pathways. This work provides a mechanistic view into how signal transduction pathways coordinate structured gene expression programs and underscores the importance of mapping transcriptional regulators to understand cell signaling pathways through single-cell CRISPR screening approaches.

Taken together, the tools and insights presented in this thesis provide high content screening methods to study gene function and network topology. Cas13d-based GI mapping and Perturb-seq-based transcriptional network analysis are complementary tools that allow us to study genetic interactions and signaling pathway regulators. Further research should build on this framework by extending these tools to larger screens to investigate larger gene panels in other signaling pathways, tissue types, and disease models. In the future, integrating results from complementary perturbation strategies such as CRISPRi and CRISPRa, as well as proteomic studies, should enable the generation of more comprehensive genetic interaction maps and models. With these tools, the CRISPR-based functional genomics can move beyond linear gene-function studies and toward a systems-level understanding of cellular regulation and its disruption in disease.

Materials and Methods

2.6 Materials

Consumables, disposables, chemicals, and reagents that were used are provided within the methods section with their respective suppliers and catalog numbers.

2.7 Methods

2.7.1 Vector maps

The plasmids pXR001-mCD4, 783-Rx-hU6, 783-Rx-mU6, 783-Rx-Dual, pMB1, and AiO-Cas12a were used to investigate the utility of Cas13d for GI screening. The all-in-one Cas9 vector pMB1, containing the 10x Genomics capture sequence 1 (5′-GCTTTAAGGCCGGTCCTAGCAA-3′) in the stem-loop of the Cas9-tracr sequence, referred to as pMB1-10x, was employed in the Perturb-seq experiments (225). The plasmids and their sequences were deposited at Addgene.

2.7.2 K562, Lenti-X 293T, and HEK293∆RAF1:ER cell culture

K562 (ATCC, CCL-243) cells were cultured in complete RPMI supplemented with 10% fetal bovine serum (Sigma-Aldrich, 11875093) and 1% pen/strep antibiotics (Sigma-Aldrich, P0781-100ML). Lenti-X 293T cells (Takara, 632180) were cultured in complete DMEM (Thermo Fisher Scientific, 11995073) supplemented with 10% fetal bovine serum and 1% pen/strep antibiotics (Sigma-Aldrich, P0781-100ML).

HEK293∆RAF1:ER cells (282) contain a tamoxifen-inducible fusion of the kinase domain of RAF1 (reviewed in Samuels *et al.*, 1993 and McMahon, 2001) (297,326), were cultured in complete DMEM low glucose without phenol red (Thermo Fisher Scientific, 11880028), supplemented with 10% fetal bovine serum (Pan Biotech, P30-1502) and 1% pen/strep antibiotics (Sigma-Aldrich, P0781-100ML).

2.7.3 K562-Cas13d-mCD4 clonal line

The pXR001 plasmid (150) was digested with NheI (NEB, R0131). The *mCD4* gene was amplified from the *S. aureus* Cas9 nuclease vector (Addgene #105998) using Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) following the manufacturer's protocol. P2A fragment and Gibson Assembly-compatible overhangs were added to the digested pXR001 plasmid site using the following primers: 5’-CGATGTTCCAGATTACGCTGGATCCGGCGCAACAAACTTCTCTCTGCTGAAACAAG CCGGAGATGTCGAAGAGAATCCTGGACCGCCGGACATGTGCCGAGC-3’ and 5’-

GCCCTCTCCACTGCCGCCCTGGCGCTGTTGGTGCCGG-3'. The amplified PCR fragment was purified from a 1% Agarose gel using NucleoSpin columns (Macherey-Nagel, 740609.250) and subsequently cloned into the digested pXR001 plasmid using Gibson Assembly (327).

The Cas13d-mCD4 gene was introduced into the K562 cell line through lentiviral transduction. The cells were stained with an anti-mouse CD4 antibody (Miltenyi, 130-116-526) using a final antibody dilution of 1:50 in 100 µL PBS with 0.5% BSA buffer, and single clones were sorted using a BD Melody Flow Cytometer 19 days post-transduction. The single clones were expanded for 5 weeks. To assess the functionality of the expanded clonal lines, we transduced the cells with a gRNA to knock down CD46 (5'-CAGACAATTGTGTCGCTGCCATC-3'). The clonal lines were screened for the functionality of the Cas13d system after 10 days via flow cytometry analysis of >10,000 cells stained with CD46 antibody (Miltenyi, 130-104-509) using a final antibody dilution of 1:100 in 100 µL PBS with 0.5% BSA buffer. The best performing clonal line out of 24 clonal lines was selected for the CRISPR screens and further experiments.

2.7.4 Cas9 gRNAs design and cloning for kinetics and dual knockout experiments

The web tool CRISPick (328,329) was utilized to select 4 *sp*Cas9 sgRNAs targeting *CD46*, *CD47*, *CD63*, and *CD71* (Supp. Table 5). The sgRNAs were ordered as complementary single-stranded oligos from Sigma-Aldrich, following this structure for the sense and antisense strands respectively: 5'-TTGGNNNNNNNNNNNNNNNNNNN-3' and 5'-AAACNNNNNNNNNNNNNNNNNNN-3'. The complementary sequences were then joined by incubating 10 µM of each oligo in T4 ligation buffer in a 10 µL reaction in a thermocycler using the following program: 37 °C for 30 min, 95 °C for 5 min, and a ramp down from 95 °C to 25 °C at 0.1 °C/s. Meanwhile, the pMB1 vector was digested with AarI (Thermo Fischer Scientific, ER1582), and the linear plasmid was purified from a 1% agarose gel. The annealed oligos and the digested plasmid were then combined via T4 ligation. To do this, the annealed oligos were diluted 1:200 with water, and 1 µL was mixed with 1 µL of 10x T4 ligation buffer and 5 U of T4 DNA Ligase (Thermo Fisher Scientific, EL0011) in a 10 µL reaction. The mixture was incubated at 16 °C overnight. The next day, 2 µL were used to transform chemo-competent *DH5α*. Single colonies were subsequently picked from the LB agar plates, and the plasmids were isolated from liquid cultures using NucleoSpin columns (Macherey-Nagel, 740588.250).

To test the functionality of the cloned gRNAs, we transduced the HEK293ΔRAF1:ER (285) cells with the *CD46* gRNAs, while the K562 cells were transduced with the *CD47*, *CD63*, and *CD71* gRNAs. After 20 days, the cells were screened for the knockout of *CD46,* and

after 7 days for the knockout of *CD47*, *CD63*, and *CD71,* using flow cytometry analysis of >10,000 cells stained with CD46-APC, CD47-APC, CD63-APC, or CD71-APC antibodies (Miltenyi, 130-104-558, 130-123-315, 130-118-151, 130-115-030) using a final antibody dilution of 1:100 in 100 µL PBS with 0.5% BSA buffer. The best performing sgRNAs were used for further experiments (Supp. Fig. 7 and Supp. Table 5).

For the dual knockout of *CD46* and *CD47*, the sgRNAs in the first position were expressed from a mU6 promoter and the one in the second position from an H1 promoter. Modified Cas9 tracr sequences WCR2 and VCR1L were selected to avoid recombination (266). The pMB1 plasmid was digested with PaqCI and NheI (NEB, R0745S, R0131), and the linear plasmid was purified from a 1% agarose gel. Gene fragments with Gibson assembly-compatible overhangs of the following structure gRNA1-WCR2-filler-VCR1L-gRNA2-H1promoter were ordered from Twist Biosciences (Supp. Table 6). The gene fragments were PCR amplified using Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) according to the manufacturer's instructions with the primers: 5'-CCCTTGGAGAACCACCTTGTTGG-3' and 5'-AATACGGTTATCCACGCGGCC-3'. The amplified PCR fragment was purified from a 1% agarose gel using NucleoSpin columns (Macherey-Nagel, 740609.250) and cloned into the digested pMB1 plasmid using Gibson Assembly (327).

2.7.5 All-in-one-Cas12a plasmid cloning

The all-in-one-Cas12a (AiO-Cas12a) plasmid consists of the en*As*Cas12a (143) gene expressed from an EF1α promoter, along with a gRNA expression cassette from a mU6 promoter. The en*As*Cas12a gene was PCR amplified from the pCAG-enAsCas12a plasmid (143) using a nested PCR strategy for cloning. The PCR reactions were conducted using the Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) according to the manufacturer's instructions, with the following primers for the first PCR reaction: 5'-ACCGGTTCTAGAGCGATGCCTGCTGCTAAGAGAGTGAAACTGGATCCTGCTGCTAAGAGAGTGAAACTGGATCCTGCTGCTAAGAGAGTGAAACTGGATACACAGTTCGAGGGCTTTACC-3' and 5'-GGATCCGCTAGCGTTGCGCAGCTCCTGGA-3'. For the second PCR reaction, Gibson Assembly compatible overhangs were added to the PCR fragments from the first PCR reaction using these primers: 5'-AACACAGGACCGGTTCTAGACTAGAGCGATGCCTGCT-3' and 5'-AGAGAGAAGTTTGTTGCGCCGGATCCGCTAGCGTTGCG-3'. The CROP-Seq Cas9 Plasmid was digested with XhoI and BamHI (NEB, R0146S, R0136S), and the linear plasmid was purified from a 1% agarose gel. The enAsCas12a gene was inserted into the digested CROP-Seq Cas9 Plasmid backbone via Gibson Assembly (327). The resulting

plasmid was subsequently digested with SacII and SnaBI (NEB, R0157S, R0130S). A gene fragment containing WPRE-3LTR with Gibson Assembly compatible overhangs was ordered from Genscript (Supp. Table 10) and cloned into the digested plasmid. The final plasmid was then digested with XhoI and the IlluminaPBS-Filler-mU6 gene fragment (Genscript), which was PCR amplified using the following primers: 5'-GATCCACTTTGGCGCCGGCCTCGAGCAG-3' and 5'-CTTTCAAGACCTAGGGCCCCCCTCGAGCCCGGGCATGCTCTTCAACCTCAATAACT GGAGTTATATGGACCATTGTTCTAGCGCTGATCCGACG-3'. The PCR reaction used the Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) according to the manufacturer's instructions. This fragment was then inserted into the digestion site using Gibson Assembly. The final plasmid included the enAsCas12a gene linked to the puromycin resistance gene under the EF1α promoter, along with the gRNA expression cassette positioned under the mU6 promoter in the reverse orientation.

2.7.6 Cas12a gRNAs design and cloning for kinetics and dual knockout experiments

Cas12a gRNAs targeting *CD63* and *CD47,* used by DeWeirdt *et al.* 2021 to assess and optimize Cas12a performance for combinatorial screens, were selected (197). Cas12a gRNAs targeting *CD46* and *CD47* were chosen using the web tool CRISPick (328,329) (Supp. Table 7). The gRNAs for single gene knockout were ordered as complementary single-stranded oligos from Sigma-Aldrich, using the following structure for the sense and the antisense strands respectively: 5'-AAAANNNNNNNNNNNNNNNNNNN-3' and 5'-TTGGNNNNNNNNNNNNNNNNNNN-3'. For dual gene knockout, the gRNAs were assembled as two gRNA arrays separated by a DR sequence, and complementary single-stranded oligos with four nucleotide T4 ligation overhangs were ordered as before from Sigma-Aldrich (Supp. Table 8). The complementary sequences were then combined by incubating 10 µM of each oligo in T4 ligation buffer in a 10 µL reaction in a thermocycler using the following program: 37 °C for 30 min, 95 °C for 5 min, and ramping down at 0.1 °C/s from 95 °C to 25 °C. The AiO-Cas12a plasmid was digested with BsmBI-v2 (NEB, R0739S) overnight, followed by purification from a 1% agarose gel using NucleoSpin columns (Macherey-Nagel, 740609.250). The annealed oligos and the digested plasmid were then combined by T4 ligation. For this, the annealed oligos were diluted 1:200 with water, and 1 µL was mixed with 1 µL of 10x T4 ligation buffer and 5 U T4 DNA Ligase (Thermo Fisher Scientific, EL0011) in a 10 µL reaction. The mixture was incubated at 16 °C overnight. The next day, 2 µL were used to transform chemo-competent *DH5α*. Single colonies were then picked from the LB agar plates, and the plasmids were isolated from liquid cultures using NucleoSpin columns (Macherey-Nagel, 740588.250).

2.7.7 Cas13d single gRNAs and arrays cloning

Cas13d gRNAs targeting *CD46*, *CD47*, *CD63*, and *CD71* were selected from cas13design.nygenome.org (234). The single gRNAs were ordered as single-stranded oligos with Gibson Assembly-compatible overhangs (Sigma-Aldrich). The oligos had the following structure: 5'-ACTGGTCGGGGTTTGAAAC-$(N)_{23}$-CAAGTAAACCCCTACCAACTGGTCGGGGTTTGAAACTTTTTTTGAATTGGCCGCG-3'. Arrays targeting one gene were designed by concatenating three single gRNAs, separated by Cas13d DR36; Gibson assembly overhangs were added as before, and the arrays were synthesized by Genscript as gene fragments. The oligos were PCR amplified using the Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) according to the manufacturer's instructions. gRNA-specific primers were used to avoid nonspecific priming (Supp. Table 9). The PCR amplified DNA fragments were purified using NucleoSpin columns (Macherey-Nagel, 740609.250). The 783-Rx-hU6 plasmid was digested with BfuAI (NEB, R0701S) and purified from a 1% agarose gel. The single gRNAs and arrays were cloned into the digested 783-Rx-hU6 plasmid backbone using Gibson Assembly (327). The gRNAs and arrays were expressed from a hU6 promoter.

2.7.8 Cas13d gRNAs design and cloning for kinetics and dual knockdown experiments

The plasmids expressing *CD46*, *CD47*, *CD63*, and *CD71* arrays were used for the Cas13d knockdown kinetics experiments (Supp. Table 2). The 783-Rx-Dual plasmid was engineered to express gRNAs from mU6 and hU6 promoters for dual gene knockdown. To clone the plasmid, we removed the *sa*Cas9 and *sp*Cas9 tracrRNA sequences from the sgLenti-orthogonal vector (Addgene #105997). For this, the plasmid was digested with NheI and BfuAI overnight (NEB, R0131, R0701S), and the digested plasmid was purified from a 1% agarose gel using NucleoSpin columns (Macherey-Nagel, 740609.250). A gene fragment without the *sa*Cas9 tracrRNA was synthesized by Genscript and cloned into the plasmid cut position through Gibson Assembly. The resulting plasmid was then digested with SphI and PaqCI (NEB, R0182S, R0745S) and purified as previously described. A gene fragment lacking the *sp*Cas9 tracrRNA was synthesized by Genscript and cloned into the plasmid cut position by Gibson Assembly (Supp. Table 10). To insert Cas13d DR36 sequences and gRNA insertion sites after the mU6 and hU6 promoters, the plasmid was digested with BfuAI (NEB, R0701S). Single-stranded DNA oligos containing the gRNA insertion sites were synthesized as oPools (IDT) (Supp. Table 9). The DNA oligos were amplified using the Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) according to the manufacturer's instructions using the following primers: 5'-GCCGTCTAATGTTCAGCTAGTATGCACAGTTGATCCGTCTC-3' and 5'-

GTGTGACGTATGATCAGATCTATGCTACAGTGAACCGTCTC-3'. After amplification, the DNA fragments were purified using NucleoSpin columns (Macherey-Nagel, 740609.250) and digested with BsmBI-v2 (NEB, R0739S) overnight. The digested fragments and plasmid were annealed together using T4 Ligation (Thermo Fisher Scientific, EL0011).

Cas13d arrays targeting CD46 and CD47 were selected for the dual knockdown of both genes. The arrays were synthesized as oligo pools (IDT). They had the following structure: 5'-AGTATGCACAGTTGATCCGTCTCAAAAC-$(N)_{23}$-DR36-$(N)_{23}$-DR36-$(N)_{23}$-CAAGAGAGACGGTTCACTGTAGCA-3'. The DNA oligos were amplified using the Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) according to the manufacturer's instructions with the following primers: 5'-GCCGTCTAATGTTCAGCTAGTATGCACAGTTGATCCGTCTC-3' and 5'-GTGTGACGTATGATCAGATCTATGCTACAGTGAACCGTCTC-3'. After amplification, the DNA fragments were purified using NucleoSpin columns (Macherey-Nagel, 740609.250) and digested with BsmBI-v2 (NEB, R0739S) overnight. Oligos were then cloned into the plasmid in both positions for combinatorial knockdown of CD46 and CD47 through 2-step digestion and T4 ligation cycles. In short, the plasmid was first digested with BfuAI (NEB, R0701S) and the first position gRNA was cloned using T4 ligation. Then the plasmid was digested again with AarI (Thermo Fisher Scientific, ER1582) and the second position gRNA was cloned.

2.7.9 Lentivirus production

Lenti-X 293T cells (Takara, 632180) were seeded at 65,000 cells per cm² in 25 mL of media (DMEM, 10% FBS, 1% P/S) in a 15 cm dish and incubated overnight at 37 °C with 5% CO_2. The following day, 15 µg of sgRNA library plasmid, 6 µg of psPAX2 (Addgene #12260), 6 µg of pMD2.G (Addgene #12259), and 108 µL of Turbofect (Thermo Fisher Scientific, R0532) were mixed into 5.4 mL of serum-free DMEM (Thermo Fischer Scientific, 11995073), vortexed briefly, incubated for 20 min at room temperature, and added to the cells. At 48 h and 72 h post-transfection, the supernatant was harvested, passed through 0.45 µm filters (Millipore), and concentrated 50x using PEG8000 lentivirus concentrator solution. The concentrator solution consisted of 80g of PEG-8000 (Sigma-Aldrich, 81268-1KG), 1.4g of NaCl (Carl Roth, 3957.1) in 80 mL of MillQ water and 20 mL of 10x PBS (Bio-Rad, 1610780) (pH 7.4) in 200 mL of water. Concentrated lentivirus aliquots were stored at −20 °C.

2.7.10 Knockout and knockdown efficiency assessment with flow cytometry

K562-CasRx cells were seeded in 12-well plates at a density of 50,000 cells/mL. They were transfected with lentiviral gRNA constructs at a low MOI of 0.2 and incubated at 37 °C with 5% CO_2 for 48 h. After incubation, the cells were selected with 2 µg/mL puromycin for 3 days. After 10 to 12 days of infection, the cells were stained with the appropriate antibodies against the target genes, and the knockout/knockdown efficiency was determined via flow cytometry analysis of >10,000 cells on a BD LSRFortessa II flow cytometer.

2.7.11 Pooled CRISPR screens libraries design

ABL1, *GAB2*, *SOS1*, *NF1*, *PTPN1*, and *SPRED2* genes were selected as target genes based on the combinatorial screen results presented in Boettcher *et al.* 2018 (198). For the Cas13d combinatorial perturbation screen, U6-g1-g2 and U6-g1-U6-g2, 27 gRNAs were chosen from the cas13design.nygenome.org algorithm (234). The gRNAs were selected based on three different criteria. The first and third groups, comprising 9 gRNAs per gene, were selected based on the predicted guide score from the cas13design algorithm. For the second group of gRNAs, we chose the sequences that target the highest number of the most expressed mRNA transcripts of a gene while considering the guide score. To accomplish this, the transcript expression data in the K562 cell line as transcripts per million (TPM) from the CCLE database were summed for the transcripts of each target gene. The gRNAs with the highest sum and guide score were selected. In each step of the gRNA selection criteria, a minimum distance of 150 base pairs between target sites on the gene mRNA transcript was maintained to avoid targeting narrow regions of the mRNA transcripts. Finally, 30 non-target control gRNAs were selected from Wessels *et al.* 2020 and added to the library (Supp. Table 1). To design the U6-a1-U6-a2 library, the same-gene targeting Cas13d arrays were created from the 27 selected gRNAs by compiling three gRNAs from each group, separated by the Cas13d DR36 sequence, to form three gRNA arrays. Non-target control arrays were generated by randomly compiling three non-target gRNAs to create 10 arrays using the same design. The U6-a1-U6-a2 library was therefore 9-fold smaller than the U6-g1-U6-g1 library (Supp. Table 1, 3).

The Cas9 sgRNA library included 4 sgRNAs per gene, along with 10 non-target control sgRNAs and 10 safe-cutter sgRNAs. The sgRNA sequences were chosen from the Brunello genome-wide library (Supp. Table 4) (328).

2.7.12 Cas9 Perturb-seq screens library design

Target genes were chosen based on RNA-Seq data after the induction of HEK293ΔRAF1:ER cells with 0.5 µM 4OHT at various time points (Fig. 20). The sgRNA

library consisted of 4 sgRNAs per gene, along with 10 non-target control sgRNAs and 10 safe cutter sgRNAs. The sgRNA sequences were selected from the Brunello genome-wide library (328). 4 positive-control sgRNAs targeting the Raf-transgene were designed using CRISPick (328,329) (Supp. Table 11).

2.7.13 Cas13d libraries cloning

The U6-g1-g2 library was cloned into the 783-Rx-mU6 plasmid. The selected 27 gRNAs per gene and 30 NTC gRNAs were cloned so that only gRNA-gNTC, gNTC-gRNA, and gNTC-gNTC combinations are obtained. For this purpose, gRNA template sequences for positions 1 and 2 of the format 5'-AGTATGCACAGTTGATCCGTCTCATTGG-DR36-$(N)_{23}$-CAAGTAAACCCCTACCAACTAGAGACGGTTCACTGTAGCA-3' and 5'-AGTATGCACAGTTGATCCGTCTCAAACTGGtCGGGGTTTGAAAC-$(N)_{23}$-DR36-GCTTTAAAGAGACGGTTCACTGTAGCA-3' were designed. The final library consisted of 11,520 elements. The U6-g1-U6-g2 and U6-a1-U6-a2 libraries were cloned into the 783-Rx-Dual plasmid. To systematically determine the effect of the different promoters used for gRNA/array expression on their activity, the libraries were cloned symmetrically so that the selected 30 gRNAs per gene and 30 NTC gRNAs were cloned into both gRNA positions. gRNA and array template sequences of the format: 5'-AGTATGCACAGTTGATCCGTCTCAAAAC-$(N)_{23}$-CAAGAGAGACGGTTCACTGTAGCA-3' and 5'-AGTATGCACAGTTGATCCGTCTCAAAAC-$(N)_{23}$-DR36-$(N)_{23}$-DR36-$(N)_{23}$-CAAGAGAGACGGTTCACTGTAGCA-3' were designed. The final libraries of U6-g1-U6-g2 and U6-a1-U6-a2 consisted of 36,864 and 4,096 elements, respectively.

The oligo pools were synthesized as oligo pools (Twist Bioscience) and PCR-amplified using Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) according to the manufacturer's protocol with 0.1 ng/µL sgRNA template DNA, 1 µM forward primer (5′-GCCGTCTAATGTTCAGCTAGTATGCACAGTTGATCCGTCTC-3′), and 1 µM reverse primer (5′-GTGTGACGTATGATCAGATCTATGCTACAGTGAACCGTCTC-3′) in a total volume of 50 µL with the following cycle numbers: 1x (98 °C for 3 min), 16x (98 °C for 1 s, 64 °C for 15 s, 72 °C for 20 s), and 1x (72 °C for 5 min). PCR products were purified using NucleoSpin columns (Macherey-Nagel, 740609.250), followed by restriction digestion with BsmBI-v2 (NEB, R0739S) at 55 °C overnight. The digested fragments from the U6-g1-g1 and U6-a1-U6-a2 libraries were then run on a 2% agarose gel, followed by excising the digested band and purifying it via NucleoSpin columns (Macherey-Nagel, 740609.250). In contrast, the U6-g1-U6-g2 library had a fragment size of only 27 bp and was run on a 20% Gradient TBE gel (Thermo Fischer Scientific, EC6315BOX). At the end of the run, the gel was stained with SYBR-Gold (Thermo Fischer Scientific, S11494), and the digested fragment

was cut. The gel pieces were passed through a microcentrifuge tube pierced with an 18-G needle. The gel slurry was resuspended in 400 µL of water and incubated at 70 °C for 45 min. The gel was then removed by passing the mixture through a SpinX column (Sigma-Aldrich, CLS8162) after centrifugation at 20,000 x g for 3 min. Finally, the DNA fragments were extracted by ethanol precipitation.

In parallel, the vectors mentioned above were prepared by restriction digestion with AarI (Thermo Fisher Scientific, ER1582) at 37 °C overnight. The digestion reaction was run on a 1% agarose gel, followed by the excision of the digested band and purification using NucleoSpin columns (Macherey-Nagel, 740609.250). 500 ng of digested vectors and the amplified sgRNA library inserts were ligated at a 2:1 insert:vector ratio using T4 ligation in a 20 µL reaction at 16 °C overnight. The reaction was purified by ethanol precipitation, and the resuspended volume was transformed into MegaX DH10β (Thermo Fisher Scientific, C640003) by electroporation using 100 ng of precipitated ligated DNA per 20 µL of bacterial suspension. *E. coli* were recovered and cultured overnight in 100 mL LB (100 µg/mL ampicillin). The plasmid library was extracted using the NucleoBond Xtra Midi kit (Macherey-Nagel, 740410.5). In parallel, a fraction of the transformation reaction was plated to determine the total number of transformed clones, ensuring the coverage remained above 1,000x. For the U6-g1-U6-g2 and U6-a1-U6-a2 libraries, the cloned plasmids were then digested again with BfuA1 (NEB, R0701S) at 37 °C overnight, and the library oligo fragments were cloned in position 2 through T4 ligation as before. The libraries U6-g1-g2, U6-g1-U6-g2, and U6-a1-U6-a2 achieved final coverages of 2,700x, 2,700x, and 11,700x per library element, respectively, ensuring even representation of all library sequences and their narrow distribution (Supp. Fig. 3B-D).

2.7.14 Cas9 library cloning for Cas9 counter screen and Perturb-seq screens

The selected 20-nucleotide target-specific sgRNA sequences were cloned using Gibson Assembly into the pMB1 library vector for the Cas9 counter screen library and the pMB1-10x library vector for the Perturb-seq screens (327). sgRNA template sequences of the format: 5′-GGAGAACCACCTTGTTGG-$(N)_{20}$-GTTTAAGAGCTAAGCTGGAAAC-3′ were synthesized as oligo pools by Integrated DNA Technologies and GenScript Biotech, Inc. for each screen, respectively.

The oligo pools were PCR-amplified using Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) according to the manufacturer's protocol with 1 ng/µL sgRNA template DNA, 1 µM forward primer (5′-GGAGAACCACCTTGTTGG-3′), and 1 µM reverse primer (5′-GTTTCCAGCTTAGCTCTTAAAC-3′) in a total volume of 50 µL with the following cycle numbers: 1x (98 °C for 3 min), 16x (98 °C for 1 s, 54 °C for 15 s, 72 °C for

20 s), and 1x (72 °C for 5 min). PCR products were purified using NucleoSpin columns (Macherey-Nagel, 740609.250). The library vectors pMB1 and pMB1-10x were prepared by restriction digestion with AarI (Thermo Fisher Scientific, ER1582) at 37 °C overnight. The digestion reaction was run on a 1% agarose gel, followed by excision of the digested band and purification via NucleoSpin columns (Macherey-Nagel, 740609.250). 100 ng of digested pMB1 and 2.4 ng of the amplified sgRNA library insert were assembled using Gibson Assembly Master Mix (NEB, E2611L) in a 20 µL reaction for 30 min. The reaction was purified using P-30 buffer exchange columns (Bio-Rad, 7326250) that were equilibrated 5x with water, and the eluted volume was transformed into 20 µL of MegaX DH10β (Thermo Fisher Scientific, C640003) via electroporation. *E. coli* were recovered and cultured overnight in 100 mL of LB with 100 µg/mL ampicillin. The plasmid libraries were extracted using the NucleoBond Xtra Midi kit (Macherey-Nagel, 740410.5). In parallel, a fraction of the transformation reaction was plated to determine the total number of transformed clones. The coverage was determined to be 3,111x clones per sgRNA for the Cas9 counter screen library and 1,643x clones per sgRNA for the Perturb-seq library, ensuring even representation of all library sgRNA sequences and their narrow distribution (Fig. 22A).

The quality of the cloned Perturb-seq sgRNA library was determined by NGS using an Illumina Miniseq (See below). MAGeCK (330) was used for library alignment. The narrow distribution of sgRNA sequences was confirmed, with read counts for 96% of sgRNA sequences falling within a single order of magnitude (Supp. Fig. 3A).

2.7.15 Pooled proliferation CRISPR screens in K562

The K562-CasRx cells were transduced with lentiviral sgRNA libraries at an MOI of >0.3 and a 1,000-fold coverage. The low MOI was used to minimize the frequency of multiple infected cells, ensuring the introduction of one copy of the gRNA expression cassette per cell. The cells were then cultured in RPMI with 10% FBS and 1x Pen/Strep (Sigma-Aldrich, P0781-100ML) in a 37 °C incubator with 5% CO_2. 24 h post-transduction, the cells underwent selection with Puromycin (Carl Roth, 0240.2) (2 µg/mL) for 96 h. After selection, aliquots of cells, representing 1,000x coverage each, were centrifuged at 1,000 x g for 5 minutes, and the pellets were frozen for later analysis using NGS (see below). The cell numbers representing 1,000x coverage for the different screens were as follows: U6-g1-g2 with 11,520,000 cells, U6-g1-U6-g2 with 36,864,000 cells, U6-a1-U6-a2 with 4,096,000 cells, and the Cas9 screen with 36,000 cells. The remaining cells were diluted to a density of 100,000 cells/mL with fresh medium. The cells were divided into two fractions. One fraction was treated with imatinib, while the other represented untreated cells. An IC50 concentration of 100 nM imatinib (MedChemExpress, HY-50946) was added to the treated

cell culture plates. Imatinib was renewed at day 3 of cell splitting (IC60 = 150 nM) and day 7 (IC80 = 300 nM) following the start of the imatinib treatment. Cells for endpoint analysis were harvested on day 19. On day 19, cells representing 1,000x coverage of the libraries per sample were harvested for downstream analysis using NGS as described below. Coverage at the cell level was maintained at over 1,000x throughout the screens, and the culture was diluted with fresh medium when the cell density reached 1 million cells/mL.

2.7.16 Direct capture Perturb-seq CRISPR screens in HEK293∆RAF1:ER

HEK293∆RAF1:ER cells were transduced with a lentiviral sgRNA library at an MOI of 0.2 and 1,000x coverage in 2 replicates. The low MOI and high coverage minimized the occurrence of multiple-infected cells, ensuring that only one gene was knocked out in each cell and facilitating an even distribution of the sgRNA library. Cells were then cultured in DMEM low glucose without phenol red, supplemented with 10% FBS (Pan Biotech, P30-1502) and 1% pen/strep (Sigma-Aldrich, P0781-100ML), in a 37 °C incubator with 5% CO_2. 48 h after transduction, transduced cells were selected with 2 µg/mL puromycin (Carl Roth, 0240.2) for 96 h. After selection, the top 20% of mCherry positive cells were sorted 8 days post-infection using a BD FACSAria II flow cytometer to increase sgRNA capture efficiency by the 10x Genomics Gel Beads. A total of 3 million cells were sorted. For the Perturb-Seq screen, 500,000 of the sorted cells were reseeded in full medium in 3 wells of a 12-well plate and incubated at 37 °C with 5% CO_2. On day 10 post-infection, the sorted cells were stimulated with 0.5 µM 4OHT (Sigma-Aldrich, H7904) for 6 h, 12 h, and 18 h. After the incubation, the cells were harvested, followed by scRNA-seq using the 10x Genomics Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (Dual Index) with Feature Barcode technology for CRISPR Screening protocol (CG000316 Rev A).

The remaining 2.5 million sorted HEK293∆RAF1:ER cells from day 8 post-infection were reseeded in full medium in a 6-well plate and incubated at 37 °C with 5% CO_2. On day 18 post-infection, 0.5 µM 4OHT was added for 48 h to induce RAF1 transgene expression and apoptosis in the induced cells. After 48 h, the dead cells were detached and removed along with the medium. The living cells were reseeded in tamoxifen-containing medium and incubated at 37 °C with 5% CO_2 for 48 h more to enhance the cell selection efficiency. Aliquots of 500,000 cells from the 48 h induction time point were taken. The cells were centrifuged, and the cell pellets were frozen for later analysis via NGS.

2.7.17 Genomic DNA extraction

For the U6-g1-g2 and the U6-g1-U6-g2 screens, the cell pellets from the baseline and day 19 time point untreated and imatinib-treated samples were resuspended in 20 ml of P1 buffer (Qiagen, 19051), which contained 100 µg/ml RNase A (Sigma-Aldrich,

10109142001) and 0.5% SDS (Sigma-Aldrich, 71725-50G), followed by incubation at 37 °C for 30 minutes. Then, Proteinase K (Sigma-Aldrich, 70663-4) was added to achieve a final concentration of 100 µg/ml and incubated at 55 °C for 30 minutes. After digestion, the samples were homogenized by passing them through an 18G needle three times and then through a 22G needle three times. The homogenized samples were mixed with 20 ml of Phenol:Chloroform:Isoamyl alcohol (Thermo Fisher Scientific, 15593031), transferred to 50 ml MaXtract tubes (Qiagen, 129073), and thoroughly mixed. The samples were centrifuged at room temperature (RT) for 5 minutes at 1,500 x g. The aqueous phase was transferred to ultracentrifuge tubes and thoroughly mixed with 2 ml of 3M sodium acetate (Sigma-Aldrich, S2889-250G) and 16 ml of isopropanol (Sigma-Aldrich, 650447-1L) at room temperature before centrifugation at 15,000g for 15 minutes. The gDNA pellets were gently washed with 10 ml of 70% ethanol (Thermo Fischer Scientific, 17740239) and dried at 37 °C. The dry pellets were resuspended in water, and the gDNA concentration was adjusted to 1 µg/µl. The degree of gDNA fragmentation was assessed on a 1% agarose gel, and the gDNA was further fragmented by boiling at 95 °C until the average size ranged between 10 and 20 kb.

For the U6-a1-U6-a2 and Cas9 screens, gDNA was extracted from the baseline and day 19 time points in untreated and imatinib-treated samples, as well as from baseline and day 21 of the HEK293ΔRAF1:ER samples, using the DNeasy Blood & Tissue DNA Purification kit (Qiagen, 69506) following the manufacturer's instructions.

2.7.18 PCR recovery of gRNA/array sequences from gDNA

Two nested PCR reactions were conducted to amplify the U6-g1-g2, U6-g1-U6-g2, and the Cas9 screens gRNA/array cassette from the extracted gDNA. For the first PCR reactions, up to 50 µg of gDNA, 0.3 µM forward (5′-GGCTTGGATTTCTATAACTTCGTATAGCA-3') and reverse (5′-CGGGGACTGTGGGCGATGTG-3′) primer, 200 µM dNTP mix (Thermo Fischer Scientific, 10297018), 1x Titanium Taq buffer, and 2 µL of Titanium Taq polymerase (Takara, 639209) were combined in a total volume of 50 µL. For the U6-g1-U6g1 screen samples, the PCR reaction was performed using the following cycles: 1x (95 °C, 3 min), 16x (95 °C, 30 s, 62°C, 30 s, 68 °C, 3 min), 1x (68 °C, 5 min). For the U6-g1-g2 and the Cas9 screens, the PCR cycling conditions were 1x (94 °C, 3 min), 16x (94 °C, 30 s, 62 °C, 10 s, 72 °C, 20 s), 1x (68 °C, 2 min). For the second PCR reactions, 2 µL of the first- round PCR, 0. 5 µM forward (5'-AATGATACGGCGACCACCGAGATCTACACACACTCTTTCCCTACACGACGCTCTTCCGATCTTGAGACTATAAGTATCCCTTGGAGAACCACCTTG- 3' and 5′-AATGATACGGCGACCACCGAGATCTACACACACTCTTTCCCTACACGACGCTCTTC

CGATCTTCCCTTGGAGAACCACCTTGTTGG-3′ for the Cas13d and Cas9 screens samples, respectively) and reverse (5'-CAAGCAGAAGACGGCATACGAGA-$(N)_6$-TGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATTGCTAGGACCGGCCTTAAA GC-3' and 5′-CAAGCAGAAGACGGCATACGAGAT-$(N)_6$-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-3′ for the Cas13d and Cas9 screens samples, respectively) primers, where $(N)_6$ is a 6 nt index for sequencing on the Illumina NGS platform, 200 µM dNTP mix (Thermo Fischer Scientific, 10297018), 1x Titanium Taq buffer, and 1.5 µL of Titanium Taq (Takara, 639209). For the U6-g1-U6g1 screen samples, the PCR cycles were: 1 x (95 °C, 3 min), 12x (95 °C, 30 s, 55 °C, 30 s, 68 °C, 3 min), 1x (68 °C, 5 min). For the U6-g1-g2 and the Cas9 screens, the PCR cycling conditions were 1x (94 °C, 3 min), 20x (94 °C, 30 s, 55 °C, 10 s, 72 °C, 20 s), 1x (68 °C, 2 min). The PCR products for the U6-g1-g2, U6-g1-U6-g2, and the Cas9 screens had a size of 344 bp, 888 bp, and 325 bp, respectively, and were purified from a 1% agarose gel via NucleoSpin columns (Macherey- Nagel, 740609.250).

For the U6-a1-U6-a2 screen samples, a one-step PCR reaction was performed using ExTaq Polymerase (Takara, RR001A). For this, 10 µg gDNA, 0.5 µM forward (5′-AATGATACGGCGACCACCGAGATCTACACACACTCTTTCCCTACACGACGCTCTTC CGATCTTGAGACTATAAGTATCCCTTGGAGAACCACCTTG-3′) and reverse (5'-CAAGCAGAAGACGGCATACGAGA-$(N)_6$-TGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATTGCTAGGACCGGCCTTAAA GC-3') primers, 800 µM dNTP mix, 1x ExTaq buffer, and 1.5 µL ExTaq (Takara) were mixed. The PCR cycles were set at 1x (95 °C, 1 min), 28x (95 °C, 30 s, 66 °C, 30 s, 72 °C, 1 min), 1x (72 °C, 10 min). The PCR product had a size of 1124 bp and was purified from a 1% agarose gel via NucleoSpin columns (Macherey-Nagel, 740609.250). NGS for all amplified samples was performed on either MiniSeq using a MiniSeq Mid Output Kit (300-cycles) or NovaSeq 6000 Illumina platforms using paired-end 150 strategy.

2.7.19 Cas9 proliferation screens data analysis

Paired-end sequencing was performed using an Illumina NovaSeq 6000 platform for the Cas9 counter screen in K562 and an Illumina Miniseq for the Cas9 Perturb-seq proliferation screen in HEK293ΔRAF1:ER. Data analysis was carried out using MAGeCK 0.5.9.2 (330). In summary, sgRNA read count files were generated from the raw CRISPR FASTQ files using the count function (Supp. Data 2). The MAGeCK MLE command was then utilized to calculate the MAGeCK beta score, Wald-P values, and false discovery rates for the enrichment and depletion of each guide compared to the baseline sample. Wald-P values from the Cas9 Perturb-seq proliferation screen were adjusted using the Bonferroni correction method in R (Supp. Data 4).

2.7.20 Cas13d screens data analysis

Paired-end sequencing was performed using an Illumina NovaSeq 6000 for the U6-g1-U6-g2, U6-a1-U6-a2, and Cas9 counter screen samples, while the Illumina NovaSeq X Plus was used for the U6-g1-g2 samples.

MAGeCK was used to analyze the proliferation screens for gRNA counting. Initially, the FASTQ files from the U6-g1-U6-g2 and U6-a1-U6-a2 screen samples were filtered to exclude sequences containing 5'-GCTTTAAGGC-3' in read 1 and 5'-CCAACAAGGT-3' in read 2, in order to remove recombined constructs expressing a single gRNA (sgRNA) or single-gene array (array). Following this, cutadapt 4.4 (331) was used to remove regions spanning the first 74 bp and the last 18 bp in read 1, and the first 59 bp and the last 33 bp in read 2, resulting in the gRNA-1 and gRNA-2 sequences in U6-g1-g2 and U6-g1-U6-g2, respectively, as well as the gRNA-1 sequence from array-1 and the gRNA-3 sequence from array-2 in U6-a1-U6-a2. FLASH v1.2.11 (332) was used to merge the resulting trimmed reads. The sequence alignment to the reference sequences was performed using Bowtie 1.2.3 (333), allowing for up to three nucleotide mismatches. Conversion of the resulting sam files to bam files was completed using SAMtools 1.18 (334), which then served as input for MAGeCK. The gRNA sequences were counted from the bam files using MAGeCK COUNT (Supp. Data 2). The distribution of the libraries was then determined from the baseline samples.

Based on the count matrix, the tau values and GI scores were calculated (Supp. Data 1). To eliminate low read counts gRNA combinations within each screen that had counts below 10% of the baseline mean counts were systematically removed. Additionally, a pseudocount was applied to each count value to mitigate the occurrence of zero counts. Normalization was then performed by dividing each count by the mean value of all count numbers of the gRNAs in the corresponding sample. Tau values were then computed using the following formula:

$$\tau_x = log_2 \left(\frac{\left(\frac{N_t^x}{N_{t_0}^x} \right)}{\left(\frac{N_t^{NTC}}{N_{t_0}^{NTC}} \right)} \right) \quad (1)$$

where N^x denotes the frequency of sgRNA *x* and N^{NTC} denotes the frequency of non-targeting control gRNAs before (t_0) or after (*t)* imatinib treatment.

To normalize the non-zero gNTC*gNTC tau values in the tau values of the sgRNA and array combinations, the average tau values of the gNTC*gNTC combinations for each sample were calculated and subtracted from all tau values within that specific sample.

Afterwards, the GI score was calculated for each gRNA/Array combination using the following formula:

$$GI = \tau(_{Gene1+Gene2}) - (\tau_{Gene1} + \tau_{Gene2}) \quad (2)$$

where $\boldsymbol{\tau}_{\text{Gene1+Gene2}}$ is the measured phenotype of the double perturbation and $\boldsymbol{\tau}_{\text{Gene1}} + \boldsymbol{\tau}_{\text{Gene2}}$ is the expected phenotype calculated from the calculated individual perturbation phenotype of each gene.

To obtain stable GI scores, sgRNAs and arrays with weak phenotypes were removed. The mean value of tau scores per sgRNA/array combination across both replicates of imatinib treatment was calculated, followed by taking the mean across all sgRNA-NTC/array-NTC phenotypes for each gene per sgRNA/array. A distinction was made between the respective orientations of sgRNA/array in position 1 (sgRNA-NTC/array-NTC) or sgRNA/array in position 2 (NTC-sgRNA/NTC-array). The individually calculated mean values for each sgRNA/array were then used to determine the overall mean value per gene, which serves as a cut-off for determining weaker phenotypes. All sgRNAs/arrays that did not reach the cut-off value were identified for each gene and orientation. Finally, GI scores of sgRNAs/arrays that did not meet the cut-off value for both orientations were removed. Similar to Aregger *et al.* 2020, all remaining GI scores per gene were mean-summarized, and their significance was calculated using the limma moderated *t*-test followed by Benjamini-Hochberg multiple testing correction (Supp. Data 3) (202).

The calculation of the “same gene GI” and the “single gene controls” was performed similarly to the calculation for the GI scores described above. “Same gene GIs” refer to “self-genetic interactions” of a gene such as ABL1-ABL1. The formula for the calculation is shown below using ABL1 as an example:

$$GI = \tau(_{ABL1+ABL1}) - (\tau_{ABL1} + \tau_{ABL1}) \quad (3)$$

In contrast, “single gene controls” represent the interaction of a gene with all non-target controls, such as ABL-NTC. While the orientation of the gene and non-target controls was previously disregarded in “same gene GI”, it is considered in “single gene controls”.

$$GI = \tau(_{ABL1+NTC}) - (\tau_{ABL1} + \tau_{NTC}) \quad (4)$$

All scripts used for data analysis were written in R 4.3.0 (335). To create the plots shown in the figures the R package ggplot2 (336) and ComplexHeatmap (337) as well as GraphPad Prism 10.3.0 were used. The network was created in BioRender.

2.7.21 Perturb-seq screen data analysis

Cell Ranger (10x Genomics) Version 6.1.1 was used for scRNA-seq data analysis (https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/using/count). NGS reads from the gene expression library were mapped to the GRCh38-1.2.0 genome reference compiled by 10x Genomics for Cell Ranger. sgRNA reads were mapped to an sgRNA feature reference. Count matrices were then used as input into the Seurat R package (338) to perform downstream analyses. Differential expression was identified based on pseudo bulks using the R-library glmGamPoi version 1.10.2 (339).

2.7.22 EGR1, TCF7, and Safe-Cutter knockout in HEK293ΔRAF1:ER Cells

sgRNAs targeting EGR1 (sgEGR1-1), TCF7 (sgTCF7-2), and a safe-cutter control (sgSafe-cutter-1) were selected from the Perturb-seq library and ordered as complementary single-stranded oligonucleotides from Sigma-Aldrich and cloned into the pMB1 sgRNA expression vector, as described previously (see Section 3.8.4). HEK293ΔRAF1:ER cells were transduced with lentiviruses produced from these vectors and selected with 2 µg/mL puromycin. Following 10 days of selection, genomic DNA was extracted using the DNeasy Blood & Tissue DNA Purification Kit (Qiagen, 69506) according to the manufacturer's instructions, to assess knockout efficiency by Sanger sequencing of the targeted loci.

The genomic regions flanking the sgRNA target sites in EGR1 and TCF7 were PCR-amplified using gene-specific primers. The sequences used were: gDNA_EGR1_for: 5'-CCAGCCAAACCACTCGACTG-3', gDNA_EGR1_rev: 5'-CGAAGAGGCCACAACACTTTTG-3', gDNA_TCF7_for: 5'-GTCGAGTCACTTCCGGTGCC-3', and gDNA_TCF7_rev: 5'-ATTGCAAAGCTCTGAGCTCAGATC-3'. PCR-amplification was performed using Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, F548L) according to the manufacturer's protocol with 250 ng template genomic DNA, 1 µM forward primer, and 1 µM reverse primer in a total volume of 50 µL with the following cycle numbers: 1x (98 °C for 5 min), 30x (98 °C for 15 s, 56 °C for 20 s, 72 °C for 1 min), and 1x (72 °C for 3 min). PCR products were purified using NucleoSpin Gel and PCR Clean-up columns (Macherey-Nagel, 740609.250) and submitted for Sanger sequencing at GENEWIZ Germany GmbH.

2.7.23 Bulk RNA-Sequencing data generation and preprocessing

For transcriptional profiling of HEK293ΔRAF1:ER cells following RAF1 induction, cells were treated with 0.5 µM 4OHT (Sigma-Aldrich, H7904) for 0.5 h, 1 h, 2 h, 4 h, and 8 h. After the incubation period, the cells were harvested and the total RNA was extracted using

TRIzol (Thermo Fisher Scientific, 15596018). Sequencing libraries were prepared using Illumina TruSeq mRNA Library Prep Kit v2 and sequenced on Illumina HiSeq 2000. Raw reads were processed using the snakemake-workflows rna-seq-star-deseq2 pipeline, v1.2.0, and counts were subsequently analyzed using DESeq2 in R (340).

For bulk RNA-sequencing of TCF7 knockout HEK293ΔRAF1:ER cells after Wnt signaling activation, TCF7 knockout and wild type HEK293ΔRAF1:ER cells were treated with 10 µM CHIR99021 for 24 h. The cells were then harvested and the total RNA was extracted using the RNeasy Mini Kit (Qiagen, 74106) according to the manufacturer's instructions. The RNA was then quantified using a NanoDrop One and sent for bulk RNA-seq at Novogene UK.

2.7.24 Processing of bulk, multiplexed QuantSeq data

bcl2fastq (v2.20.0 by Illumina) was used to demultiplex and convert raw sequencing data to fastq files. A Snakemake (v7.18.2) (341) workflow was designed in which BBMap's BBDuk (v39.01) was used to trim adapters, STAR (v2.7.10b) (342) to align reads to the GENCODE GRCh38.p13 (v39) geneset, UMI-tools (v1.1.4) (343) to extract and deduplicate UMIs, and subread's featureCounts (v2.0.6) (344) to count mapped reads on the gene level.

2.7.25 RNA isolation, cDNA synthesis and RT-qPCR

For the MAPK and Wnt signaling pathway activation, HEK293ΔRAF1:ER cells were treated with 0.5 µM 4OHT for 12 h and 10 µM CHIR99021 for 24 h. To validate the EGR1-TCF7 positive feedback loop, the cells were treated with 0.5 µM 4OHT for 12 h. Finally, to examine the effect of MAPK and Wnt signaling activation on DKK1, cells were treated with 0.5 µM 4OHT for 12 h and 10 µM CHIR99021 for 6 h either separately or in combination.

Total RNA was isolated for qRT-PCR using a RNeasy Mini Kit (Qiagen, 74106) according to the manufacturer's instructions and quantified with a Nanodrop One. cDNA was synthesized from 500 ng total RNA using a RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Fischer Scientific, K1632) using random hexamer primers. Quantitative PCR (qPCR) was performed with ORA qPCR Green ROX L Mix (highQu, QPD0105) on a LightCycler 480 System (Roche) following the manufacturer's protocol with 12.5 ng template cDNA, 0.2 µM forward primer, and 0.2 µM reverse primer in a total volume of 5 µL with the following cycle numbers: 1x (95 °C for 5 min), 45x (95 °C for 10 s, 60 °C for 10 s, 72 °C for 10 s), and 1x (95 °C for 5 s, 65 °C for 1 min, ramp to 97 °C at 0.11 °C/s). The primer pairs used for each gene were as follows: GAPDH_for: CTGGTAAAGTGGATATTGTTGCCAT, GAPDH_rev: TGGAATCATATTGGAACATGTAAACC, EGR1_for: CTTCAACCCTCAGGCGGACA,

EGR1_rev: GGAAAAGCGGCCAGTATAGGT, TCF7_for: CTGACCTCTCTGGCTTCTACTC, TCF7_rev: CAGAACCTAGCATCAAGGATGGG, AXIN2_for: CAAACTTTCGCCAACCGTGGTTG, AXIN2_rev: GGTGCAAAGACATAGCCAGAACC, DKK1_for: CACACCAAAGGACAAGAAGG, and DKK1_rev: CAAGACAGACCTTCTCCACA. Analysis was conducted using the double delta Ct ($2^{-}\Delta\Delta Ct$) method to quantify the relative gene expression (345). GAPDH served as the reference housekeeping gene. The expected RT-qPCR gene product was validated using a 2% agarose gel.

2.7.26 Statistics and reproducibility

The number of technical and/or biological replicates for all experiments is listed in the Figure legends or text. Unless otherwise indicated, statistical significance was calculated using limma moderated *t*-test followed by Benjamini-Hochberg multiple testing correction. Pearson's correlation was used to determine the *r* values. Statistical analyses were performed using GraphPad Prism 9 (GraphPad Software) or the R language programming environment.

2.7.27 Data availability

Raw FASTQ files from the K562 screens have been deposited on the SRA database under the following accession code PRJNA1092399 [https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1092399]. Plasmids and their sequences are deposited at Addgene: pXR001-mCD4 (Addgene #228359), 783-Rx-hU6 (Addgene #228360), 783-Rx-mU6 (Addgene #228361), 783-Rx-Dual (Addgene #228362), pMB1 (Addgene #228363), AiO-Cas12a (Addgene#228364).

Raw and processed transcriptome data from the HEK293ΔRAF1:ER Perturb-seq and proliferation screens is available at GEO under the accession number GSE250559.

2.7.28 Code availability

Data processing scripts and raw input data for the data processing scripts are available at Zenodo [doi: https://zenodo.org/records/14160184 and doi: https://zenodo.org/records/10493550].

References

1. Unger Avila P, Padvitski T, Leote AC, Chen H, Saez-Rodriguez J, Kann M, et al. Gene regulatory networks in disease and ageing. Nat Rev Nephrol. 2024 Sep;20(9):616–33.

2. Przybyla L, Gilbert LA. A new era in functional genomics screens. Nat Rev Genet. 2022 Feb;23(2):89–103.

3. Bizuayehu HM, Dadi AF, Ahmed KY, Tegegne TK, Hassen TA, Kibret GD, et al. Burden of 30 cancers among men: Global statistics in 2022 and projections for 2050 using population-based estimates. Cancer. 2024 Nov 1;130(21):3708–23.

4. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000 Jan 7;100(1):57–70.

5. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar 4;144(5):646–74.

6. Luo J, Solimini NL, Elledge SJ. Principles of cancer therapy: oncogene and non-oncogene addiction. Cell. 2009 Mar 6;136(5):823–37.

7. Kinzler KW, Vogelstein B. Cancer-susceptibility genes. Gatekeepers and caretakers. Nature. 1997 Apr 24;386(6627):761, 763.

8. Michor F, Iwasa Y, Nowak MA. Dynamics of cancer progression. Nat Rev Cancer. 2004 Mar;4(3):197–205.

9. Ashworth A, Lord CJ, Reis-Filho JS. Genetic interactions in cancer progression and treatment. Cell. 2011 Apr 1;145(1):30–8.

10. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2010 Jun 25;141(7):1117–34.

11. Rawlings JS, Rosler KM, Harrison DA. The JAK/STAT signaling pathway. J Cell Sci. 2004 Mar 15;117(Pt 8):1281–3.

12. Weng AP, Aster JC. Multiple niches for Notch in cancer: context is everything. Curr Opin Genet Dev. 2004 Feb;14(1):48–54.

13. Clevers H, Nusse R. Wnt/β-catenin signaling and disease. Cell. 2012 Jun 8;149(6):1192–205.

14. Blume-Jensen P, Hunter T. Oncogenic kinase signalling. Nature. 2001 May 17;411(6835):355–65.

15. Cargnello M, Roux PP. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. Microbiol Mol Biol Rev. 2011 Mar;75(1):50–83.

16. Shaul YD, Seger R. The MEK/ERK cascade: from signaling specificity to diverse functions. Biochim Biophys Acta. 2007 Aug;1773(8):1213–26.

17. Lavoie H, Gagnon J, Therrien M. ERK signalling: a master regulator of cell behaviour, life and fate. Nat Rev Mol Cell Biol. 2020 Oct;21(10):607–32.

18. Kolch W. Coordinating ERK/MAPK signalling through scaffolds and inhibitors. Nat Rev Mol Cell Biol. 2005 Nov;6(11):827–37.

19. Kyriakis JM, Avruch J. Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. Physiol Rev. 2001 Apr;81(2):807–69.

20. Chang L, Karin M. Mammalian MAP kinase signalling cascades. Nature. 2001 Mar 1;410(6824):37–40.

21. Plotnikov A, Zehorai E, Procaccia S, Seger R. The MAPK cascades: signaling components, nuclear roles and mechanisms of nuclear translocation. Biochim Biophys Acta. 2011 Sep;1813(9):1619–33.

22. Wortzel I, Seger R. The ERK Cascade: Distinct Functions within Various Subcellular Organelles. Genes Cancer. 2011 Mar;2(3):195–209.

23. Bandaru P, Kondo Y, Kuriyan J. The Interdependent Activation of Son-of-Sevenless and Ras. Cold Spring Harb Perspect Med. 2019 Feb 1;9(2).

24. Avruch J, Khokhlatchev A, Kyriakis JM, Luo Z, Tzivion G, Vavvas D, et al. Ras activation of the Raf kinase: tyrosine kinase recruitment of the MAP kinase cascade. Recent Prog Horm Res. 2001;56:127–55.

25. Rushworth LK, Hindley AD, O'Neill E, Kolch W. Regulation and role of Raf-1/B-Raf heterodimerization. Mol Cell Biol. 2006 Mar;26(6):2262–72.

26. Morrison DK, Cutler RE. The complexity of Raf-1 regulation. Curr Opin Cell Biol. 1997 Apr;9(2):174–9.

27. Adachi M, Fukuda M, Nishida E. Nuclear export of MAP kinase (ERK) involves a MAP kinase kinase (MEK)-dependent active transport mechanism. J Cell Biol. 2000 Mar 6;148(5):849–56.

28. Roskoski R. ERK1/2 MAP kinases: structure, function, and regulation. Pharmacol Res. 2012 Aug;66(2):105–43.

29. Dhillon AS, Hagan S, Rath O, Kolch W. MAP kinase signalling pathways in cancer. Oncogene. 2007 May 14;26(22):3279–90.

30. Kim EK, Choi E-J. Pathological roles of MAPK signaling pathways in human diseases. Biochim Biophys Acta. 2010 Apr;1802(4):396–405.

31. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic signaling pathways in the cancer genome atlas. Cell. 2018 Apr 5;173(2):321-337.e10.

32. Maik-Rachline G, Hacohen-Lev-Ran A, Seger R. Nuclear ERK: mechanism of translocation, substrates, and role in cancer. Int J Mol Sci. 2019 Mar 8;20(5).

33. Holderfield M, Deuker MM, McCormick F, McMahon M. Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond. Nat Rev Cancer. 2014 Jul;14(7):455–67.

34. Burotto M, Chiou VL, Lee J-M, Kohn EC. The MAPK pathway across different malignancies: a new perspective. Cancer. 2014 Nov 15;120(22):3446–56.

35. Bahar ME, Kim HJ, Kim DR. Targeting the RAS/RAF/MAPK pathway for cancer therapy: from mechanism to clinical studies. Signal Transduct Target Ther. 2023 Dec 18;8(1):455.

36. De Luca A, Maiello MR, D'Alessio A, Pergameno M, Normanno N. The RAS/RAF/MEK/ERK and the PI3K/AKT signalling pathways: role in cancer pathogenesis and implications for therapeutic approaches. Expert Opin Ther Targets. 2012 Apr;16 Suppl 2:S17-27.

37. Lei Z-N, Teng Q-X, Tian Q, Chen W, Xie Y, Wu K, et al. Signaling pathways and therapeutic interventions in gastric cancer. Signal Transduct Target Ther. 2022 Oct 8;7(1):358.

38. Amarante-Mendes GP, Rana A, Datoguia TS, Hamerschlak N, Brumatti G. BCR-ABL1 Tyrosine Kinase Complex Signaling Transduction: Challenges to Overcome Resistance in Chronic Myeloid Leukemia. Pharmaceutics. 2022 Jan 17;14(1).

39. Witte ON, Dasgupta A, Baltimore D. Abelson murine leukaemia virus protein is phosphorylated in vitro to form phosphotyrosine. Nature. 1980 Feb 28;283(5750):826–31.

40. Steelman LS, Abrams SL, Whelan J, Bertrand FE, Ludwig DE, Bäsecke J, et al. Contributions of the Raf/MEK/ERK, PI3K/PTEN/Akt/mTOR and Jak/STAT pathways to leukemia. Leukemia. 2008 Apr;22(4):686–707.

41. Skorski T, Kanakaraj P, Nieborowska-Skorska M, Ratajczak MZ, Wen SC, Zon G, et al. Phosphatidylinositol-3 kinase activity is regulated by BCR/ABL and is required for the growth of Philadelphia chromosome-positive cells. Blood. 1995 Jul 15;86(2):726–36.

42. Sattler M, Mohi MG, Pride YB, Quinnan LR, Malouf NA, Podar K, et al. Critical role for Gab2 in transformation by BCR/ABL. Cancer Cell. 2002 Jun;1(5):479–92.

43. Pendergast AM, Quilliam LA, Cripe LD, Bassing CH, Dai Z, Li N, et al. BCR-ABL-induced oncogenesis is mediated by direct interaction with the SH2 domain of the GRB-2 adaptor protein. Cell. 1993 Oct 8;75(1):175–85.

44. Notari M, Neviani P, Santhanam R, Blaser BW, Chang J-S, Galietta A, et al. A MAPK/HNRPK pathway controls BCR/ABL oncogenic potential by regulating MYC mRNA translation. Blood. 2006 Mar 15;107(6):2507–16.

45. Platanias LC. Map kinase signaling pathways and hematologic malignancies. Blood. 2003 Jun 15;101(12):4667–79.

46. Kang CD, Yoo SD, Hwang BW, Kim KW, Kim DW, Kim CM, et al. The inhibition of ERK/MAPK not the activation of JNK/SAPK is primarily required to induce apoptosis in chronic myelogenous leukemic K562 cells. Leuk Res. 2000 Jun;24(6):527–34.

47. Dance M, Montagner A, Salles J-P, Yart A, Raynal P. The molecular functions of Shp2 in the Ras/Mitogen-activated protein kinase (ERK1/2) pathway. Cell Signal. 2008 Mar;20(3):453–9.

48. Montagner A, Yart A, Dance M, Perret B, Salles J-P, Raynal P. A novel role for Gab1 and SHP2 in epidermal growth factor-induced Ras activation. J Biol Chem. 2005 Feb 18;280(7):5350–60.

49. Druker BJ, Tamura S, Buchdunger E, Ohno S, Segal GM, Fanning S, et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. Nat Med. 1996 May;2(5):561–6.

50. Cereda M, Mourikis TP, Ciccarelli FD. Genetic redundancy, functional compensation, and cancer vulnerability. Trends Cancer. 2016 Apr;2(4):160–2.

51. Lavi O. Redundancy: a critical obstacle to improving cancer therapy. Cancer Res. 2015 Mar 1;75(5):808–12.

52. Sonnhammer ELL, Koonin EV. Orthology, paralogy and proposed classification for paralog subtypes. Trends Genet. 2002 Dec;18(12):619–20.

53. De Kegel B, Ryan CJ. Paralog buffering contributes to the variable essentiality of genes in cancer cell lines. PLoS Genet. 2019 Oct 25;15(10):e1008466.

54. Wang T, Yu H, Hughes NW, Liu B, Kendirli A, Klein K, et al. Gene Essentiality Profiling Reveals Gene Networks and Synthetic Lethal Interactions with Oncogenic Ras. Cell. 2017 Feb 23;168(5):890-903.e15.

55. O'Neil NJ, Bailey ML, Hieter P. Synthetic lethality and cancer. Nat Rev Genet. 2017 Oct;18(10):613–23.

56. Wright S, Dobzhansky T. Genetics of natural populations; experimental reproduction of some of the changes caused by natural selection in certain populations of Drosophila pseudoobscura. Genetics. 1946 Mar;31(2):125–56.

57. Dobzhansky T. Genetics of natural populations; recombination and variability in populations of Drosophila pseudoobscura. Genetics. 1946 May;31:269–90.

58. Lucchesi JC. Synthetic lethality and semi-lethality among functionally related mutants of Drosophila melanfgaster. Genetics. 1968 May;59(1):37–44.

59. Nijman SMB. Synthetic lethality: general principles, utility and detection using genetic screens in human cells. FEBS Lett. 2011 Jan 3;585(1):1–6.

60. Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature. 2005 Apr 14;434(7035):913–7.

61. Farmer H, McCabe N, Lord CJ, Tutt ANJ, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature. 2005 Apr 14;434(7035):917–21.

62. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. Science. 2017 Mar 17;355(6330):1152–8.

63. McLornan DP, List A, Mufti GJ. Applying synthetic lethality for the selective targeting of cancer. N Engl J Med. 2014 Oct 30;371(18):1725–35.

64. Domingo J, Baeza-Centurion P, Lehner B. The causes and consequences of genetic interactions (epistasis). Annu Rev Genomics Hum Genet. 2019 Aug 31;20:433–60.

65. Huang A, Garraway LA, Ashworth A, Weber B. Synthetic lethality as an engine for cancer drug target discovery. Nat Rev Drug Discov. 2020 Jan;19(1):23–38.

66. Collins SR, Schuldiner M, Krogan NJ, Weissman JS. A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol. 2006;7(7):R63.

67. Boucher B, Jenna S. Genetic interaction networks: better understand to better predict. Front Genet. 2013 Dec 17;4:290.

68. Horlbeck MA, Xu A, Wang M, Bennett NK, Park CY, Bogdanoff D, et al. Mapping the genetic landscape of human cells. Cell. 2018 Aug 9;174(4):953-967.e22.

69. Norman TM, Horlbeck MA, Replogle JM, Ge AY, Xu A, Jost M, et al. Exploring genetic interaction manifolds constructed from rich single-cell phenotypes. Science. 2019 Aug 23;365(6455):786–93.

70. Han K, Jeng EE, Hess GT, Morgens DW, Li A, Bassik MC. Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. Nat Biotechnol. 2017 May;35(5):463–74.

71. Costanzo M, Kuzmin E, van Leeuwen J, Mair B, Moffat J, Boone C, et al. Global Genetic Networks and the Genotype-to-Phenotype Relationship. Cell. 2019 Mar 21;177(1):85–100.

72. Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Pagé N, et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science. 2001 Dec 14;294(5550):2364–8.

73. Tong AHY, Lesage G, Bader GD, Ding H, Xu H, Xin X, et al. Global mapping of the yeast genetic interaction network. Science. 2004 Feb 6;303(5659):808–13.

74. Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, et al. The genetic landscape of a cell. Science. 2010 Jan 22;327(5964):425–31.

75. Schuldiner M, Collins SR, Thompson NJ, Denic V, Bhamidipati A, Punna T, et al. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell. 2005 Nov 4;123(3):507–19.

76. Roller DG, Axelrod M, Capaldo BJ, Jensen K, Mackey A, Weber MJ, et al. Synthetic lethal screening with small-molecule inhibitors provides a pathway to rational combination therapies for melanoma. Mol Cancer Ther. 2012 Nov;11(11):2505–15.

77. Ji Z, Mei FC, Lory PL, Gilbertson SR, Chen Y, Cheng X. Chemical genetic screening of KRAS-based synthetic lethal inhibitors for pancreatic cancer. Front Biosci (Landmark Ed). 2009 Jan 1;14:2904–10.

78. Mohr S, Bakal C, Perrimon N. Genomic screening with RNAi: results and challenges. Annu Rev Biochem. 2010;79:37–64.

79. Boettcher M, McManus MT. Choosing the right tool for the job: rnai, TALEN, or CRISPR. Mol Cell. 2015 May 21;58(4):575–85.

80. Silva JM, Marran K, Parker JS, Silva J, Golding M, Schlabach MR, et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. Science. 2008 Feb 1;319(5863):617–20.

81. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. 2009 Nov 5;462(7269):108–12.

82. Lord CJ, Quinn N, Ryan CJ. Integrative analysis of large-scale loss-of-function screens identifies robust cancer-associated genetic interactions. eLife. 2020 May 28;9.

83. Kaelin WG. Molecular biology. Use and abuse of RNAi to study mammalian gene function. Science. 2012 Jul 27;337(6093):421–2.

84. Barrangou R, Birmingham A, Wiemann S, Beijersbergen RL, Hornung V, Smith A van B. Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference. Nucleic Acids Res. 2015 Apr 20;43(7):3407–19.

85. Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, et al. Defining a cancer dependency map. Cell. 2017 Jul 27;170(3):564-576.e16.

86. Evers B, Jastrzebski K, Heijmans JPM, Grernrum W, Beijersbergen RL, Bernards R. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. Nat Biotechnol. 2016 Jun;34(6):631–3.

87. Wood AJ, Lo T-W, Zeitler B, Pickle CS, Ralston EJ, Lee AH, et al. Targeted genome editing across species using ZFNs and TALENs. Science. 2011 Jul 15;333(6040):307.

88. Carroll D. Genome engineering with zinc-finger nucleases. Genetics. 2011 Aug;188(4):773–82.

89. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009 Dec 11;326(5959):1509–12.

90. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol. 2013 Jan;14(1):49–55.

91. Stoddard BL. Homing endonuclease structure and function. Q Rev Biophys. 2005 Feb;38(1):49–95.

92. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007 Mar 23;315(5819):1709–12.

93. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012 Aug 17;337(6096):816–21.

94. Garneau JE, Dupuis M-È, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 2010 Nov 4;468(7320):67–71.

95. Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu Rev Genet. 2011;45:273–97.

96. Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat Rev Genet. 2010 Mar;11(3):181–90.

97. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. Science. 2010 Jan 8;327(5962):167–70.

98. Deveau H, Garneau JE, Moineau S. CRISPR/Cas system and its role in phage-bacteria interactions. Annu Rev Microbiol. 2010;64:475–93.

99. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology (Reading, Engl). 2009 Mar;155(3):733–40.

100. van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJJ. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem Sci. 2009 Aug;34(8):401–7.

101. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008 Aug 15;321(5891):960–4.

102. Carte J, Wang R, Li H, Terns RM, Terns MP. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev. 2008 Dec 15;22(24):3489–96.

103. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013 Feb 15;339(6121):819–23.

104. Cho SW, Kim S, Kim JM, Kim J-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol. 2013 Mar;31(3):230–2.

105. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol. 2013 Mar;31(3):233–9.

106. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. eLife. 2013 Jan 29;2:e00471.

107. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013 Feb 15;339(6121):823–6.

108. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 2013 May 9;153(4):910–8.

109. Mullard A. CRISPR pioneers win Nobel prize. Nat Rev Drug Discov. 2020 Dec;19(12):827.

110. Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature. 2014 Sep 25;513(7519):569–73.

111. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014 Nov 28;346(6213):1258096.

112. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014 Feb 27;156(5):935–49.

113. Xue C, Greene EC. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. Trends Genet. 2021 Jul;37(7):639–56.

114. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. 2014 Jan 3;343(6166):84–7.

115. Shalem O, Sanjana NE, Zhang F. High-throughput functional genomics using CRISPR-Cas9. Nat Rev Genet. 2015 May;16(5):299–311.

116. Smith I, Greenside PG, Natoli T, Lahr DL, Wadden D, Tirosh I, et al. Evaluation of RNAi and CRISPR technologies by large-scale gene expression profiling in the Connectivity Map. PLoS Biol. 2017 Nov 30;15(11):e2003213.

117. Kampmann M, Horlbeck MA, Chen Y, Tsai JC, Bassik MC, Gilbert LA, et al. Next-generation libraries for robust RNA interference-based genome-wide screens. Proc Natl Acad Sci USA. 2015 Jun 30;112(26):E3384-91.

118. Morgens DW, Deans RM, Li A, Bassik MC. Systematic comparison of CRISPR/Cas9 and RNAi screens for essential genes. Nat Biotechnol. 2016 Jun;34(6):634–6.

119. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011 Mar 31;471(7340):602–7.

120. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science. 2014 Mar 14;343(6176):1247997.

121. Nishimasu H, Cong L, Yan WX, Ran FA, Zetsche B, Li Y, et al. Crystal Structure of Staphylococcus aureus Cas9. Cell. 2015 Aug 27;162(5):1113–26.

122. Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol. 2011 Jun;9(6):467–77.

123. Kovalev MA, Davletshin AI, Karpov DS. Engineering Cas9: next generation of genomic editors. Appl Microbiol Biotechnol. 2024 Feb 14;108(1):209.

124. Truong D-JJ, Kühner K, Kühn R, Werfel S, Engelhardt S, Wurst W, et al. Development of an intein-mediated split-Cas9 system for gene therapy. Nucleic Acids Res. 2015 Jul 27;43(13):6450–8.

125. Zetsche B, Volz SE, Zhang F. A split-Cas9 architecture for inducible genome editing and transcription modulation. Nat Biotechnol. 2015 Feb;33(2):139–42.

126. Ibraheim R, Tai PWL, Mir A, Javeed N, Wang J, Rodríguez TC, et al. Self-inactivating, all-in-one AAV vectors for precision Cas9 genome editing via homology-directed repair in vivo. Nat Commun. 2021 Nov 1;12(1):6267.

127. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013 Feb 28;152(5):1173–83.

128. Alerasool N, Segal D, Lee H, Taipale M. An efficient KRAB domain for CRISPRi applications in human cells. Nat Methods. 2020 Nov;17(11):1093–6.

129. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol. 2015 May;33(5):510–7.

130. Kearns NA, Pham H, Tabak B, Genga RM, Silverstein NJ, Garber M, et al. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nat Methods. 2015 May;12(5):401–3.

131. Yeo NC, Chavez A, Lance-Byrne A, Chan Y, Menn D, Milanova D, et al. An enhanced CRISPR repressor for targeted mammalian gene regulation. Nat Methods. 2018 Aug;15(8):611–6.

132. Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat Methods. 2013 Oct;10(10):973–6.

133. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013 Jul 18;154(2):442–51.

134. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, et al. Genome-scale CRISPR-mediated control of gene repression and activation. Cell. 2014 Oct 23;159(3):647–61.

135. Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, et al. Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature. 2015 Aug 13;524(7564):230–3.

136. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell. 2014 Oct 23;159(3):635–46.

137. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, P R Iyer E, et al. Highly efficient Cas9-mediated transcriptional programming. Nat Methods. 2015 Apr;12(4):326–8.

138. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature. 2015 Jan 29;517(7536):583–8.

139. Gupta R, Ghosh A, Chakravarti R, Singh R, Ravichandiran V, Swarnakar S, et al. Cas13d: A new molecular scissor for transcriptome engineering. Front Cell Dev Biol. 2022 Mar 31;10:866800.

140. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015 Oct 22;163(3):759–71.

141. Kim HK, Song M, Lee J, Menon AV, Jung S, Kang Y-M, et al. In vivo high-throughput profiling of CRISPR-Cpf1 activity. Nat Methods. 2017 Feb;14(2):153–9.

142. Kim HK, Min S, Song M, Jung S, Choi JW, Kim Y, et al. Deep learning improves prediction of CRISPR-Cpf1 guide RNA activity. Nat Biotechnol. 2018 Mar;36(3):239–41.

143. Kleinstiver BP, Sousa AA, Walton RT, Tak YE, Hsu JY, Clement K, et al. Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. Nat Biotechnol. 2019 Mar;37(3):276–82.

144. Tak YE, Kleinstiver BP, Nuñez JK, Hsu JY, Horng JE, Gong J, et al. Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. Nat Methods. 2017 Dec;14(12):1163–6.

145. Liu Y, Han J, Chen Z, Wu H, Dong H, Nie G. Engineering cell signaling using tunable CRISPR-Cpf1-based transcription factors. Nat Commun. 2017 Dec 13;8(1):2095.

146. Li X, Wang Y, Liu Y, Yang B, Wang X, Wei J, et al. Base editing with a Cpf1-cytidine deaminase fusion. Nat Biotechnol. 2018 Apr;36(4):324–7.

147. Kellner MJ, Koob JG, Gootenberg JS, Abudayyeh OO, Zhang F. SHERLOCK: nucleic acid detection with CRISPR nucleases. Nat Protoc. 2019 Oct;14(10):2986–3012.

148. Chen JS, Ma E, Harrington LB, Da Costa M, Tian X, Palefsky JM, et al. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. Science. 2018 Apr 27;360(6387):436–9.

149. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DBT, et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science. 2016 Aug 5;353(6299):aaf5573.

150. Konermann S, Lotfy P, Brideau NJ, Oki J, Shokhirev MN, Hsu PD. Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. Cell. 2018 Apr 19;173(3):665-676.e14.

151. Zhang C, Konermann S, Brideau NJ, Lotfy P, Wu X, Novick SJ, et al. Structural Basis for the RNA-Guided Ribonuclease Activity of CRISPR-Cas13d. Cell. 2018 Sep 20;175(1):212-223.e17.

152. Wessels H-H, Méndez-Mancilla A, Hao Y, Papalexi E, Mauck WM, Lu L, et al. Efficient combinatorial targeting of RNA transcripts in single cells with Cas13 RNA Perturb-seq. Nat Methods. 2023 Jan;20(1):86–94.

153. Iorio MV, Visone R, Di Leva G, Donati V, Petrocca F, Casalini P, et al. MicroRNA signatures in human ovarian cancer. Cancer Res. 2007 Sep 15;67(18):8699–707.

154. Alvarez-Garcia I, Miska EA. MicroRNA functions in animal development and human disease. Development. 2005 Nov;132(21):4653–62.

155. Schmitz SU, Grote P, Herrmann BG. Mechanisms of long noncoding RNA function in development and disease. Cell Mol Life Sci. 2016 Jul;73(13):2491–509.

156. Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, et al. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science. 2017 Jan 6;355(6320).

157. Bester AC, Lee JD, Chavez A, Lee Y-R, Nachmani D, Vora S, et al. An Integrated Genome-wide CRISPRa Approach to Functionalize lncRNAs in Drug Resistance. Cell. 2018 Apr 19;173(3):649-664.e20.

158. Ai Y, Liang D, Wilusz JE. CRISPR/Cas13 effectors have differing extents of off-target effects that limit their utility in eukaryotic cells. Nucleic Acids Res. 2022 Jun 24;50(11):e65.

159. Shi P, Murphy MR, Aparicio AO, Kesner JS, Fang Z, Chen Z, et al. Collateral activity of the CRISPR/RfxCas13d system in human cells. Commun Biol. 2023 Mar 28;6(1):334.

160. Wei J, Lotfy P, Faizi K, Baungaard S, Gibson E, Wang E, et al. Deep learning and CRISPR-Cas13d ortholog discovery for optimized RNA targeting. Cell Syst. 2023 Dec 20;14(12):1087-1102.e13.

161. Wan Y, Huang C, Feng D, Wang L, Lin X, Zhao X, et al. Characterizing the collateral activity of CRISPR/Cas13 in mammalian cells: Implications for RNA editing and therapeutic applications. Int J Biol Macromol. 2024 Dec;283(Pt 4):137861.

162. Kelley CP, Haerle MC, Wang ET. Negative autoregulation mitigates collateral RNase activity of repeat-targeting CRISPR-Cas13d in mammalian cells. Cell Rep. 2022 Aug 16;40(7):111226.

163. Tong H, Huang J, Xiao Q, He B, Dong X, Liu Y, et al. High-fidelity Cas13 variants for targeted RNA degradation with minimal collateral effects. Nat Biotechnol. 2023 Jan;41(1):108–19.

164. Li Y, Xu J, Guo X, Li Z, Cao L, Liu S, et al. The collateral activity of RfxCas13d can induce lethality in a RfxCas13d knock-in mouse model. Genome Biol. 2023 Feb 1;24(1):20.

165. Tieu V, Sotillo E, Bjelajac JR, Chen C, Malipatlolla M, Guerrero JA, et al. A versatile CRISPR-Cas13d platform for multiplexed transcriptomic regulation and metabolic engineering in primary human T cells. Cell. 2024 Feb 29;187(5):1278-1295.e20.

166. Wessels H-H, Stirn A, Méndez-Mancilla A, Kim EJ, Hart SK, Knowles DA, et al. Prediction of on-target and off-target activity of CRISPR-Cas13d guide RNAs using deep learning. Nat Biotechnol. 2024 Apr;42(4):628–37.

167. Guo X, Rahman JA, Wessels H-H, Méndez-Mancilla A, Haro D, Chen X, et al. Transcriptome-wide Cas13 guide RNA design for model organisms and viral RNA pathogens. Cell Genomics. 2021 Oct 13;1(1).

168. Kushawah G, Hernandez-Huertas L, Abugattas-Nuñez Del Prado J, Martinez-Morales JR, DeVore ML, Hassan H, et al. CRISPR-Cas13d Induces Efficient mRNA Knockdown in Animal Embryos. Dev Cell. 2020 Sep 28;54(6):805-817.e7.

169. Buchman AB, Brogan DJ, Sun R, Yang T, Hsu PD, Akbari OS. Programmable RNA targeting using casrx in flies. The CRISPR Journal. 2020 Jun;3(3):164–76.

170. He B, Peng W, Huang J, Zhang H, Zhou Y, Yang X, et al. Modulation of metabolic functions through Cas13d-mediated gene knockdown in liver. Protein Cell. 2020 Jul;11(7):518–24.

171. Jiang W, Li H, Liu X, Zhang J, Zhang W, Li T, et al. Precise and efficient silencing of mutant KrasG12D by CRISPR-CasRx controls pancreatic cancer progression. Theranostics. 2020 Sep 16;10(25):11507–19.

172. Zhuang C, Liu Y, Fu S, Yuan C, Luo J, Huang X, et al. Silencing of lncRNA MIR497HG via CRISPR/Cas13d Induces Bladder Cancer Progression Through Promoting the Crosstalk Between Hippo/Yap and TGF-β/Smad Signaling. Front Mol Biosci. 2020 Dec 9;7:616768.

173. Zhuang C, Zhuang C, Zhou Q, Huang X, Gui Y, Lai Y, et al. Engineered CRISPR/Cas13d Sensing hTERT Selectively Inhibits the Progression of Bladder Cancer In Vitro. Front Mol Biosci. 2021 Mar 19;8:646412.

174. Du M, Jillette N, Zhu JJ, Li S, Cheng AW. CRISPR artificial splicing factors. Nat Commun. 2020 Jun 12;11(1):2973.

175. Bock C, Datlinger P, Chardon F, Coelho MA, Dong MB, Lawson KA, et al. High-content CRISPR screening. Nat Rev Methods Primers. 2022 Feb 10;2(1).

176. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science. 2014 Jan 3;343(6166):80–4.

177. Koike-Yusa H, Li Y, Tan E-P, Velasco-Herrera MDC, Yusa K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol. 2014 Mar;32(3):267–73.

178. Doench JG. Am I ready for CRISPR? A user's guide to genetic screens. Nat Rev Genet. 2018 Feb;19(2):67–80.

179. Hart T, Chandrashekhar M, Aregger M, Steinhart Z, Brown KR, MacLeod G, et al. High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell. 2015 Dec 3;163(6):1515–26.

180. Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, et al. Identification and characterization of essential genes in the human genome. Science. 2015 Nov 27;350(6264):1096–101.

181. Sharma S, Petsalaki E. Application of CRISPR-Cas9 Based Genome-Wide Screening Approaches to Study Cellular Signalling Mechanisms. Int J Mol Sci. 2018 Mar 21;19(4).

182. Lopes R, Sprouffske K, Sheng C, Uijttewaal ECH, Wesdorp AE, Dahinden J, et al. Systematic dissection of transcriptional regulatory networks by genome-scale and single-cell CRISPR screens. Sci Adv. 2021 Jul 2;7(27).

183. Dixon G, Pan H, Yang D, Rosen BP, Jashari T, Verma N, et al. QSER1 protects DNA methylation valleys from de novo methylation. Science. 2021 Apr 9;372(6538).

184. Breslow DK, Hoogendoorn S, Kopp AR, Morgens DW, Vu BK, Kennedy MC, et al. A CRISPR-based screen for Hedgehog signaling provides insights into ciliary function and ciliopathies. Nat Genet. 2018 Mar;50(3):460–71.

185. Pusapati GV, Kong JH, Patel BB, Krishnan A, Sagner A, Kinnebrew M, et al. CRISPR Screens Uncover Genes that Regulate Target Cell Sensitivity to the Morphogen Sonic Hedgehog. Dev Cell. 2018 Jan 8;44(1):113-129.e8.

186. Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. Cell. 2015 Mar 12;160(6):1246–60.

187. Li QV, Dixon G, Verma N, Rosen BP, Gordillo M, Luo R, et al. Genome-scale screens identify JNK-JUN signaling as a barrier for pluripotency exit and endoderm differentiation. Nat Genet. 2019 Jun;51(6):999–1010.

188. Liu Y, Yu C, Daley TP, Wang F, Cao WS, Bhate S, et al. CRISPR Activation Screens Systematically Identify Factors that Drive Neuronal Fate and Reprogramming. Cell Stem Cell. 2018 Nov 1;23(5):758-771.e8.

189. Kuhn M, Santinha AJ, Platt RJ. Moving from in vitro to in vivo CRISPR screens. Gene and Genome Editing. 2021 May;100008.

190. Yau EH, Kummetha IR, Lichinchi G, Tang R, Zhang Y, Rana TM. Genome-Wide CRISPR Screen for Essential Cell Growth Mediators in Mutant KRAS Colorectal Cancers. Cancer Res. 2017 Nov 15;77(22):6330–9.

191. Dede M, McLaughlin M, Kim E, Hart T. Multiplex enCas12a screens detect functional buffering among paralogs otherwise masked in monogenic Cas9 knockout screens. Genome Biol. 2020 Oct 15;21(1):262.

192. Bassik MC, Kampmann M, Lebbink RJ, Wang S, Hein MY, Poser I, et al. A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. Cell. 2013 Feb 14;152(4):909–22.

193. Du D, Roguev A, Gordon DE, Chen M, Chen S-H, Shales M, et al. Genetic interaction mapping in mammalian cells using CRISPR interference. Nat Methods. 2017 Jun;14(6):577–80.

194. Roguev A, Talbot D, Negri GL, Shales M, Cagney G, Bandyopadhyay S, et al. Quantitative genetic-interaction mapping in mammalian cells. Nat Methods. 2013 May 1;10(5):432–7.

195. Shen JP, Zhao D, Sasik R, Luebeck J, Birmingham A, Bojorquez-Gomez A, et al. Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic interactions. Nat Methods. 2017 Jun;14(6):573–6.

196. Wong ASL, Choi GCG, Cui CH, Pregernig G, Milani P, Adam M, et al. Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. Proc Natl Acad Sci USA. 2016 Mar 1;113(9):2544–9.

197. DeWeirdt PC, Sanson KR, Sangree AK, Hegde M, Hanna RE, Feeley MN, et al. Optimization of AsCas12a for combinatorial genetic screens in human cells. Nat Biotechnol. 2021 Jan;39(1):94–104.

198. Boettcher M, Tian R, Blau JA, Markegard E, Wagner RT, Wu D, et al. Dual gene activation and knockout screen reveals directional dependencies in genetic networks. Nat Biotechnol. 2018 Feb;36(2):170–8.

199. Najm FJ, Strand C, Donovan KF, Hegde M, Sanson KR, Vaimberg EW, et al. Orthologous CRISPR-Cas9 enzymes for combinatorial genetic screens. Nat Biotechnol. 2018 Feb;36(2):179–89.

200. DeWeirdt PC, Sangree AK, Hanna RE, Sanson KR, Hegde M, Strand C, et al. Genetic screens in isogenic mammalian cell lines without single cell cloning. Nat Commun. 2020 Feb 6;11(1):752.

201. Dixit A, Kuksenko O, Feldman D, Regev A. Shuffle-Seq: *En masse* combinatorial encoding for *n* -way genetic interaction screens. BioRxiv. 2019 Dec 2;

202. Aregger M, Lawson KA, Billmann M, Costanzo M, Tong AHY, Chan K, et al. Systematic mapping of genetic interactions for de novo fatty acid synthesis identifies C12orf49 as a regulator of lipid metabolism. Nat Metab. 2020 Jun 1;2(6):499–513.

203. Gonatopoulos-Pournatzis T, Aregger M, Brown KR, Farhangmehr S, Braunschweig U, Ward HN, et al. Genetic interaction mapping and exon-resolution functional genomics with a hybrid Cas9-Cas12a platform. Nat Biotechnol. 2020 May;38(5):638–48.

204. Chow RD, Wang G, Ye L, Codina A, Kim HR, Shen L, et al. In vivo profiling of metastatic double knockouts through CRISPR-Cpf1 screens. Nat Methods. 2019 May;16(5):405–8.

205. Thompson NA, Ranzani M, van der Weyden L, Iyer V, Offord V, Droop A, et al. Combinatorial CRISPR screen identifies fitness effects of gene paralogues. Nat Commun. 2021 Feb 26;12(1):1302.

206. Parrish PCR, Thomas JD, Gabel AM, Kamlapurkar S, Bradley RK, Berger AH. Discovery of synthetic lethal and tumor suppressor paralog pairs in the human genome. Cell Rep. 2021 Aug 31;36(9):109597.

207. Zhou P, Chan BKC, Wan YK, Yuen CTL, Choi GCG, Li X, et al. A Three-Way Combinatorial CRISPR Screen for Analyzing Interactions among Druggable Targets. Cell Rep. 2020 Aug 11;32(6):108020.

208. Gier RA, Budinich KA, Evitt NH, Cao Z, Freilich ES, Chen Q, et al. High-performance CRISPR-Cas12a genome editing for combinatorial genetic screening. Nat Commun. 2020 Jul 13;11(1):3455.

209. Campa CC, Weisbach NR, Santinha AJ, Incarnato D, Platt RJ. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. Nat Methods. 2019 Sep;16(9):887–93.

210. Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell. 2015 Jan 15;160(1–2):339–50.

211. Ito T, Young MJ, Li R, Jain S, Wernitznig A, Krill-Burger JM, et al. Paralog knockout profiling identifies DUSP4 and DUSP6 as a digenic dependence in MAPK pathway-driven cancers. Nat Genet. 2021 Dec 2;53(12):1664–72.

212. Esmaeili Anvar N, Lin C, Ma X, Wilson LL, Steger R, Sangree AK, et al. Efficient gene knockout and genetic interaction screening using the in4mer CRISPR/Cas12a multiplex knockout platform. Nat Commun. 2024 Apr 27;15(1):3577.

213. Mani R, St Onge RP, Hartman JL, Giaever G, Roth FP. Defining genetic interaction. Proc Natl Acad Sci USA. 2008 Mar 4;105(9):3461–6.

214. Adamson B, Norman TM, Jost M, Cho MY, Nuñez JK, Chen Y, et al. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. Cell. 2016 Dec 15;167(7):1867-1882.e21.

215. Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. Cell. 2016 Dec 15;167(7):1853-1866.e17.

216. Jaitin DA, Weiner A, Yofe I, Lara-Astiaso D, Keren-Shaul H, David E, et al. Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. Cell. 2016 Dec 15;167(7):1883-1896.e15.

217. Chen G, Ning B, Shi T. Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. Front Genet. 2019 Apr 5;10:317.

218. Replogle JM, Saunders RA, Pogson AN, Hussmann JA, Lenail A, Guna A, et al. Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. Cell. 2022 Jul 7;185(14):2559-2575.e28.

219. Stuart T, Satija R. Integrative single-cell analysis. Nat Rev Genet. 2019 May;20(5):257–72.

220. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017 Jan 16;8:14049.

221. Arimbasseri AG, Rijal K, Maraia RJ. Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. Transcription. 2014 Mar;5(1):e27369.

222. Hill AJ, McFaline-Figueroa JL, Starita LM, Gasperini MJ, Matreyek KA, Packer J, et al. On the design of CRISPR-based single-cell molecular screens. Nat Methods. 2018 Apr;15(4):271–4.

223. Datlinger P, Rendeiro AF, Schmidl C, Krausgruber T, Traxler P, Klughammer J, et al. Pooled CRISPR screening with single-cell transcriptome readout. Nat Methods. 2017 Mar;14(3):297–301.

224. Zetsche B, Heidenreich M, Mohanraju P, Fedorova I, Kneppers J, DeGennaro EM, et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat Biotechnol. 2017 Jan;35(1):31–4.

225. Replogle JM, Norman TM, Xu A, Hussmann JA, Chen J, Cogan JZ, et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. Nat Biotechnol. 2020 Aug;38(8):954–61.

226. Cheng J, Lin G, Wang T, Wang Y, Guo W, Liao J, et al. Massively Parallel CRISPR-Based Genetic Perturbation Screening at Single-Cell Resolution. Adv Sci (Weinh). 2023 Feb;10(4):e2204484.

227. Choo XY, Lim YM, Katwadi K, Yap L, Tryggvason K, Sun AX, et al. Evaluating Capture Sequence Performance for Single-Cell CRISPR Activation Experiments. ACS Synth Biol. 2021 Mar 19;10(3):640–5.

228. Schraivogel D, Gschwind AR, Milbank JH, Leonce DR, Jakob P, Mathur L, et al. Targeted Perturb-seq enables genome-scale genetic screens in single cells. Nat Methods. 2020 Jun 1;17(6):629–35.

229. El Kassem G, Hillmer J, Boettcher M. Evaluation of Cas13d as a tool for genetic interaction mapping. Nat Commun. 2025 Feb 14;16(1):1631.

230. El Kassem G, Sieber A, Klinger B, Uhlitz F, Steinbrecht D, van Bentum M, et al. Targeted Perturb-seq Reveals EGR1 and FOS as Key Regulators of the Transcriptional RAF-MAPK Response. BioRxiv. 2024 Jan 15;

231. Mair B, Moffat J, Boone C, Andrews BJ. Genetic interaction networks in cancer cells. Curr Opin Genet Dev. 2019 Feb;54:64–72.

232. Costanzo M, VanderSluis B, Koch EN, Baryshnikova A, Pons C, Tan G, et al. A global genetic interaction network maps a wiring diagram of cellular function. Science. 2016 Sep 23;353(6306).

233. Granados-Riveron JT, Aquino-Jarquin G. CRISPR-Cas13 Precision Transcriptome Engineering in Cancer. Cancer Res. 2018 Aug 1;78(15):4107–13.

234. Wessels H-H, Méndez-Mancilla A, Guo X, Legut M, Daniloski Z, Sanjana NE. Massively parallel Cas13 screens reveal principles for guide RNA design. Nat Biotechnol. 2020 Jun;38(6):722–7.

235. Montero JJ, Trozzo R, Sugden M, Öllinger R, Belka A, Zhigalova E, et al. Genome-scale pan-cancer interrogation of lncRNA dependencies using CasRx. Nat Methods. 2024 Apr;21(4):584–96.

236. Allen F, Crepaldi L, Alsinet C, Strong AJ, Kleshchevnikov V, De Angeli P, et al. Predicting the mutations generated by repair of Cas9-induced double-strand breaks. Nat Biotechnol. 2018 Nov 27;37:64–72.

237. Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature. 2016 Apr 28;532(7600):517–21.

238. Sattler M, Griffin JD. Molecular mechanisms of transformation by the BCR-ABL oncogene. Semin Hematol. 2003 Apr;40(2 Suppl 2):4–10.

239. Cortes J, Pavlovsky C, Saußele S. Chronic myeloid leukaemia. Lancet. 2021 Nov 20;398(10314):1914–26.

240. Bot JF, van der Oost J, Geijsen N. The double life of CRISPR-Cas13. Curr Opin Biotechnol. 2022 Dec;78:102789.

241. LaMontagne KR, Flint AJ, Franza BR, Pandergast AM, Tonks NK. Protein tyrosine phosphatase 1B antagonizes signalling by oncoprotein tyrosine kinase p210 bcr-abl in vivo. Mol Cell Biol. 1998 May;18(5):2965–75.

242. LaMontagne KR, Hannon G, Tonks NK. Protein tyrosine phosphatase PTP1B suppresses p210 bcr-abl-induced transformation of rat-1 fibroblasts and promotes differentiation of K562 cells. Proc Natl Acad Sci USA. 1998 Nov 24;95(24):14094–9.

243. Tang Y, Liao S, Liu G, Xiong X, Liu H, Li F, et al. Advanced single-cell pooled CRISPR screening identifies C19orf53 required for cell proliferation based on mTORC1 regulators. Cell Biol Toxicol. 2022 Feb;38(1):43–68.

244. Mattick JS, Amaral PP, Carninci P, Carpenter S, Chang HY, Chen L-L, et al. Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat Rev Mol Cell Biol. 2023 Jun;24(6):430–47.

245. Statello L, Guo C-J, Chen L-L, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol. 2021 Jan 8;22(2):96–118.

246. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat Protoc. 2013 Nov;8(11):2180–96.

247. Handelmann CR, Tsompana M, Samudrala R, Buck MJ. The impact of nucleosome structure on CRISPR/Cas9 fidelity. Nucleic Acids Res. 2023 Mar 21;51(5):2333–44.

248. Tycko J, Wainberg M, Marinov GK, Ursu O, Hess GT, Ego BK, et al. Mitigation of off-target toxicity in CRISPR-Cas9 screens for essential non-coding elements. Nat Commun. 2019 Sep 6;10(1):4063.

249. Haapaniemi E, Botla S, Persson J, Schmierer B, Taipale J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. Nat Med. 2018 Jul;24(7):927–30.

250. Ihry RJ, Worringer KA, Salick MR, Frias E, Ho D, Theriault K, et al. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. Nat Med. 2018 Jul;24(7):939–46.

251. Aguirre AJ, Meyers RM, Weir BA, Vazquez F, Zhang C-Z, Ben-David U, et al. Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting. Cancer Discov. 2016 Aug;6(8):914–29.

252. Munoz DM, Cassiani PJ, Li L, Billy E, Korn JM, Jones MD, et al. CRISPR Screens Provide a Comprehensive Assessment of Cancer Vulnerabilities but Generate False-Positive Hits for Highly Amplified Genomic Regions. Cancer Discov. 2016 Aug;6(8):900–13.

253. Wang Q, Liu X, Zhou J, Yang C, Wang G, Tan Y, et al. The CRISPR-Cas13a Gene-Editing System Induces Collateral Cleavage of RNA in Glioma Cells. Adv Sci (Weinh). 2019 Oct 16;6(20):1901299.

254. Powell JE, Lim CKW, Krishnan R, McCallister TX, Saporito-Magriña C, Zeballos MA, et al. Targeted gene silencing in the nervous system with CRISPR-Cas13. Sci Adv. 2022 Jan 21;8(3):eabk2485.

255. da Silva Pescador G, Baia Amaral D, Varberg JM, Zhang Y, Hao Y, Florens L, et al. Protein profiling of zebrafish embryos unmasks regulatory layers during early embryogenesis. Cell Rep. 2024 Oct 22;43(10):114769.

256. Liu H, Zhao X-F, Lu Y-N, Hayes LR, Wang J. CRISPR/Cas13d targeting suppresses repeat-associated non-AUG translation of C9orf72 hexanucleotide repeat RNA. J Clin Invest. 2024 Sep 17;134(21).

257. Gruber C, Krautner L, Bergant V, Grass V, Ma Z, Rheinemann L, et al. Engineered, nucleocytoplasmic shuttling Cas13d enables highly efficient cytosolic RNA targeting. Cell Discov. 2024 Apr 12;10(1):42.

258. Liang W-W, Müller S, Hart SK, Wessels H-H, Méndez-Mancilla A, Sookdeo A, et al. Transcriptome-scale RNA-targeting CRISPR screens reveal essential lncRNAs in human cells. Cell. 2024 Dec 26;187(26):7637-7654.e29.

259. Li JD, Taipale M, Blencowe BJ. Efficient, specific, and combinatorial control of endogenous exon splicing with dCasRx-RBM25. Mol Cell. 2024 Jul 11;84(13):2573-2589.e5.

260. Kuo H-C, Prupes J, Chou C-W, Finkelstein IJ. Massively parallel profiling of RNA-targeting CRISPR-Cas13d. Nat Commun. 2024 Jan 12;15(1):498.

261. Liao C, Ttofali F, Slotkowski RA, Denny SR, Cecil TD, Leenay RT, et al. Modular one-pot assembly of CRISPR arrays enables library generation and reveals factors influencing crRNA biogenesis. Nat Commun. 2019 Jul 3;10(1):2948.

262. Magnusson JP, Rios AR, Wu L, Qi LS. Enhanced Cas12a multi-gene regulation using a CRISPR array separator. eLife. 2021 Sep 9;10.

263. Zoephel J, Randau L. RNA-Seq analyses reveal CRISPR RNA processing and regulation patterns. Biochem Soc Trans. 2013 Dec;41(6):1459–63.

264. Liao C, Slotkowski RA, Achmedov T, Beisel CL. The Francisella novicida Cas12a is sensitive to the structure downstream of the terminal repeat in CRISPR arrays. RNA Biol. 2019 Apr;16(4):404–12.

265. Cetin R, Wegner M, Luwisch L, Saud S, Achmedov T, Süsser S, et al. Optimized metrics for orthogonal combinatorial CRISPR screens. Sci Rep. 2023 May 6;13(1):7405.

266. Li R, Klingbeil O, Monducci D, Young MJ, Rodriguez DJ, Bayyat Z, et al. Comparative optimization of combinatorial CRISPR screens. Nat Commun. 2022 May 5;13(1):2469.

267. Cheng X, Li Z, Shan R, Li Z, Wang S, Zhao W, et al. Modeling CRISPR-Cas13d on-target and off-target effects using machine learning approaches. Nat Commun. 2023 Feb 10;14(1):752.

268. Quintás-Cardama A, Cortes J. Molecular biology of bcr-abl1-positive chronic myeloid leukemia. Blood. 2009 Feb 19;113(8):1619–30.

269. Martin GA, Viskochil D, Bollag G, McCabe PC, Crosier WJ, Haubruck H, et al. The GAP-related domain of the neurofibromatosis type 1 gene product interacts with ras p21. Cell. 1990 Nov 16;63(4):843–9.

270. Kholodenko BN, Hancock JF, Kolch W. Signalling ballet in space and time. Nat Rev Mol Cell Biol. 2010 Jun;11(6):414–26.

271. Oda K, Matsuoka Y, Funahashi A, Kitano H. A comprehensive pathway map of epidermal growth factor receptor signaling. Mol Syst Biol. 2005 May 25;1:2005.0010.

272. Parikh JR, Klinger B, Xia Y, Marto JA, Blüthgen N. Discovering causal signaling pathways through gene-expression patterns. Nucleic Acids Res. 2010 Jul;38(Web Server issue):W109-17.

273. Hibshman PS, Der CJ. The RAS signaling network and cancer. In: Rauen KA, editor. The rasopathies: genetic syndromes of the RAS/MAPK pathway. Cham: Springer Nature Switzerland; 2024. p. 363–95.

274. Bugter JM, Fenderico N, Maurice MM. Mutations and mechanisms of WNT pathway tumour suppressors in cancer. Nat Rev Cancer. 2021 Jan;21(1):5–21.

275. de Bruijn I, Kundra R, Mastrogiacomo B, Tran TN, Sikina L, Mazor T, et al. Analysis and Visualization of Longitudinal Genomic and Clinical Data from the AACR Project GENIE Biopharma Collaborative in cBioPortal. Cancer Res. 2023 Dec 1;83(23):3861–7.

276. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013 Apr 2;6(269):pl1.

277. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012 May;2(5):401–4.

278. Zeller E, Hammer K, Kirschnick M, Braeuning A. Mechanisms of RAS/β-catenin interactions. Arch Toxicol. 2013 Apr;87(4):611–32.

279. Saegusa M, Hashimura M, Kuwata T, Hamano M, Watanabe J, Kawaguchi M, et al. Transcription factor Egr1 acts as an upstream regulator of beta-catenin signalling through up-regulation of TCF4 and p300 expression during trans-differentiation of endometrial carcinoma cells. J Pathol. 2008 Dec;216(4):521–32.

280. Schulze A, Lehmann K, Jefferies HB, McMahon M, Downward J. Analysis of the transcriptional program induced by Raf in epithelial cells. Genes Dev. 2001 Apr 15;15(8):981–94.

281. Weber CK, Slupsky JR, Kalmes HA, Rapp UR. Active Ras induces heterodimerization of cRaf and BRaf. Cancer Res. 2001 May 1;61(9):3595–8.

282. Cagnol S, Van Obberghen-Schilling E, Chambard JC. Prolonged activation of ERK1,2 induces FADD-independent caspase 8 activation and cell death. Apoptosis. 2006 Mar;11(3):337–46.

283. Marshall CJ. Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. Cell. 1995 Jan 27;80(2):179–85.

284. Santos SDM, Verveer PJ, Bastiaens PIH. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. Nat Cell Biol. 2007 Mar;9(3):324–30.

285. Uhlitz F, Sieber A, Wyler E, Fritsche-Guenther R, Meisig J, Landthaler M, et al. An immediate-late gene expression module decodes ERK signal duration. Mol Syst Biol. 2017 May 3;13(5):928.

286. Tullai JW, Schaffer ME, Mullenbrock S, Kasif S, Cooper GM. Identification of transcription factor binding sites upstream of human genes regulated by the phosphatidylinositol 3-kinase and MEK/ERK signaling pathways. J Biol Chem. 2004 May 7;279(19):20167–77.

287. Jürchott K, Kuban R-J, Krech T, Blüthgen N, Stein U, Walther W, et al. Identification of Y-box binding protein 1 as a core regulator of MEK/ERK pathway-dependent gene signatures in colorectal cancer cells. PLoS Genet. 2010 Dec 2;6(12):e1001231.

288. Dijkmans TF, van Hooijdonk LWA, Schouten TG, Kamphorst JT, Fitzsimons CP, Vreugdenhil E. Identification of new Nerve Growth Factor-responsive immediate-early genes. Brain Res. 2009 Jan 16;1249:19–33.

289. Saeki Y, Endo T, Ide K, Nagashima T, Yumoto N, Toyoda T, et al. Ligand-specific sequential regulation of transcription factors for differentiation of MCF-7 cells. BMC Genomics. 2009 Nov 20;10:545.

290. Stelniec-Klotz I, Legewie S, Tchernitsa O, Witzel F, Klinger B, Sers C, et al. Reverse engineering a hierarchical regulatory network downstream of oncogenic KRAS. Mol Syst Biol. 2012;8:601.

291. Herschman HR. Primary response genes induced by growth factors and tumor promoters. Annu Rev Biochem. 1991;60:281–319.

292. Tullai JW, Schaffer ME, Mullenbrock S, Sholder G, Kasif S, Cooper GM. Immediate-early and delayed primary response genes are distinct in function and genomic architecture. J Biol Chem. 2007 Aug 17;282(33):23981–95.

293. Amit I, Citri A, Shay T, Lu Y, Katz M, Zhang F, et al. A module of negative feedback regulators defines growth factor signaling. Nat Genet. 2007 Apr;39(4):503–12.

294. Legewie S, Herzel H, Westerhoff HV, Blüthgen N. Recurrent design patterns in the feedback regulation of the mammalian signalling network. Mol Syst Biol. 2008 May 6;4:190.

295. Avraham R, Yarden Y. Feedback regulation of EGFR signalling: decision making by early and delayed loops. Nat Rev Mol Cell Biol. 2011 Feb;12(2):104–17.

296. Ünal EB, Uhlitz F, Blüthgen N. A compendium of ERK targets. FEBS Lett. 2017 Sep;591(17):2607–15.

297. Samuels ML, Weber MJ, Bishop JM, McMahon M. Conditional transformation of cells and rapid activation of the mitogen-activated protein kinase cascade by an estradiol-dependent human raf-1 protein kinase. Mol Cell Biol. 1993 Oct;13(10):6241–52.

298. Kervestin S, Jacobson A. NMD: a multifaceted response to premature translational termination. Nat Rev Mol Cell Biol. 2012 Nov;13(11):700–12.

299. Ram A, Murphy D, DeCuzzi N, Patankar M, Hu J, Pargett M, et al. A guide to ERK dynamics, part 1: mechanisms and models. Biochem J. 2023 Dec 13;480(23):1887–907.

300. Blum Y, Mikelson J, Dobrzyński M, Ryu H, Jacques M-A, Jeon NL, et al. Temporal perturbation of ERK dynamics reveals network architecture of FGF2/MAPK signaling. Mol Syst Biol. 2019 Nov;15(11):e8947.

301. von Kriegsheim A, Baiocchi D, Birtwistle M, Sumpton D, Bienvenut W, Morrice N, et al. Cell fate decisions are specified by the dynamic ERK interactome. Nat Cell Biol. 2009 Dec;11(12):1458–64.

302. Vaudry D, Stork PJS, Lazarovici P, Eiden LE. Signaling pathways for PC12 cell differentiation: making the right connections. Science. 2002 May 31;296(5573):1648–9.

303. Ram A, Murphy D, DeCuzzi N, Patankar M, Hu J, Pargett M, et al. A guide to ERK dynamics, part 2: downstream decoding. Biochem J. 2023 Dec 13;480(23):1909–28.

304. Thiel G, Cibelli G. Regulation of life and death by the zinc finger transcription factor Egr-1. J Cell Physiol. 2002 Dec;193(3):287–92.

305. Baron V, Adamson ED, Calogero A, Ragona G, Mercola D. The transcription factor Egr1 is a direct regulator of multiple tumor suppressors including TGFbeta1, PTEN, p53, and fibronectin. Cancer Gene Ther. 2006 Feb;13(2):115–24.

306. Wang B, Guo H, Yu H, Chen Y, Xu H, Zhao G. The role of the transcription factor EGR1 in cancer. Front Oncol. 2021 Mar 24;11:642547.

307. Zhou C, Martinez E, Di Marcantonio D, Solanki-Patel N, Aghayev T, Peri S, et al. JUN is a key transcriptional regulator of the unfolded protein response in acute myeloid leukemia. Leukemia. 2017 May;31(5):1196–205.

308. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics. 2006 Mar 20;7 Suppl 1(Suppl 1):S7.

309. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. Science. 2002 Oct 25;298(5594):824–7.

310. Glauser DA, Schlegel W. Mechanisms of transcriptional regulation underlying temporal integration of signals. Nucleic Acids Res. 2006 Sep 22;34(18):5175–83.

311. Murphy LO, Smith S, Chen R-H, Fingar DC, Blenis J. Molecular interpretation of ERK signal duration by immediate early gene products. Nat Cell Biol. 2002 Aug;4(8):556–64.

312. Murphy LO, MacKeigan JP, Blenis J. A network of immediate early gene products propagates subtle differences in mitogen-activated protein kinase signal amplitude and duration. Mol Cell Biol. 2004 Jan;24(1):144–53.

313. Whitmarsh AJ. Regulation of gene transcription by mitogen-activated protein kinase signaling pathways. Biochim Biophys Acta. 2007 Aug;1773(8):1285–98.

314. Dessauges C, Mikelson J, Dobrzyński M, Jacques M-A, Frismantiene A, Gagliardi PA, et al. Optogenetic actuator - ERK biosensor circuits identify MAPK network nodes that shape ERK dynamics. Mol Syst Biol. 2022 Jun;18(6):e10670.

315. Genolet O, Ravid Lustig L, Schulz EG. Dissecting Molecular Phenotypes Through FACS-Based Pooled CRISPR Screens. Methods Mol Biol. 2022;2520:1–24.

316. Kim D, Rath O, Kolch W, Cho KH. A hidden oncogenic positive feedback loop caused by crosstalk between Wnt and ERK pathways. Oncogene. 2007 Jul 5;26(31):4571–9.

317. Zhao Y, Xia Q, Liu Y, Bai W, Yao Y, Ding J, et al. TCF7L2 and EGR1 synergistic activation of transcription of LCN2 via an ERK1/2-dependent pathway in esophageal squamous cell carcinoma cells. Cell Signal. 2019 Mar;55:8–16.

318. Horst D, Chen J, Morikawa T, Ogino S, Kirchner T, Shivdasani RA. Differential WNT activity in colorectal cancer confers limited tumorigenic potential and is regulated by MAPK signaling. Cancer Res. 2012 Mar 15;72(6):1547–56.

319. Biechele TL, Kulikauskas RM, Toroni RA, Lucero OM, Swift RD, James RG, et al. Wnt/β-catenin signaling and AXIN1 regulate apoptosis triggered by inhibition of the mutant kinase BRAFV600E in human melanoma. Sci Signal. 2012 Jan 10;5(206):ra3.

320. Harmston N, Lim JYS, Arqués O, Palmer HG, Petretto E, Virshup DM, et al. Widespread Repression of Gene Expression in Cancer by a Wnt/β-Catenin/MAPK Pathway. Cancer Res. 2021 Jan 15;81(2):464–75.

321. Guardavaccaro D, Clevers H. Wnt/β-catenin and MAPK signaling: allies and enemies in different battlefields. Sci Signal. 2012 Apr 10;5(219):pe15.

322. Schmidt R, Steinhart Z, Layeghi M, Freimer JW, Bueno R, Nguyen VQ, et al. CRISPR activation and interference screens decode stimulation responses in primary human T cells. Science. 2022 Feb 4;375(6580):eabj4008.

323. Roberts PJ, Der CJ. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. Oncogene. 2007 May 14;26(22):3291–310.

324. Santarpia L, Lippman SM, El-Naggar AK. Targeting the MAPK-RAS-RAF signaling pathway in cancer therapy. Expert Opin Ther Targets. 2012 Jan 12;16(1):103–19.

325. Tedesco D, Liu T, Makhanov M, Hu D, Isachenko N, Diehl P, et al. 165 A comparison of three single-cell CRISPR-Seq protocols using the 10X Genomics platform. Regular and Young Investigator Award Abstracts. BMJ Publishing Group Ltd; 2024. p. A188–A188.

326. McMahon M. Steroid receptor fusion proteins for conditional activation of Raf-MEK-ERK signaling pathway. Meth Enzymol. 2001;332:401–17.

327. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009 May;6(5):343–5.

328. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol. 2016 Feb;34(2):184–91.

329. Sanson KR, Hanna RE, Hegde M, Donovan KF, Strand C, Sullender ME, et al. Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. Nat Commun. 2018 Dec 21;9(1):5416.

330. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. 2014;15(12):554.

331. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet j. 2011 May 2;17(1):10.

332. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011 Nov 1;27(21):2957–63.

333. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009 Mar 4;10(3):R25.

334. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021 Feb 16;10(2).

335. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.

336. Wickham H. ggplot2: Elegant Graphics for Data Analysis (Use R!). 2nd ed. Cham: Springer; 2016.

337. Gu Z. Complex heatmap visualization. iMeta. 2022 Sep;1(3):e43.

338. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021 Jun 24;184(13):3573-3587.

339. Ahlmann-Eltze C, Huber W. glmGamPoi: fitting Gamma-Poisson generalized linear models on single cell count data. Bioinformatics. 2021 Apr 5;36(24):5701–2.

340. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

341. Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-Tinch CH, Sochat V, et al. Sustainable data analysis with Snakemake. F1000Res. 2021 Jan 18;10:33.

342. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013 Jan 1;29(1):15–21.

343. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. 2017 Mar;27(3):491–9.

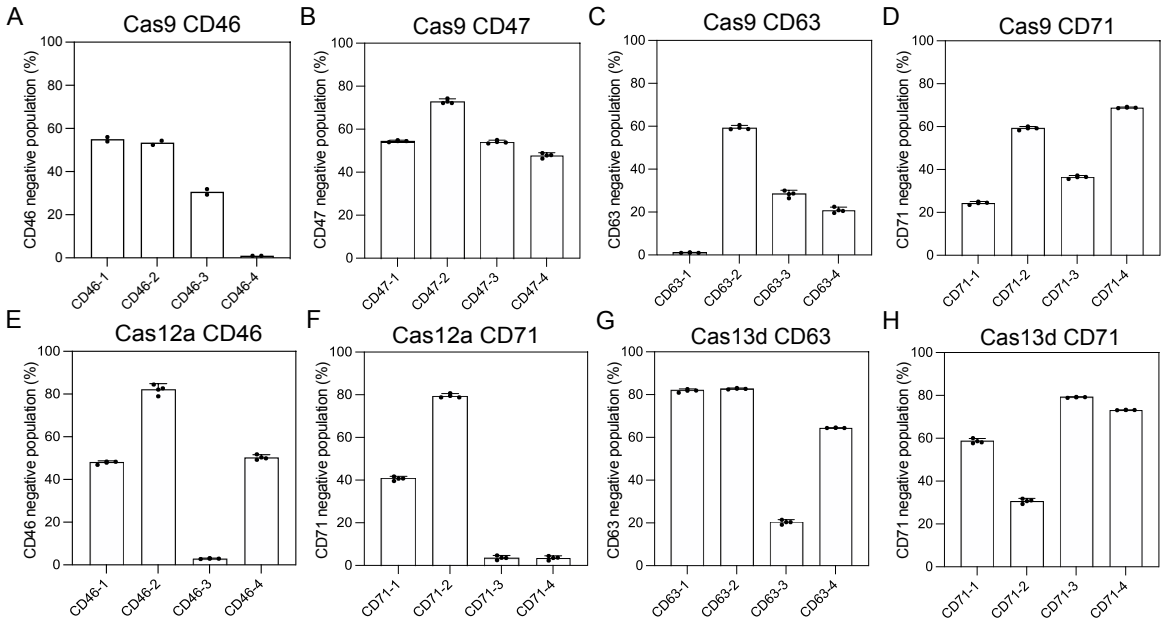
344. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014 Apr 1;30(7):923–30.

345. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001 Dec;25(4):402–8.

Appendix

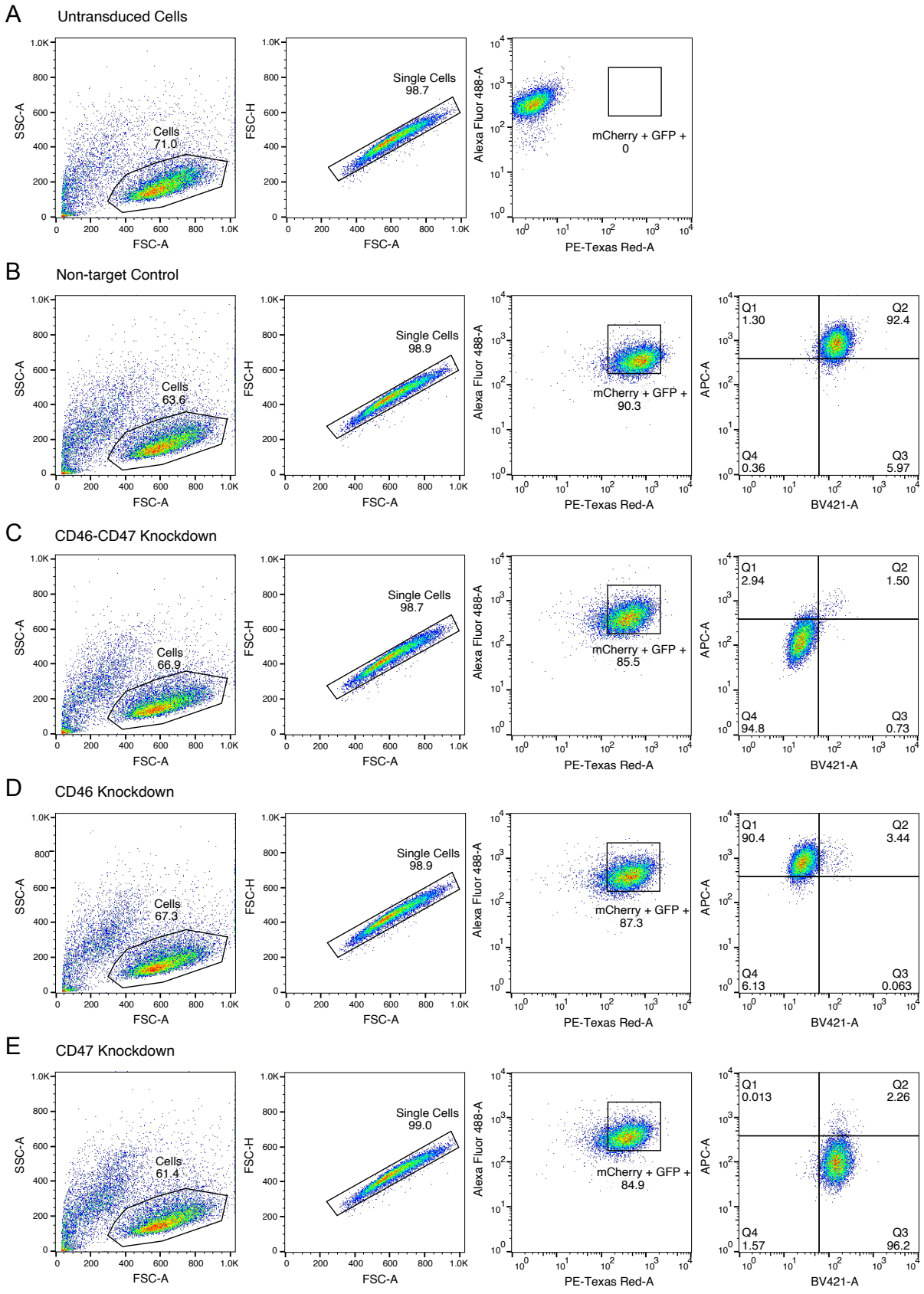
2.8 Appendix Chapter 1

2.8.1 Supplementary Figures



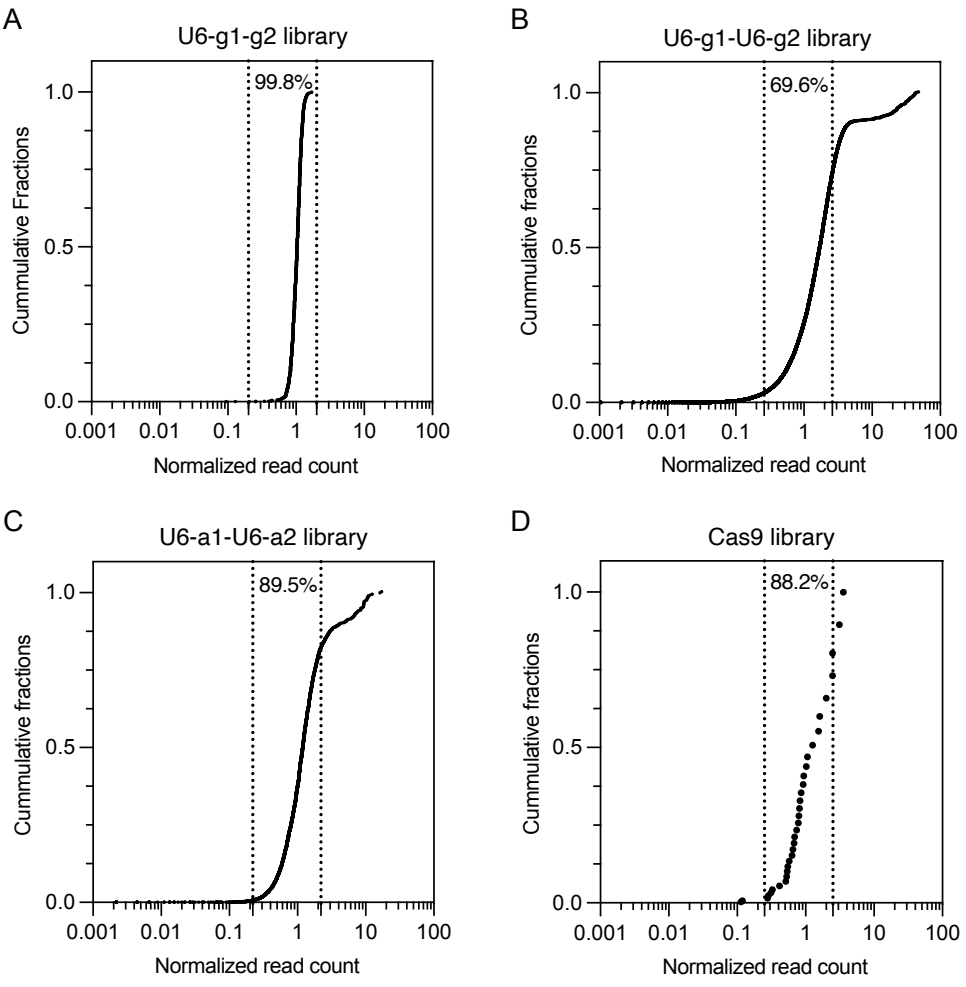
Supplementary Figure 1 - Knockout efficiency of Cas9 and Cas12a sgRNAs and Cas13d arrays per gene targeting different cellular surface markers.

(A-H) *The percentage of gene knockout cells using the indicated CRISPR system was determined via flow cytometry analysis of >10,000 cells stained with CD46-APC antibodies (Miltenyi, 130-130-362).* ***(A)*** *CD46 knockout with 4 Cas9 sgRNAs in HEK293ΔRAF1:ER cells (285).* ***(B)*** *CD47 knockout with 4 Cas9 sgRNAs in K562 cells.* ***(C)*** *CD63 knockout with 4 Cas9 sgRNAs in K562 cells.* ***(D)*** *CD71 knockout with 4 Cas9 sgRNAs in K562 cells.* ***(E)*** *CD46 knockout with 4 Cas12a gRNAs in K562 cells.* ***(F)*** *CD71 knockout with 4 Cas12a gRNAs in K562 cells.* ***(G)*** *CD63 knockout with 4 Cas13d arrays in K562 cells.* ***(H)*** *CD71 knockout with 4 Cas13d arrays in K562 cells. Values represent the mean of biological replicates; error bars show SD (n=3 except in (A) where n=2).*



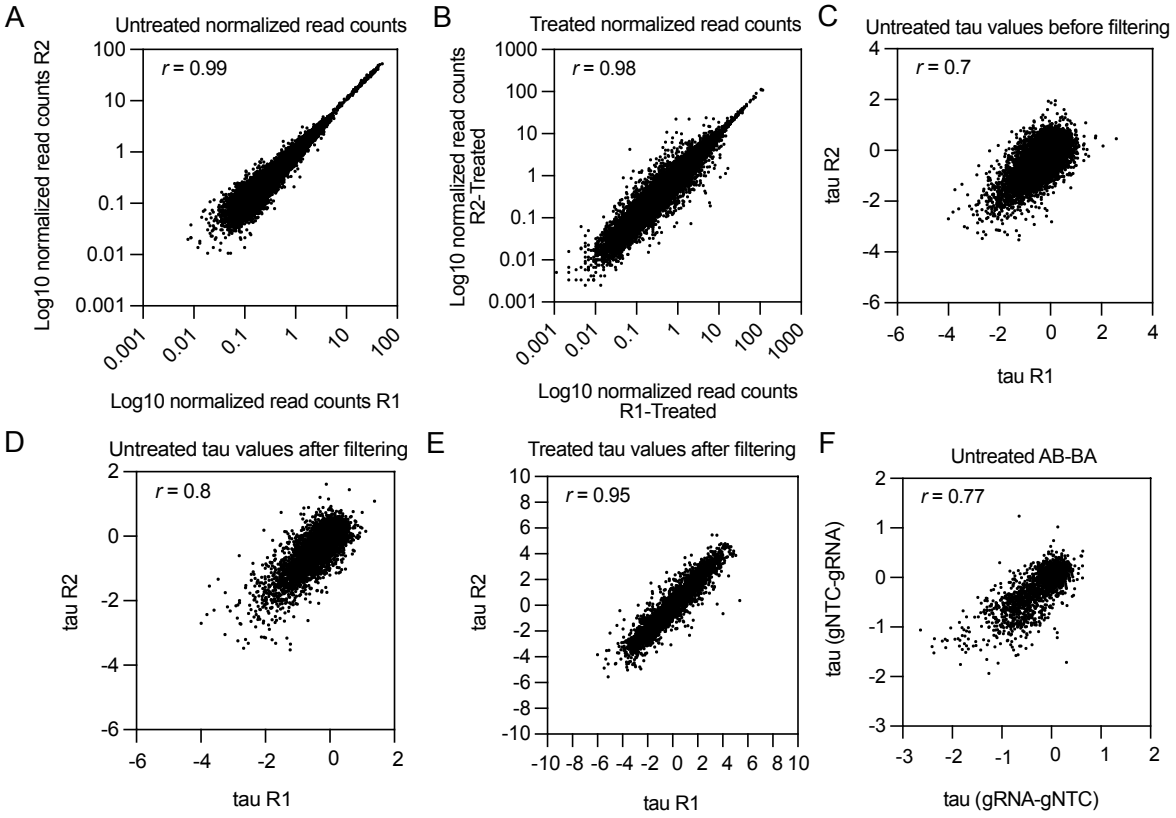
Supplementary Figure 2 - Flow cytometry gating strategy used in the comparison of single gene and double gene perturbation properties of Cas9, Cas12a, and Cas13d.

***(A**) Untransduced control cells. **(B)** Non-target control cells. **(C)** CD46-CD47 dual knockdown cells. **(D)** CD46 knockdown cells. **(E)** CD47 Knockdown cells. The gating strategy for the Cas13d double perturbation of CD47 and CD46 is illustrated here. The process begins by identifying events that represent cell populations using the FSC-Area vs SSC-Area scatter plots. Next, single cells are isolated from FSC-Area vs FSC-High scatter plots. Finally, GFP and mCherry positive cells, which represent those carrying the Cas13d-EGFP and the gRNA/Array construct, are selected. These cells are then divided into four bins representing non-perturbed, single-perturbed, and double-perturbed cells*



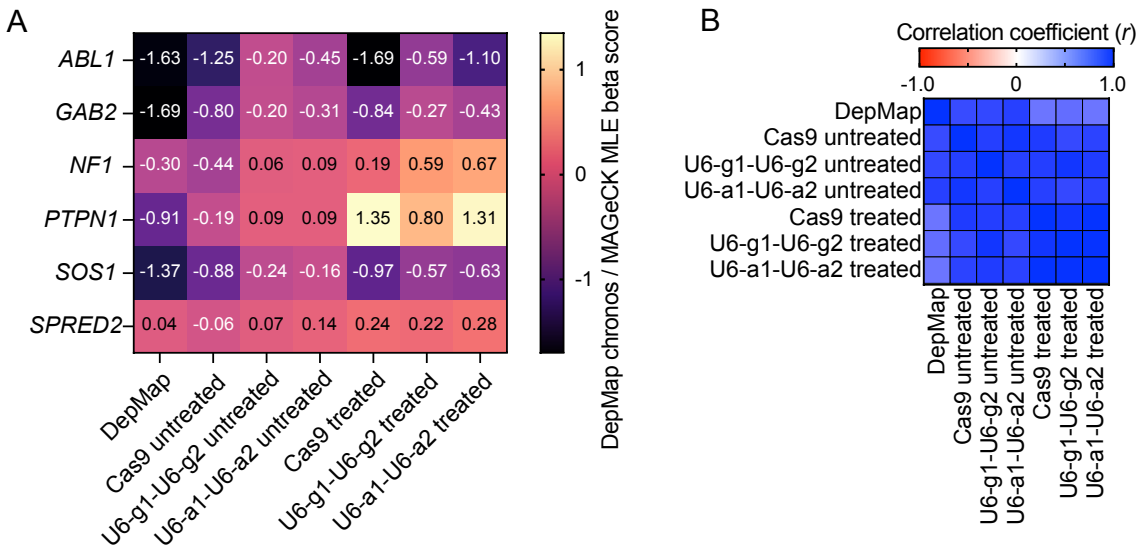
Supplementary Figure 3 - Narrow distribution of screening library elements after cloning.

***(A)** U6-g1-g2 library distribution. **(B)** U6-g1-U6-g2 library distribution. **(C)** U6-a1-U6-a2 library distribution. **(D)** Cas9 library distribution. The percentage of elements within one order of magnitude is shown between the dotted lines*.



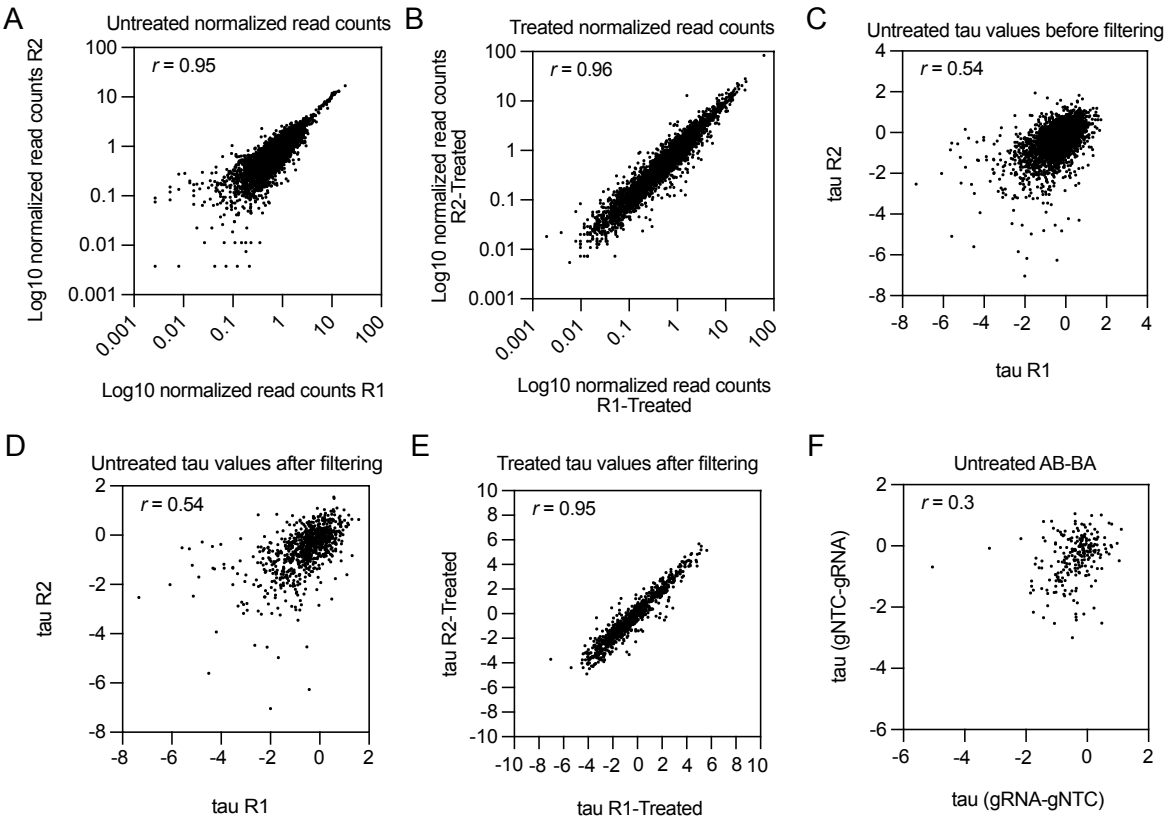
Supplementary Figure 4 - High technical reproducibility of the U6-g1-U6-g2 screens in K562 at different levels of data analysis.

(A-B) *Correlation between normalized read counts from two technical screen replicates under untreated (A) and imatinib-treated (B) conditions.* ***(C)*** *Correlation between tau values from two technical screen replicates in the untreated condition prior to filtering for functional gRNAs.* ***(D-E)*** *Correlation between tau values from two technical screen replicates in the untreated (D) and imatinib-treated (E) conditions following the filtering for functional gRNAs.* ***(F)*** *Correlation between tau values from gRNA-gNTC and gNTC-gRNA combinations in the untreated condition. Pearson's correlation was used to calculate the r values.*



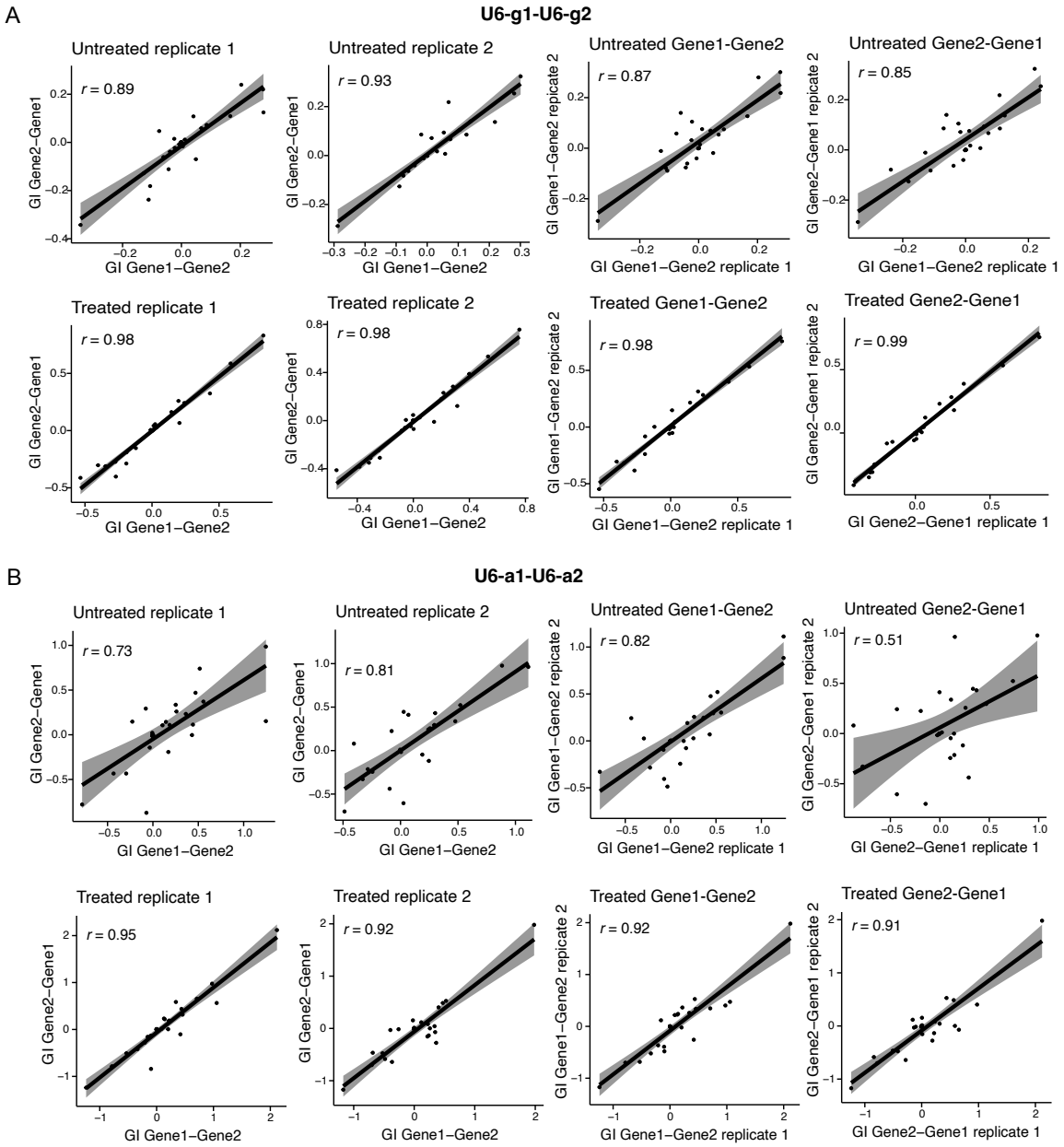
Supplementary Figure 5 - Cas13d screens phenotypes correlate with DepMap data.

***(A)** Heatmap of DepMap Chronos scores for the six candidate genes in K562, along with the MAGeCK MLE beta scores from the Cas9 counter screens and all Cas13d screens. **(B)** Pearson's correlation (r) between the beta scores presented in (A).*



Supplementary Figure 6 - High technical reproducibility of the U6-a1-U6-a2 screens in K562 at different levels of data analysis.

(A-B) *Correlation between normalized read counts from two technical screen replicates under untreated (A) and imatinib-treated (B) conditions.* ***(C)*** *Correlation between tau values from two technical screen replicates in the untreated condition prior to filtering for functional gRNAs.* ***(D-E)*** *Correlation between tau values from two technical screen replicates in the untreated (D) and imatinib-treated (E) conditions following the filtering for functional gRNAs.* ***(F)*** *Correlation between tau values from gRNA-gNTC and gNTC-gRNA combinations in the untreated condition. Pearson's correlation was used to calculate the r values.*



Supplementary Figure 7 - Cas13d enables the highly reproducible identification of GIs.

*Pearson's correlation of GI scores obtained from **(A)** single gRNA and **(B)** single-gene array screens, displayed without (top) and with imatinib treatment (bottom). The left panels show the correlation of GI scores between Gene1-Gene2 and Gene2-Gene1 orientations. The right panels show the correlation of GI scores between screen replicates.*

2.8.2 Supplementary Tables

Supplementary Table 1 - List of gRNA sequences in the U6-g1-g2 and U6-g1-U6-g2 Cas13d libraries.

gRNA on-target scores predicted by the cas13design.nygenome.org algorithm (234), and the gRNA TIGER score (166,234) are included.

Gene	gRNA	gRNA sequence	gRNA score	TIGER score
ABL1	ABL1_1	CCACAGATAAAATTCGAGGACAG	1.000	0.982
ABL1	ABL1_10	ACGTGTAAAATATAGTGGCACAT	0.918	0.981
ABL1	ABL1_11	CACAAAATCATACAGTGCAACGA	0.857	0.992
ABL1	ABL1_12	AATTAAATAAGACATTCGCCCAG	0.879	0.984
ABL1	ABL1_13	ATCTCCAACATTTTCCCGCGCCA	0.917	0.712
ABL1	ABL1_14	AGCAATACTCCAAATGCCCAGAC	0.853	0.978
ABL1	ABL1_15	CGCCACTTAGAAAAGAGCGTCTA	0.898	0.978
ABL1	ABL1_16	GAACCGCATAAAACGATCCAGAG	0.916	0.326
ABL1	ABL1_17	AATGATGATGAACCAACTCGGCC	0.848	0.120
ABL1	ABL1_18	GAAAAAGACAGTCCACATCAGCC	0.868	0.906
ABL1	ABL1_19	CAGAAACAAAATGCACTGGACCC	0.914	0.436
ABL1	ABL1_2	CAAAGTCAGATGCTACTGGCCGC	0.747	0.686
ABL1	ABL1_20	GTCACTGATTTCCTTCACCGAAC	0.794	0.569
ABL1	ABL1_21	AGAACTCAGAAAAGTGAAGGCTG	0.865	0.973
ABL1	ABL1_22	AAAACATTGCAGTGTGGACCCCC	0.911	n.a.
ABL1	ABL1_23	TGATTATAGCCTAAGACCCGGAG	0.791	0.059
ABL1	ABL1_24	CAAACGTAGGTAGTAGAGCTGGA	0.846	0.727
ABL1	ABL1_25	CACCCAAATCAAGAGCTGTCCTG	0.908	0.845
ABL1	ABL1_26	TTCACAGCATCAACCAGACTCGT	0.789	0.869
ABL1	ABL1_27	CCTTCTTGGATTTGCAGCCCACC	0.826	0.897
ABL1	ABL1_3	ACAAAATGCACTGGACCCCGAGA	0.951	0.928
ABL1	ABL1_4	ATTAAATAAGACATTCGCCCAGT	0.978	0.983
ABL1	ABL1_5	CCACAAAATCATACAGTGCAACG	0.874	0.992
ABL1	ABL1_6	GTGTAAAATATAGTGGCACATCA	0.919	0.979
ABL1	ABL1_7	CACAGATAAAATTCGAGGACAGA	0.954	0.943
ABL1	ABL1_8	AAAGTCAGATGCTACTGGCCGCT	0.710	0.243
ABL1	ABL1_9	GAAACATGTCTTCAGCACACGAC	0.905	0.938
GAB2	GAB2_1	ATAACTTAAGTGATTAGGCACCA	1.000	0.955
GAB2	GAB2_10	TCAGACAGAAGGTAATAGCGCCT	0.972	0.409
GAB2	GAB2_11	TCAGCCTGATTGAAGCCACAGAT	0.866	0.496
GAB2	GAB2_12	ATCAAGGACTTAAGGAGACACCA	0.882	0.959
GAB2	GAB2_13	CCATAAATCACATGACAGCCTGG	0.930	0.993
GAB2	GAB2_14	AAACACAAAACTATCCTGCAGCT	0.828	0.305
GAB2	GAB2_15	AGAAACTGGAGACAGCAGCCATT	0.864	0.991
GAB2	GAB2_16	CAAACAGTGACAATTCTGCGTGA	0.930	0.459
GAB2	GAB2_17	GTCCTCTTCTGTCTCAGCCACCA	0.805	0.703
GAB2	GAB2_18	AGATTCGATTTGAACACCTCAGT	0.864	0.320
GAB2	GAB2_19	ACCACAGCAAAAACATGACCCAC	0.902	0.984
GAB2	GAB2_2	AGTTACCAGAATAGGGGGCTGGA	0.601	0.511

GAB2	GAB2_20	AAGTTCAGGTTGATGATCCGCAG	0.729	0.947
GAB2	GAB2_21	ATTTCAAGACACAAGTAGCCAAC	0.862	0.979
GAB2	GAB2_22	TAAAAACATTCAGATCCCCCACC	0.898	0.282
GAB2	GAB2_23	GGAAACATTTCTCAGGGAGTCTG	0.653	0.888
GAB2	GAB2_24	TTAACTTCTACATTGGAAACCAC	0.845	0.757
GAB2	GAB2_25	GGAACATCCATATTGTCTACCAG	0.897	0.933
GAB2	GAB2_26	CCGCTCATCCGGCCACTCCGCAG	0.531	0.474
GAB2	GAB2_27	GCCAGATAATCAACGCTGCCGGT	0.841	0.474
GAB2	GAB2_3	CACCATAAATCACATGACAGCCT	0.963	0.932
GAB2	GAB2_4	CAGACAGAAGGTAATAGCGCCTC	0.981	0.964
GAB2	GAB2_5	CAAACACAAAACTATCCTGCAGC	0.879	0.715
GAB2	GAB2_6	CCACAGCAAAAACATGACCCACT	0.948	0.988
GAB2	GAB2_7	AACTTAAGTGATTAGGCACCAAC	1.000	0.980
GAB2	GAB2_8	TAGTTACCAGAATAGGGGGCTGG	0.538	0.116
GAB2	GAB2_9	TCCATTCACTTATTGGAACACCA	0.894	0.885
NF1	NF1_1	AAAGCAAGAAACAAGGCAGTCAA	1.000	0.989
NF1	NF1_10	CACACAGAAGATTATAGGCAGCT	1.000	0.339
NF1	NF1_11	TTAGAAAGGTTAAGGTTGGCAGT	0.802	0.957
NF1	NF1_12	TTATAAAACAGGAAGTGCAGCAT	0.995	0.261
NF1	NF1_13	AACGAACTAGATTTGACAGCCAT	1.000	0.993
NF1	NF1_14	GGATTCTTCATGGTACACCACAC	0.781	0.962
NF1	NF1_15	CACCATTAAGGACAACAGCCGAT	0.980	0.221
NF1	NF1_16	AAAAAGTTAAAGAAAAGGCAGCA	1.000	0.827
NF1	NF1_17	TGTTTGCTTTGAAAACGGTCCTG	0.763	0.474
NF1	NF1_18	AGCAAAATAGCATAACACCCACT	0.974	0.991
NF1	NF1_19	ACAGCAAGAAATGTTGAGACCCA	1.000	0.980
NF1	NF1_2	ATTAGAAAGGTTAAGGTTGGCAG	0.879	0.736
NF1	NF1_20	GTTTCTTCATCAATTCCAGGCAG	0.751	0.820
NF1	NF1_21	TTTAAAATAGTAGTGAGGCCGCT	0.972	0.165
NF1	NF1_22	AGCACAAACAAGTCACAGCACCG	1.000	0.993
NF1	NF1_23	CATTAAAACCAAAACTTTGCAGG	0.694	0.077
NF1	NF1_24	CGAAAGCAAGAAACAAGGCAGTC	0.970	0.697
NF1	NF1_25	AAGCAAGAATGAAGACAGTCAGC	1.000	0.985
NF1	NF1_26	TGGGAAGTTGCAAGTGAGGTCAT	0.575	0.611
NF1	NF1_27	TAGCAGCAATTCTAACAGCCCCC	0.969	0.311
NF1	NF1_3	CTAGAACAGTAAGAAGCAGCGCC	1.000	0.989
NF1	NF1_4	TTAGAAAACATGTTCCAGAGCAG	1.000	0.980
NF1	NF1_5	GAAACTTAACAATAAGCTCAGCA	0.875	0.854
NF1	NF1_6	AAACGAACTAGATTTGACAGCCA	1.000	0.869
NF1	NF1_7	TCACAAAATCATTGAAAGGCCGC	1.000	0.958
NF1	NF1_8	AAACTTAACAATAAGCTCAGCAT	0.871	0.899
NF1	NF1_9	GCTAGAACAGTAAGAAGCAGCGC	0.999	0.940
NTC	NTC_1	ACGCTACGAAGCTTATGGCACTG	n.a.	n.a.
NTC	NTC_10	CGGGGTCGTGCAATGTCCGGCTC	n.a.	n.a.
NTC	NTC_11	GGCAAGTGTGGTTACGCCGTATT	n.a.	n.a.
NTC	NTC_12	GCTGTGGCGGCCGGCTGTGAAAC	n.a.	n.a.

NTC	NTC_13	AATGAACCAGTTGAGCGCCAGGA	n.a.	n.a.
NTC	NTC_14	ATAATACTACCACCGTCTGCGTA	n.a.	n.a.
NTC	NTC_15	AAGGTTCGGCACTACCACCGCAA	n.a.	n.a.
NTC	NTC_16	CAATTTATCTTTGGAACGGAGCG	n.a.	n.a.
NTC	NTC_17	TTCTTCGGCTTAGACCAGTGCGG	n.a.	n.a.
NTC	NTC_18	GGGGATGCATAGTGCCAAGAGCA	n.a.	n.a.
NTC	NTC_19	TCGGGCCCCAGAGTACTAGGCAC	n.a.	n.a.
NTC	NTC_2	CATCCGCCCTGAATATTCACAAT	n.a.	n.a.
NTC	NTC_20	AACTGTGACGATAACGCCACATG	n.a.	n.a.
NTC	NTC_21	GGTGTACCGCATTTAAGGTACGA	n.a.	n.a.
NTC	NTC_22	GTCGGTTACTGCCCGGGAAGTTG	n.a.	n.a.
NTC	NTC_23	ATCGACCGACCGTGAAGATACAC	n.a.	n.a.
NTC	NTC_24	TCGTCGTAACGAGATGTATTCTA	n.a.	n.a.
NTC	NTC_25	CACCTGACTCAACTTAGTTGGTA	n.a.	n.a.
NTC	NTC_26	AAATTACCAAGCGGGCGAGACTA	n.a.	n.a.
NTC	NTC_27	CCGTCGCTATGCTCTATAGGGTT	n.a.	n.a.
NTC	NTC_28	TCTATCTGACACGGGATCAGCCT	n.a.	n.a.
NTC	NTC_29	GGAGTAAGCTGTCGTGGACAAAT	n.a.	n.a.
NTC	NTC_3	CTACCATATTGACCCGCGCCCGG	n.a.	n.a.
NTC	NTC_30	TACTTACCTAATGCGCGGTAGAT	n.a.	n.a.
NTC	NTC_4	CTCGGAAGACCGGAACGCGGGTG	n.a.	n.a.
NTC	NTC_5	ACATATGCGTCAATACTCCCCAG	n.a.	n.a.
NTC	NTC_6	GCTGGGTGACGCGATTCATGTTA	n.a.	n.a.
NTC	NTC_7	AAAGCTAGTCGGCTTAGGGTCGG	n.a.	n.a.
NTC	NTC_8	GTCGTAGCGCTAGGGTGGCCTTC	n.a.	n.a.
NTC	NTC_9	GCTGTAGAGCGACGACCTGTTAC	n.a.	n.a.
PTPN1	PTPN1_1	GACTTAAAAACAACAAGCCCAGC	0.986	0.798
PTPN1	PTPN1_10	TGCTACTATATATACACACCCAA	0.951	0.755
PTPN1	PTPN1_11	ATGTCTGAGTTACAGCAAGACCC	0.862	0.981
PTPN1	PTPN1_12	TCAGCCAGACAGAAGGTTCCAGA	0.820	0.448
PTPN1	PTPN1_13	CTTAAGGCATATAGCAGAGCAGC	0.916	0.985
PTPN1	PTPN1_14	CGACAATGACTTCAGCAACAGGC	0.847	0.799
PTPN1	PTPN1_15	AACCCATCATGAATATGGCAAGT	0.815	0.851
PTPN1	PTPN1_16	AAAATGTGAAAATGCTGGCAAGA	0.884	0.976
PTPN1	PTPN1_17	AAAATGGTTTATTCCATGGCCAT	0.845	0.918
PTPN1	PTPN1_18	TCCCTAAATCATGTCCAGAGCGT	0.812	0.939
PTPN1	PTPN1_19	ACCATCTCCCAAAAGTGACCGCA	0.879	0.992
PTPN1	PTPN1_2	TTAAGGCATATAGCAGAGCAGCT	0.937	0.827
PTPN1	PTPN1_20	TACATGCATTCTAATACACACGG	0.842	0.640
PTPN1	PTPN1_21	ACACCTCGAATGTTCCCACCACA	0.814	0.976
PTPN1	PTPN1_22	GAAGTCACTAGAGTGTCATGCCA	0.876	0.458
PTPN1	PTPN1_23	ATGTGTAACAAGGTGGAGAGCCA	0.830	0.542
PTPN1	PTPN1_24	ACAGAAGACCCTGAGACACTCCC	0.809	0.229
PTPN1	PTPN1_25	CTCTGTTGAGCATGACGACACCC	0.871	0.833
PTPN1	PTPN1_26	AGAAAGTTCAAGAATGAGGCTGG	0.826	0.991
PTPN1	PTPN1_27	TAAGCCTTATCCACACCTCACCA	0.808	0.799

PTPN1	PTPN1_3	AAGAAAGTTCAAGAATGAGGCTG	0.890	0.982
PTPN1	PTPN1_4	ATGCTACTATATATACACACCCA	0.960	0.980
PTPN1	PTPN1_5	AGCGACAATGACTTCAGCAACAG	0.911	0.691
PTPN1	PTPN1_6	CAAAATGTGAAAATGCTGGCAAG	0.906	0.991
PTPN1	PTPN1_7	ACTTAAAAACAACAAGCCCAGCT	0.957	0.209
PTPN1	PTPN1_8	TGTTTCTTATTAAATACCCACGT	0.869	0.582
PTPN1	PTPN1_9	GCCATGTAATGATCAGGTCATGC	0.825	0.913
SOS1	SOS1_1	AAGTTCAAAGTAATGGAGACAGT	1.000	0.990
SOS1	SOS1_10	AAAGTTCAAAGTAATGGAGACAG	0.988	0.990
SOS1	SOS1_11	GAGCATCATCATTAGACTCGAGA	0.905	0.282
SOS1	SOS1_12	TAAACCTATTCACAGCAGCAGAA	0.955	0.764
SOS1	SOS1_13	TTCTTACATCAAAGACCGACCGA	0.987	0.989
SOS1	SOS1_14	TGCAAATAAAGTGCTGCCCCAGG	0.936	0.944
SOS1	SOS1_15	AAATTCCAAAGAAAGGCACACAT	0.954	0.994
SOS1	SOS1_16	TAAACATAATTTGATGCCCAACA	0.985	0.432
SOS1	SOS1_17	CGATCATGAAAACCAGGTCGCAA	0.915	0.969
SOS1	SOS1_18	CTCAATAATTCGACTCACCACAG	0.951	0.946
SOS1	SOS1_19	GACGTAAAAGTGTCGCCCACCAA	0.979	0.948
SOS1	SOS1_2	CCATCAATACCTTGTCAGCACAC	0.986	0.994
SOS1	SOS1_20	AAACAATTTTGAATTGGAGACAA	0.890	0.075
SOS1	SOS1_21	ATGCCAGAAAAATAACAGCCGTT	0.949	0.973
SOS1	SOS1_22	AATAATTTAATATATGCCAGCCA	0.974	0.542
SOS1	SOS1_23	TCCATCAATACCTTGTCAGCACA	0.862	0.990
SOS1	SOS1_24	TGAAATTAATGTTTTAGGACGCT	0.940	0.307
SOS1	SOS1_25	CATATTAGATGTATAGACCACAC	0.961	0.898
SOS1	SOS1_26	CTGACTATAAAACCGACATGCAG	0.859	0.989
SOS1	SOS1_27	CTGTAAAGATATCAATGCTGCCA	0.934	0.759
SOS1	SOS1_3	TTCACCAGATATTTGCACCCCAA	1.000	0.978
SOS1	SOS1_4	ATAATTTAATATATGCCAGCCAA	1.000	n.a.
SOS1	SOS1_5	AGAGCATCATCATTAGACTCGAG	0.963	0.629
SOS1	SOS1_6	AATGCCAGAAAAATAACAGCCGT	0.997	0.995
SOS1	SOS1_7	ACTTATAGATGATACAGAAGCAG	0.993	0.278
SOS1	SOS1_8	TTTCTTACATAATTCCCAACCAG	0.954	0.893
SOS1	SOS1_9	CGACATACATTTAATACTCGCAG	0.958	0.978
SPRED2	SPRED2_1	CCCATGAAGAAAATGCAACCCAC	1.000	0.910
SPRED2	SPRED2_10	CATACAGTGATAGAGCATGCTGT	0.963	0.987
SPRED2	SPRED2_11	CAAACTTCCTATTATCGACCTTC	0.622	0.826
SPRED2	SPRED2_12	TACAGATTGTTAATAGTACAGGA	0.884	0.281
SPRED2	SPRED2_13	CATAGCATTCCAATACCACCAGT	0.959	0.992
SPRED2	SPRED2_14	CATTTTCTTGTTCACCTAGACGC	0.678	0.098
SPRED2	SPRED2_15	GCTAAAGATAGAAACTGCAGCTT	0.875	0.762
SPRED2	SPRED2_16	CAAGGCAATAAGAGCCATCCACC	0.919	0.319
SPRED2	SPRED2_17	TGAAACGTTGGATTGGCTTTGGT	0.579	0.821
SPRED2	SPRED2_18	TAACAACTAAATAGAGGTGACAA	0.874	0.557
SPRED2	SPRED2_19	AAAATCAACTACAAAACCCACCG	0.913	0.976
SPRED2	SPRED2_2	ACCATGGATGAGAAAGCCGCTTC	0.852	0.981

SPRED2	SPRED2_20	TTACTCCCCTGTCAAAGGCTCGG	0.562	0.634
SPRED2	SPRED2_21	TACAAGTTTGTTAACTAGCTCAC	0.874	0.393
SPRED2	SPRED2_22	ACACGCACAATATAGCTGTCATC	0.910	0.960
SPRED2	SPRED2_23	CCAAAGTCTGAGGAGTCCACGTA	0.833	0.980
SPRED2	SPRED2_24	GACGTATTTACATAAATGTCCCC	0.871	0.645
SPRED2	SPRED2_25	GAATGAATTATGCTTCAGGCACT	0.895	0.932
SPRED2	SPRED2_26	TGCATGACCTTACAGACCCCGAC	0.724	0.356
SPRED2	SPRED2_27	CCCATTAATATAAAATAGCCAGG	0.859	0.840
SPRED2	SPRED2_3	ACATACAGTGATAGAGCATGCTG	0.966	0.983
SPRED2	SPRED2_4	ACAAGGCAATAAGAGCCATCCAC	0.977	0.965
SPRED2	SPRED2_5	ACTTCCTATTATCGACCTTCCAG	0.685	0.951
SPRED2	SPRED2_6	CCATTAATATAAAATAGCCAGGG	0.959	0.194
SPRED2	SPRED2_7	CCATGAAGAAAATGCAACCCACC	1.000	0.835
SPRED2	SPRED2_8	TCACCATGGATGAGAAAGCCGCT	0.823	0.353
SPRED2	SPRED2_9	ATACATAATAACTGACTGCAGGC	0.884	0.979

Supplementary Table 2 - List of Cas13d gRNA and array sequences for cell surface marker knockdown.

Best performing arrays against each gene are highlighted in bold.

Gene	gRNA/Array	gRNA/Array sequence
CD46	CD46-1	CAGACAATTGTGTCGCTGCCATC
CD46	CD46-2	AGCCATGTATGATTCCGATCACA
CD46	CD46-3	ACAATCACAGCAATGACCCAAAC
CD46	**CD46-Array**	CAGACAATTGTGTCGCTGCCATCcaagtaaaccctaccaactggtcggggtttgaaacAGCCATGTATGATTCCG ATCACAcaagtaaaccctaccaactggtcggggtttgaaacACAATCACAGCAATGACCCAAACcaagtaaaccccta ccaactggtcggggtttgaaac
CD71	CD71-1	AGGAGAGCTGTGCCTACACCGGATTTTGCA
CD71	CD71-2	TTCTGCAGCAGCTCTGGAGATTGTCTGGAC
CD71	CD71-3	ACTCAGGAATCCTCTCAATCAGTTCCTTAT
CD71	CD71-Array-1	AGGAGAGCTGTGCCTACACCGGATTTTGCAcaagtaaaccctaccaactggtcggggtttgaaacTTCTGCAGC AGCTCTGGAGATTGTCTGGACcaagtaaaccctaccaactggtcggggtttgaaacACTCAGGAATCCTCTCAAT CAGTTCCTTATcaagtaaaccctaccaactggtcggggtttgaaac
CD71	CD71-Array-2	TTAATCACGAACTGACCAGCGACcaagtaaaccctaccaactggtcggggtttgaaacTCACAAATGAAAGCAG TTGGCTGcaagtaaaccctaccaactggtcggggtttgaaacAGCCACTGTAAACTCAGGCCCATcaagtaaacccc taccaactggtcggggtttgaaac
CD71	**CD71-Array-3**	CACAAATGAAAGCAGTTGGCTGTcaagtaaaccctaccaactggtcggggtttgaaacCCTCACAAATGAAAGC AGTTGGCcaagtaaaccctaccaactggtcggggtttgaaacCACGAGCAGAATACAGCCACTGTcaagtaaacccc taccaactggtcggggtttgaaac
CD71	CD71-Array-4	ATAAGGATAATCTGTGTCCTCGCcaagtaaaccctaccaactggtcggggtttgaaacAGACATGTCGGAAAGG AGACTCTcaagtaaaccctaccaactggtcggggtttgaaacCAGATTAATATAAGTGAAAGCCTcaagtaaacccc taccaactggtcggggtttgaaac
CD47	CD47-1	CCGCAATACAGAGACTCAGTCCA
CD47	CD47-2	AACCTGAAATCAGAAGAGGGCCA
CD47	CD47-3	GCATTGGTATACACGCCGCAATA
CD47	**CD47-Array**	CCGCAATACAGAGACTCAGTCCAcaagtaaaccctaccaactggtcggggtttgaaacAACCTGAAATCAGAAG AGGGCCAcaagtaaaccctaccaactggtcggggtttgaaacGCATTGGTATACACGCCGCAATAcaagtaaacccc taccaactggtcggggtttgaaac

CD63	CD63-1	CACAGTAACATTAATGCAGCAGG
CD63	CD63-2	AATCTGTGTAGTTAGCAGCCCCA
CD63	CD63-3	TAAGACAATAGTTCTCCTTGCAG
CD63	CD63-Array-1	CACAGTAACATTAATGCAGCAGGcaagtaaaccccctaccaactggtcggggtttgaaacAATCTGTGTAGTTAGC AGCCCCAcaagtaaaccccctaccaactggtcggggtttgaaacTAAGACAATAGTTCTCCTTGCAGcaagtaaaccccct accaactggtcggggtttgaaac
CD63	**CD63-Array-2**	CCACCAACATGATAAGAGACAGAcaagtaaaccccctaccaactggtcggggtttgaaacAGCCCCCAATCTTCTCC ACACAGcaagtaaaccccctaccaactggtcggggtttgaaacCCCTCATGTGACGCGGTAACAGCcaagtaaaccccct accaactggtcggggtttgaaac
CD63	CD63-Array-3	AGTACACATGCATTAAGGTCCCAcaagtaaaccccctaccaactggtcggggtttgaaacTTTTCGGGTAATTCTCC ATCTGCcaagtaaaccccctaccaactggtcggggtttgaaacCTACCCGGAAGAAGGCGCCCACCcaagtaaaccccct accaactggtcggggtttgaaac
CD63	CD63-Array-4	TATCTCTAAACACATAGCCAGCAcaagtaaaccccctaccaactggtcggggtttgaaacACAAAAGCAATTCCAAG GGCTGCcaagtaaaccccctaccaactggtcggggtttgaaacCTCCAGCCACGGTCCCCTCCCCCcaagtaaaccccta ccaactggtcggggtttgaaac
NTC	NTC-1	GTATCCCATAGTCCTTAAATTGG
NTC	NTC-2	CCTGTGCTTCCTTTGATCTGGGG
NTC	NTC-3	GACGATATATACCAAGCGAACGT
NTC	**NTC-Array**	GTATCCCATAGTCCTTAAATTGGcaagtaaaccccctaccaactggtcggggtttgaaacCCTGTGCTTCCTTTGATC TGGGGcaagtaaaccccctaccaactggtcggggtttgaaacGACGATATATACCAAGCGAACGTcaagtaaaccccta ccaactggtcggggtttgaaac

Supplementary Table 3 - List of gRNA sequences in U6-a1-U6-a2 Cas13d library.

Gene	gRNA	gRNA 1 sequence	gRNA 2 sequence	gRNA 3 sequence
ABL1	ABL1_1	CCACAGATAAAATTCGAGGACAG	CAAAGTCAGATGCTACTGGCCGC	ACAAAATGCACTGGACCCCGAGA
ABL1	ABL1_2	ATTAAATAAGACATTCGCCCAGT	CCACAAAATCATACAGTGCAACG	GTGTAAAATATAGTGGCACATCA
ABL1	ABL1_3	CACAGATAAAATTCGAGGACAGA	AAAGTCAGATGCTACTGGCCGCT	GAAACATGTCTTCAGCACACGAC
ABL1	ABL1_4	ACGTGTAAAATATAGTGGCACAT	CACAAAATCATACAGTGCAACGA	AATTAAATAAGACATTCGCCCAG
ABL1	ABL1_5	ATCTCCAACATTTTCCCGCGCCA	AGCAATACTCCAAATGCCCAGAC	CGCCACTTAGAAAAGAGCGTCTA
ABL1	ABL1_6	GAACCGCATAAAACGATCCAGAG	AATGATGATGAACCAACTCGGCC	GAAAAAGACAGTCCACATCAGCC
ABL1	ABL1_7	CAGAAACAAAATGCACTGGACCC	GTCACTGATTTCCTTCACCGAAC	AGAACTCAGAAAAGTGAAGGCTG
ABL1	ABL1_8	AAAACATTGCAGTGTGGACCCCC	TGATTATAGCCTAAGACCCGGAG	CAAACGTAGGTAGTAGAGCTGGA
ABL1	ABL1_9	CACCCAAATCAAGAGCTGTCCTG	TTCACAGCATCAACCAGACTCGT	CCTTCTTGGATTTGCAGCCCACC
GAB2	GAB2_1	ATAACTTAAGTGATTAGGCACCA	AGTTACCAGAATAGGGGGCTGGA	CACCATAAATCACATGACAGCCT
GAB2	GAB2_2	CAGACAGAAGGTAATAGCGCCTC	CAAACACAAAACTATCCTGCAGC	CCACAGCAAAAACATGACCCACT
GAB2	GAB2_3	AACTTAAGTGATTAGGCACCAAC	TAGTTACCAGAATAGGGGGCTGG	TCCATTCACTTATTGGAACACCA
GAB2	GAB2_4	TCAGACAGAAGGTAATAGCGCCT	TCAGCCTGATTGAAGCCACAGAT	ATCAAGGACTTAAGGAGACACCA
GAB2	GAB2_5	CCATAAATCACATGACAGCCTGG	AAACACAAAACTATCCTGCAGCT	AGAAACTGGAGACAGCAGCCATT
GAB2	GAB2_6	CAAACAGTGACAATTCTGCGTGA	GTCCTCTTCTGTCTCAGCCACCA	AGATTCGATTTGAACACCTCAGT
GAB2	GAB2_7	ACCACAGCAAAAACATGACCCAC	AAGTTCAGGTTGATGATCCGCAG	ATTTCAAGACACAAGTAGCCAAC
GAB2	GAB2_8	TAAAAACATTCAGATCCCCCACC	GGAAACATTTCTCAGGGAGTCTG	TTAACTTCTACATTGGAAACCAC
GAB2	GAB2_9	GGAACATCCATATTGTCTACCAG	CCGCTCATCCGGCCACTCCGCAG	GCCAGATAATCAACGCTGCCGGT
NF1	NF1_1	AAAGCAAGAAACAAGGCAGTCAA	ATTAGAAAGGTTAAGGTTGGCAG	CTAGAACAGTAAGAAGCAGCGCC
NF1	NF1_2	TTAGAAAACATGTTCCAGAGCAG	GAAACTTAACAATAAGCTCAGCA	AAACGAACTAGATTTGACAGCCA
NF1	NF1_3	TCACAAAATCATTGAAAGGCCGC	AAACTTAACAATAAGCTCAGCAT	GCTAGAACAGTAAGAAGCAGCGC
NF1	NF1_4	CACACAGAAGATTATAGGCAGCT	TTAGAAAGGTTAAGGTTGGCAGT	TTATAAAACAGGAAGTGCAGCAT
NF1	NF1_5	AACGAACTAGATTTGACAGCCAT	GGATTCTTCATGGTACACCACAC	CACCATTAAGGACAACAGCCGAT
NF1	NF1_6	AAAAAGTTAAAGAAAAGGCAGCA	TGTTTGCTTTGAAAACGGTCCTG	AGCAAAATAGCATAACACCCACT

NF1	NF1_7	ACAGCAAGAAATGTTGAGACCCA	GTTTCTTCATCAATTCCAGGCAG	TTTAAAATAGTAGTGAGGCCGCT
NF1	NF1_8	AGCACAAACAAGTCACAGCACCG	CATTAAAACCAAAACTTTGCAGG	CGAAAGCAAGAAACAAGGCAGTC
NF1	NF1_9	AAGCAAGAATGAAGACAGTCAGC	TGGGAAGTTGCAAGTGAGGTCAT	TAGCAGCAATTCTAACAGCCCCC
PTPN1	PTPN1_1	GACTTAAAAACAACAAGCCCAGC	TTAAGGCATATAGCAGAGCAGCT	AAGAAAGTTCAAGAATGAGGCTG
PTPN1	PTPN1_2	ATGCTACTATATATACACACCCA	AGCGACAATGACTTCAGCAACAG	CAAAATGTGAAAATGCTGGCAAG
PTPN1	PTPN1_3	ACTTAAAAACAACAAGCCCAGCT	TGTTTCTTATTAAATACCCACGT	GCCATGTAATGATCAGGTCATGC
PTPN1	PTPN1_4	TGCTACTATATATACACACCCAA	ATGTCTGAGTTACAGCAAGACCC	TCAGCCAGACAGAAGGTTCCAGA
PTPN1	PTPN1_5	CTTAAGGCATATAGCAGAGCAGC	CGACAATGACTTCAGCAACAGGC	AACCCATCATGAATATGGCAAGT
PTPN1	PTPN1_6	AAAATGTGAAAATGCTGGCAAGA	AAAATGGTTTATTCCATGGCCAT	TCCCTAAATCATGTCCAGAGCGT
PTPN1	PTPN1_7	ACCATCTCCCAAAAGTGACCGCA	TACATGCATTCTAATACACACGG	ACACCTCGAATGTTCCCACCACA
PTPN1	PTPN1_8	GAAGTCACTAGAGTGTCATGCCA	ATGTGTAACAAGGTGGAGAGCCA	ACAGAAGACCCTGAGACACTCCC
PTPN1	PTPN1_9	CTCTGTTGAGCATGACGACACCC	AGAAAGTTCAAGAATGAGGCTGG	TAAGCCTTATCCACACCTCACCA
SOS1	SOS1_1	AAGTTCAAAGTAATGGAGACAGT	CCATCAATACCTTGTCAGCACAC	TTCACCAGATATTTGCACCCCAA
SOS1	SOS1_2	ATAATTTAATATATGCCAGCCAA	AGAGCATCATCATTAGACTCGAG	AATGCCAGAAAAATAACAGCCGT
SOS1	SOS1_3	ACTTATAGATGATACAGAAGCAG	TTTCTTACATAATTCCCAACCAG	CGACATACATTTAATACTCGCAG
SOS1	SOS1_4	AAAGTTCAAAGTAATGGAGACAG	GAGCATCATCATTAGACTCGAGA	TAAACCTATTCACAGCAGCAGAA
SOS1	SOS1_5	TTCTTACATCAAAGACCGACCGA	TGCAAATAAAGTGCTGCCCCAGG	AAATTCCAAAGAAAGGCACACAT
SOS1	SOS1_6	TAAACATAATTTGATGCCCAACA	CGATCATGAAAACCAGGTCGCAA	CTCAATAATTCGACTCACCACAG
SOS1	SOS1_7	GACGTAAAAGTGTCGCCCACCAA	AAACAATTTTGAATTGGAGACAA	ATGCCAGAAAAATAACAGCCGTT
SOS1	SOS1_8	AATAATTTAATATATGCCAGCCA	TCCATCAATACCTTGTCAGCACA	TGAAATTAATGTTTTAGGACGCT
SOS1	SOS1_9	CATATTAGATGTATAGACCACAC	CTGACTATAAAACCGACATGCAG	CTGTAAAGATATCAATGCTGCCA
SPRED2	SPRED2_1	CCCATGAAGAAAATGCAACCCAC	ACCATGGATGAGAAAGCCGCTTC	ACATACAGTGATAGAGCATGCTG
SPRED2	SPRED2_2	ACAAGGCAATAAGAGCCATCCAC	ACTTCCTATTATCGACCTTCCAG	CCATTAATATAAAATAGCCAGGG
SPRED2	SPRED2_3	CCATGAAGAAAATGCAACCCACC	TCACCATGGATGAGAAAGCCGCT	ATACATAATAACTGACTGCAGGC
SPRED2	SPRED2_4	CATACAGTGATAGAGCATGCTGT	CAAACTTCCTATTATCGACCTTC	TACAGATTGTTAATAGTACAGGA
SPRED2	SPRED2_5	CATAGCATTCCAATACCACCAGT	CATTTTCTTGTTCACCTAGACGC	GCTAAAGATAGAAACTGCAGCTT
SPRED2	SPRED2_6	CAAGGCAATAAGAGCCATCCACC	TGAAACGTTGGATTGGCTTTGGT	TAACAACTAAATAGAGGTGACAA
SPRED2	SPRED2_7	AAAATCAACTACAAAACCCACCG	TTACTCCCCTGTCAAAGGCTCGG	TACAAGTTTGTTAACTAGCTCAC
SPRED2	SPRED2_8	ACACGCACAATATAGCTGTCATC	CCAAAGTCTGAGGAGTCCACGTA	GACGTATTTACATAAATGTCCCC
SPRED2	SPRED2_9	GAATGAATTATGCTTCAGGCACT	TGCATGACCTTACAGACCCCGAC	CCCATTAATATAAAATAGCCAGG
NTC	NTC_16	CAATTTATCTTTGGAACGGAGCG	TTCTTCGGCTTAGACCAGTGCGG	GGGGATGCATAGTGCCAAGAGCA
NTC	NTC_19	TCGGGCCCCAGAGTACTAGGCAC	AACTGTGACGATAACGCCACATG	GGTGTACCGCATTTAAGGTACGA
NTC	NTC_22	GTCGGTTACTGCCCGGGAAGTTG	ATCGACCGACCGTGAAGATACAC	TCGTCGTAACGAGATGTATTCTA
NTC	NTC_25	CACCTGACTCAACTTAGTTGGTA	AAATTACCAAGCGGGCGAGACTA	CCGTCGCTATGCTCTATAGGGTT
NTC	NTC_28	TCTATCTGACACGGGATCAGCCT	GGAGTAAGCTGTCGTGGACAAAT	TACTTACCTAATGCGCGGTAGAT
NTC	NTC_1	ACGCTACGAAGCTTATGGCACTG	CATCCGCCCTGAATATTCACAAT	CTACCATATTGACCCGCGCCCGG
NTC	NTC_4	CTCGGAAGACCGGAACGCGGGTG	ACATATGCGTCAATACTCCCCAG	GCTGGGTGACGCGATTCATGTTA
NTC	NTC_7	AAAGCTAGTCGGCTTAGGGTCGG	GTCGTAGCGCTAGGGTGGCCTTC	GCTGTAGAGCGACGACCTGTTAC
NTC	NTC_10	CGGGGTCGTGCAATGTCCGGCTC	GGCAAGTGTGGTTACGCCGTATT	GCTGTGGCGGCCGGCTGTGAAAC
NTC	NTC_13	AATGAACCAGTTGAGCGCCAGGA	ATAATACTACCACCGTCTGCGTA	AAGGTTCGGCACTACCACCGCAA

Supplementary Table 4 - List of sgRNA sequences in Cas9 counter screen library.

Gene	sgRNA	sgRNA sequence
ABL1	ABL1_1	CTTAGGCTATAATCACAATG
ABL1	ABL1_2	GGTTCATCATCATTCAACGG
ABL1	ABL1_3	TCAGTGATGATATAGAACGG
ABL1	ABL1_4	TTGCTCCCTCGAAAAGAGCG
GAB2	GAB2_1	CAAACGTGAACAGAGTTGGC
GAB2	GAB2_2	GAATGTCCTAGGGATCTGGT
GAB2	GAB2_3	GACAACAGCCGACTTCACCG
GAB2	GAB2_4	TGCTGAGGACTGCCCCATCG
NF1	NF1_1	CATGAGACCACTGTCTACGT
NF1	NF1_2	CCACCACCTAGAATCGAAAG
NF1	NF1_3	GAGGAAGCAGATATCCGGTG
NF1	NF1_4	GGTCCAGTCAGTGAACGTAA
PTPN1	PTPN1_1	AAGGTGCCAAATTCATCATG
PTPN1	PTPN1_2	AGGGCCTCCTTACCAGCAAG
PTPN1	PTPN1_3	GAAGCTTGGCCACTCTACAT
PTPN1	PTPN1_4	GGCCCTTTGCCTAACACATG
SOS1	SOS1_1	AGTACCGGAGTACACTGGAA
SOS1	SOS1_2	AGTGGCATATAAGCAGACCT
SOS1	SOS1_3	ATTAAATGTATGTCGGCACT
SOS1	SOS1_4	CTTACGTACCATATGTACGC
SPRED2	SPRED2_1	AGGCGGGATCAGTCGCGTCG
SPRED2	SPRED2_2	AGTCTGAGGAGTCCACGTAG
SPRED2	SPRED2_3	CTGTGGGGTATGAGTCGTGG
SPRED2	SPRED2_4	TCCAACGTTTCATCACTGGA
non_target	NTC-1	AGTATGAGACTCATAGGGTG
non_target	NTC-2	ATGTAACGAGTTGTAAGTCA
non_target	NTC-3	AGAAGACCGAGGCGCTTCAA
non_target	NTC-4	CATTAGCAGCCCAGCGCCCA
non_target	NTC-5	CGGGATGGTCCCTGCCGAGA
safe_cutter	SAFE_CUTTER-1	GCCATTAATCATGATCTGGA
safe_cutter	SAFE_CUTTER-2	GACAATAGAAATTCGAAATG
safe_cutter	SAFE_CUTTER-3	GAAATCTAAAGATGTACACT
safe_cutter	SAFE_CUTTER-4	GTCCATGTATTTGACCATCA
safe_cutter	SAFE_CUTTER-5	GCGTAGTACTGTGGGGTAAA

Supplementary Table 5 - List of Cas9 sgRNA sequences for cell surface marker knockout.

Best performing sgRNAs against each gene are highlighted in bold.

Gene	gRNA	gRNA sequence
CD46	**CD46-1**	ACTCGTAAGTCCCATTTGCA
CD46	CD46-2	ATTGTGGTGACAATTCAGTG
CD46	CD46-3	GGAGTACAGCAGCAACACCA
CD46	CD46-4	GGATCAGTAGCAATTTGGAG
CD71	CD71_1	AAATTCATATGTCCCTCGTG
CD71	CD71_2	AATTGGTGTGTTGATATACA
CD71	CD71_3	CAGGAACCGAGTCTCCAGTG
CD71	**CD71_4**	CTATACGCCACATAACCCCC
CD63	CD63-1	GGCTGCTAACTACACAGATT
CD63	**CD63-2**	TGTCCAGGATCGAAGCAGTG
CD63	CD63-3	AAGACAATAGTTCTCCTTGC
CD63	CD63-4	CCTGAGTCAGACCATAATCC
CD47	CD47-1	TAAGCACTTAAATATAGATC
CD47	**CD47-2**	ATCGAGCTAAAATATCGTGT
CD47	CD47-3	ACAGGAGTATAGCAAAAATT
CD47	CD47-4	ATGCTTTGTTACTAATATGG
NTC	NTC1	AGTATGAGACTCATAGGGTG

Supplementary Table 6 - List of Cas9 dual targeting oligo sequences for CD46 and CD47 cell surface marker dual knockout.

Oligo name	Oligo sequence
pMB1D-CD46-1-NTC2	CCCTTGGAGAACCACCTTGTTGGACTCGTAAGTCCCATTTGCAGTTTGAGAGCTAAGCAGAAAGCTGCATAGCAAG TTCAAATAAGGCTAGTCCGTACACAACTTGAAAAAGTGGCAGCCGAGTCGGCTGCTTTTTTGAACCGACGGATGAT CTCGTGCACCGGTTCAAAAAAGCACCGACTCGGTGCCACTTTGCTGTTTCCAGCAAAGTTGATAACGGACTAGCCTT ATTTAAACTTGCTATGCTGTTTCCAGCATAGCTCTTAAACATTGAAGCTCAACCAAAGTCCGGGAAAGAGTGGTCTC ATACAGAACTTATAAGATTCCCAAATCCAAAGACATTTCACGTTTATGGTGATTTCCCAGAACACATAGCGACATGC AAATATTGCAGGGCGCCACTCCCCTGTCCCTCACAGCCATCTTCCTGCCAGGGCGCACGCGCGCTGGGTGTTCCCG CCTAGTGACACTGGGCCCGCGATTCCTTGGAGCGGGTTGATGACGTCAGCGTTCGCGGCCGCGTGGATAACCGTA TT
pMB1D-NTC1-CD46-1	CCCTTGGAGAACCACCTTGTTGGAGTATGAGACTCATAGGGTGGTTTGAGAGCTAAGCAGAAAGCTGCATAGCAA GTTCAAATAAGGCTAGTCCGTACACAACTTGAAAAAGTGGCAGCCGAGTCGGCTGCTTTTTTGAACCGACGGATGA TCTCGTGCACCGGTTCAAAAAAGCACCGACTCGGTGCCACTTTGCTGTTTCCAGCAAAGTTGATAACGGACTAGCCT TATTTAAACTTGCTATGCTGTTTCCAGCATAGCTCTTAAACTGCAAATGGGACTTACGAGTCGGGAAAGAGTGGTCT CATACAGAACTTATAAGATTCCCAAATCCAAAGACATTTCACGTTTATGGTGATTTCCCAGAACACATAGCGACATG CAAATATTGCAGGGCGCCACTCCCCTGTCCCTCACAGCCATCTTCCTGCCAGGGCGCACGCGCGCTGGGTGTTCCC GCCTAGTGACACTGGGCCCGCGATTCCTTGGAGCGGGTTGATGACGTCAGCGTTCGCGGCCGCGTGGATAACCGT ATT
pMB1D-CD47-2-NTC2	CCCTTGGAGAACCACCTTGTTGGATCGAGCTAAAATATCGTGTGTTTGAGAGCTAAGCAGAAAGCTGCATAGCAAG TTCAAATAAGGCTAGTCCGTACACAACTTGAAAAAGTGGCAGCCGAGTCGGCTGCTTTTTTGAACCGACGGATGAT CTCGTGCACCGGTTCAAAAAAGCACCGACTCGGTGCCACTTTGCTGTTTCCAGCAAAGTTGATAACGGACTAGCCTT ATTTAAACTTGCTATGCTGTTTCCAGCATAGCTCTTAAACATTGAAGCTCAACCAAAGTCCGGGAAAGAGTGGTCTC ATACAGAACTTATAAGATTCCCAAATCCAAAGACATTTCACGTTTATGGTGATTTCCCAGAACACATAGCGACATGC AAATATTGCAGGGCGCCACTCCCCTGTCCCTCACAGCCATCTTCCTGCCAGGGCGCACGCGCGCTGGGTGTTCCCG CCTAGTGACACTGGGCCCGCGATTCCTTGGAGCGGGTTGATGACGTCAGCGTTCGCGGCCGCGTGGATAACCGTA TT

pMB1D-NTC1-CD47-2	CCCTTGGAGAACCACCTTGTTGGAGTATGAGACTCATAGGGTGGTTTGAGAGCTAAGCAGAAAGCTGCATAGCAA GTTCAAATAAGGCTAGTCCGTACACAACTTGAAAAAGTGGCAGCCGAGTCGGCTGCTTTTTTGAACCGACGGATGA TCTCGTGCACCGGTTCAAAAAAGCACCGACTCGGTGCCACTTTGCTGTTTCCAGCAAAGTTGATAACGGACTAGCCT TATTTAAACTTGCTATGCTGTTTCCAGCATAGCTCTTAAACACACGATATTTTAGCTCGATCGGGAAAGAGTGGTCT CATACAGAACTTATAAGATTCCCAAATCCAAAGACATTTCACGTTTATGGTGATTTCCCAGAACACATAGCGACATG CAAATATTGCAGGGCGCCACTCCCCTGTCCCTCACAGCCATCTTCCTGCCAGGGCGCACGCGCGCTGGGTGTTCCC GCCTAGTGACACTGGGCCCGCGATTCCTTGGAGCGGGTTGATGACGTCAGCGTTCGCGGCCGCGTGGATAACCGT ATT
pMB1D-CD46-1-CD47-2	CCCTTGGAGAACCACCTTGTTGGACTCGTAAGTCCCATTTGCAGTTTGAGAGCTAAGCAGAAAGCTGCATAGCAAG TTCAAATAAGGCTAGTCCGTACACAACTTGAAAAAGTGGCAGCCGAGTCGGCTGCTTTTTTGAACCGACGGATGAT CTCGTGCACCGGTTCAAAAAAGCACCGACTCGGTGCCACTTTGCTGTTTCCAGCAAAGTTGATAACGGACTAGCCTT ATTTAAACTTGCTATGCTGTTTCCAGCATAGCTCTTAAACACACGATATTTTAGCTCGATCGGGAAAGAGTGGTCTC ATACAGAACTTATAAGATTCCCAAATCCAAAGACATTTCACGTTTATGGTGATTTCCCAGAACACATAGCGACATGC AAATATTGCAGGGCGCCACTCCCCTGTCCCTCACAGCCATCTTCCTGCCAGGGCGCACGCGCGCTGGGTGTTCCCG CCTAGTGACACTGGGCCCGCGATTCCTTGGAGCGGGTTGATGACGTCAGCGTTCGCGGCCGCGTGGATAACCGTA TT
pMB1D-CD47-2-CD46-1	CCCTTGGAGAACCACCTTGTTGGATCGAGCTAAAATATCGTGTGTTTGAGAGCTAAGCAGAAAGCTGCATAGCAAG TTCAAATAAGGCTAGTCCGTACACAACTTGAAAAAGTGGCAGCCGAGTCGGCTGCTTTTTTGAACCGACGGATGAT CTCGTGCACCGGTTCAAAAAAGCACCGACTCGGTGCCACTTTGCTGTTTCCAGCAAAGTTGATAACGGACTAGCCTT ATTTAAACTTGCTATGCTGTTTCCAGCATAGCTCTTAAACTGCAAATGGGACTTACGAGTCGGGAAAGAGTGGTCTC ATACAGAACTTATAAGATTCCCAAATCCAAAGACATTTCACGTTTATGGTGATTTCCCAGAACACATAGCGACATGC AAATATTGCAGGGCGCCACTCCCCTGTCCCTCACAGCCATCTTCCTGCCAGGGCGCACGCGCGCTGGGTGTTCCCG CCTAGTGACACTGGGCCCGCGATTCCTTGGAGCGGGTTGATGACGTCAGCGTTCGCGGCCGCGTGGATAACCGTA TT
pMB1D-NTC1-NTC2	CCCTTGGAGAACCACCTTGTTGGAGTATGAGACTCATAGGGTGGTTTGAGAGCTAAGCAGAAAGCTGCATAGCAA GTTCAAATAAGGCTAGTCCGTACACAACTTGAAAAAGTGGCAGCCGAGTCGGCTGCTTTTTTGAACCGACGGATGA TCTCGTGCACCGGTTCAAAAAAGCACCGACTCGGTGCCACTTTGCTGTTTCCAGCAAAGTTGATAACGGACTAGCCT TATTTAAACTTGCTATGCTGTTTCCAGCATAGCTCTTAAACATTGAAGCTCAACCAAAGTCCGGGAAAGAGTGGTCT CATACAGAACTTATAAGATTCCCAAATCCAAAGACATTTCACGTTTATGGTGATTTCCCAGAACACATAGCGACATG CAAATATTGCAGGGCGCCACTCCCCTGTCCCTCACAGCCATCTTCCTGCCAGGGCGCACGCGCGCTGGGTGTTCCC GCCTAGTGACACTGGGCCCGCGATTCCTTGGAGCGGGTTGATGACGTCAGCGTTCGCGGCCGCGTGGATAACCGT ATT

Supplementary Table 7 - List of Cas12a gRNA sequences for cell surface marker knockout.

Best performing gRNAs against each gene are highlighted in bold.

Gene	gRNA	gRNA sequence
CD47	**enAsCas12a-CD47**	ATACCAAACTGTCCCCAGAACAG
CD63	**enAsCas12a-CD63**	TCCCAATCTGTGTAGTTAGCAGC
Safe-cutter	enAsCas12a-Safe-1	CCGACAAGAGACGGTCCTAATGT
Safe-cutter	enAsCas12a-Safe-2	GCTTACGATCGTTGGAGGGCAGC
CD46	enAs12a-CD46-1	CTTACTATAACAGGCGTCATCTG
CD46	**enAs12a-CD46-2**	CAGTGGTCAAATGTCGATTTCCA
CD46	enAs12a-CD46-3	GGAGGTGGTGTACACAAAACCTC
CD46	enAs12a-CD46-4	AATGGCCAAGCAGTCCCTGCAAA
CD71	enAsCas12a-CD71-1	GGAATATGGAAGGAGACTGTCCC
CD71	**enAsCas12a-CD71-2**	GTGGCACCAACCGATCCAAAGTC
CD71	enAsCas12a-CD71-3	TAGAACCAGATCACTATGTTGTA
CD71	enAsCas12a-CD71-4	CGAGGACACAGATTATCCTTATT

Supplementary Table 8 - List of Cas12a dual targeting oligo sequences for CD46 and CD47 cell surface marker dual knockout.

Oligo name	Oligo sequence
enAs12a-CD47-Safe2_s	AAAAGCTGCCCTCCAACGATCGTAAGCATCTACAAGAGTAGAAATTACTGTTCTGGGGACAGTTTGGTATATCTACAAGAGTAGAAATTA
enAs12a-CD47-Safe2_as	TTGGTAATTTCTACTCTTGTAGATATACCAAACTGTCCCCAGAACAGTAATTTCTACTCTTGTAGATGCTTACGATCGTTGGAGGGCAGC
enAs12a-Safe1-CD47_s	AAAACTGTTCTGGGGACAGTTTGGTATATCTACAAGAGTAGAAATTAACATTAGGACCGTCTCTTGTCGGATCTACAAGAGTAGAAATTA
enAs12a-Safe1-CD47_as	TTGGTAATTTCTACTCTTGTAGATCCGACAAGAGACGGTCCTAATGTTAATTTCTACTCTTGTAGATATACCAAACTGTCCCCAGAACAG
enAs12a-CD46-Safe2_s	aaaaGCTGCCCTCCAACGATCGTAAGCatctacaagagtagaaattaTGGAAATCGACATTTGACCACTGatctacaagagtagaaatta
enAs12a-CD46-Safe2_as	ttggtaatttctactcttgtagatCAGTGGTCAAATGTCGATTTCCAtaatttctactcttgtagatGCTTACGATCGTTGGAGGGCAGC
enAs12a-Safe1-CD46_s	aaaaTGGAAATCGACATTTGACCACTGatctacaagagtagaaattaACATTAGGACCGTCTCTTGTCGGatctacaagagtagaaatta
enAs12a-Safe1-CD46_as	ttggtaatttctactcttgtagatCCGACAAGAGACGGTCCTAATGTtaatttctactcttgtagatCAGTGGTCAAATGTCGATTTCCA
enAs12a-CD47-CD46_s	aaaaTGGAAATCGACATTTGACCACTGatctacaagagtagaaattaCTGTTCTGGGGACAGTTTGGTATatctacaagagtagaaatta
enAs12a-CD47-CD46_as	ttggtaatttctactcttgtagatATACCAAACTGTCCCCAGAACAGtaatttctactcttgtagatCAGTGGTCAAATGTCGATTTCCA
enAs12a-CD46-CD47_s	aaaaCTGTTCTGGGGACAGTTTGGTATatctacaagagtagaaattaTGGAAATCGACATTTGACCACTGatctacaagagtagaaatta
enAs12a-CD46-CD47_as	ttggtaatttctactcttgtagatCAGTGGTCAAATGTCGATTTCCAtaatttctactcttgtagatATACCAAACTGTCCCCAGAACAG
enAs12a-Safe1-Safe2_s	AAAAGCTGCCCTCCAACGATCGTAAGCATCTACAAGAGTAGAAATTAACATTAGGACCGTCTCTTGTCGGATCTACAAGAGTAGAAATTA
enAs12a-Safe1-Safe2_as	TTGGTAATTTCTACTCTTGTAGATCCGACAAGAGACGGTCCTAATGTTAATTTCTACTCTTGTAGATGCTTACGATCGTTGGAGGGCAGC

Supplementary Table 9 - List of ssDNA primers and oligo sequences.

Primer name	Primer sequence
enAsCas12a-FOR	ACCGGTTCTAGAGCGATGCCTGCTGCTAAGAGAGTGAAACTGGATCCTGCTGCTAAGAGAGTGAAACTGGATCCTGCTGCTAAGAGAGTGAAACTGGATACACAGTTCGAGGGCTTTACC
enAsCas12a-REV	GGATCCGCTAGCGTTGCGCAGCTCCTGGA
enAsCas12a-GA-FOR	AACACAGGACCGGTTCTAGACTAGAGCGATGCCTGCT
enAsCas12a-GA-REV	AAGTTTGTTGCGCCGGATCCGCTAGCGTTGCG
S1-GA-FOR	GATCCACTTTGGCGCCGGCCTCGAGCAG
S1-GA-REV	CTTTCAAGACCTAGGGCCCCCCTCGAGCCCGGGCATGCTCTTCAACCTCAATAACTGGAGTTATATGGACCATTGTTCTAGCGCTGATCCGACG
GG_Dual_gRNA_For	GCCGTCTAATGTTCAGCTAGTATGCACAGTTGATCCGTCTC
GG_Dual_gRNA_Rev	GTGTGACGTATGATCAGATCTATGCTACAGTGAACCGTCTC
pLV_CD71-Array-T4-3x Diff-For	actggtcggggtttgaaacCACCTGCGATCaaacAGGAGAGCTGTGCCTACACC
pLV_CD71-Array-T4-3x Diff-Rev	cgcggccaattcaaaaaaagtttcaaaccccgaccagttggtaggggtttacttgATAAGGAACTGATTGAGAGGATTCCTGAG
783-Rx-hU6-CD46-Array-1-For	actggtcggggtttgaaacCAGACAATTGTGTCGCTGCCATC
783-Rx-hU6-CD46-Array-1-Rev	cgcggccaattcaaaaaaagtttcaaaccccgaccagttggtaggggtttacttgGTTTGGGTCATTGCTGTGATTGTG
783-Rx-hU6-CD47-Array-1-For	actggtcggggtttgaaacCCGCAATACAGAGACTCAGTCCAc
783-Rx-hU6-CD47-Array-1-Rev	cgcggccaattcaaaaaaagtttcaaaccccgaccagttggtaggggtttacttgTATTGCGGCGTGTATACCAATGC

Supplementary Table 10 - List of synthetic DNA fragment sequences.

Fragment name	Fragment sequence
Illumina PBS-Filler-mU6-EF1a Ready for Gibson	CACTTTGGCGCCGGCCTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCAT TATAAGCTGCAATAAACAAGTTGGATTTGGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC AAAAAAAGAGACGTACAAAAAAGAGCAAGAAGCTAAAAAAGATTTAAAAATTATTTTTAGCGCA GTTAATGGAACAGGAACTAAATTTACCCCAAAGAATCAGGATATGAAGATGAAACATTTAAAAA TGTTGTAAATCCAAATCCAGAATTTGATCCTGCATGAAAAATACCGCTTGAATATGTTTATTTTGC TCATTTTTATTAAAAGTTTGCAAAAAGGAACATAAAAATTCTAATTATTGATACTAAAGTTATTAA AAAGAAGGGTCATAGAATATAATATTGCTCATTTACAACCGTCTCCCCAACAAGGTGGTTCTCCA AGGGATACTTATAGTCTCAAAACACACAATTACTTTACAGTTAGGGTGAGTTTCCTTTTGTGCTGT TTATAACTTCGTATAATGTATGCTATACGAAGTTATAGAAATCCAAGCCTATCATGTAAAATGTA GCTAGTATTAAAAAGAACAGATTATCTGTCTTTTATCGCACATTAAGCCTCTATAGTTACTAGGAA ATATTATATGCAAATTAACCGGGGCAGGGGAGTAGCCGAGCTTCTCCCACAAGTCCGTGCGAGG GGGCCGGCGCGGGCCTAGAGATGGCGCGTCGGATCAGCGCTAGAACAATGGTCCTGTGTTCTG GCGGCAAACCCGTTGCGAAAAAGAACGTTCACGGCGACTACTGCACTTATATACGGTTCTCCCCC ACCCTCGGGAAAAAGGCGGAGCCAGTACACGACATCACTTTCCCAGTTTACCCCGCGCCACCTTC TCTAGGCACCGGATCAATTGCCGACCCCTCCCCCCAACTTCTCGGGGACTGTGGGCGATGTGCGC TCTGCCCACTGACGGGCACCGGAGCCAATTCCCACTCCTTTCAAGACCTACCCGGGCTCGAGGG GGGCCCTAGGTCTTG
WPRE-3LTR Ready for Gibson	GACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAG ACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCGTCGACTTTAAGACCAATGACTTACAAGGCA GCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAC GAAGACAAGATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAG CTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGT AGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGT GGAAAATCTCTAGCAGTACGTATAGTAGTTCATGTCATCTTATTATTC
sgLenti orthogonal vector Fragment	cttggagaaccaccttgttggagacgcaggtgtcactggcgctgttggtgccggcacctgacacagcagaggatgcagagccc aaggggcgcgccaaagccgaaggagccacccagcacgcaagccaggaaccacctgcgatcGTTTAAGAGCTAAGCT GGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTgaaaAAGTGGCACCGAGTC GGTGCGCTCACCTATTAGCGGCTAAGGttttttttgcatgcacagcatccccaaaggaggtgagcactttcttgccgt gggcgcggaccatggggttgccaaggatggcagtggggctggagaggttcccaaaggacgcaaagaacctctgggtccagg ggtagacgatcagcagcctggccagggcttcggccccacattcggccacattgaccttgccccagaggccggtgatgagccct aggtagcggtcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggtctgttagagagata attagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcag ttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaa aggacgaaacaccgtgagacgcaggtacttccccagcaccagcgtcttcccttgagtgacagctaggagttgtgaggcgcgcc gcagcagcagcagcaagcgcctaagagagatggctcggcacatgtacctgcgatcGCTTTAAGGCCGGTCCTAGCA Atttttagatcggaagagcacacgtctgaactccagtcacgcgtggataaccgtattaccgccatgcatgtgcccgtcagtgg
sgLenti orthogonal vector Fragment 2	cttggagaaccaccttgttggagacgcaggtgtcactggcgctgttggtgccggcacctgacacagcagaggatgcagagccc aaggggcgcgccaaagccgaaggagccacccagcacgcaagccaggaaccacctgcgatcGCTCACCTATTAGCGG CTAAGGttttttttgcatgcacagcatccccaaaggagg
sgLenti-Orthogonal-CasRx-CasRx-DR36_hU6-Frag	AGTATGCACAGTTGATCCGTCTCAtgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttga aagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccgtgcaagtaaacccctaccaaagacgcaggtact tccccagcaccagcgtcttcccttgagtgacagctaggagttgtgaggcgcgccgcagcgcagcagcaagcgcctaagagag atggctcggcacatgtacctgcgatcctggtcggggtttgaaacGCTTAGAGACGGTTCACTGTAGCA
sgLenti-Orthogonal-CasRx-CasRx-DR36_Intermediate-Frag	AGTATGCACAGTTGATCCGTCTCAcagtggggctggagaggttcccaaaggacgcaaagaacctctgggtccagg ggtagacgatcagcagcctggccagggcttcggccccacattcggccacattgaccttgccccagaggccggtgatgagccct aggtagcggtcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggtctgttagagagata attagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgca AGAGACGGTTCACTGTAGCA
sgLenti-Orthogonal-CasRx-CasRx-DR36_mU6-Frag	AGTATGCACAGTTGATCCGTCTCAttggcaagtaaacccctaccaaagacgcaggtgtcactggcgctgttggtgcc ggcacctgacacagcagaggatgcagagcccaaggggcgcgccaaagccgaaggagccacccagcacgcaagccaggaa ccacctgcgatcctggtcggggtttgaaacGCTCACCTATTAGCGGCTAAGGttttttttgcatgcacagcatccccaaa ggaggtgagcactttcttgccgtgggcgcggaccatggggttgccaaggatggcagtAGAGACGGTTCACTGTAGCA

2.8.3 Supplementary Data

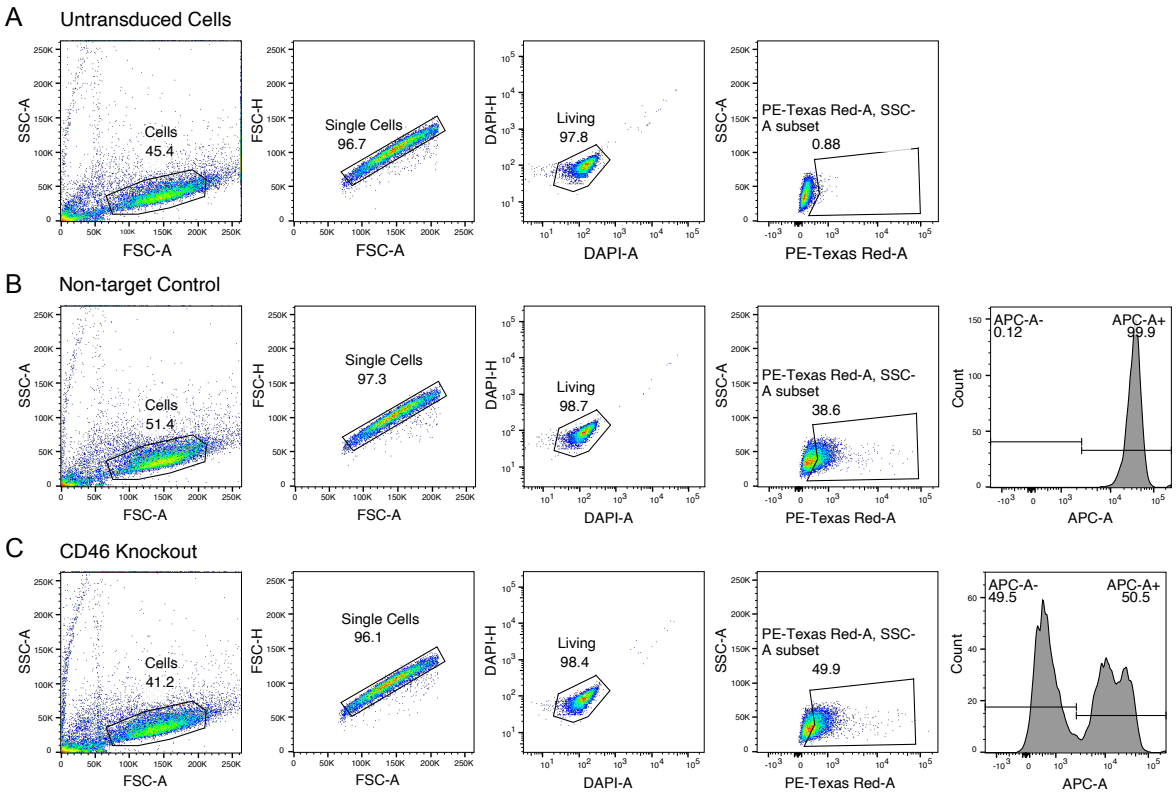
Supplementary Data 1: Calculated tau values and GI scores of gRNA and array sequences from Cas9 and Cas13d screens samples.

Supplementary Data 2: Raw count numbers of gRNA and array sequences from Cas9 and Cas13d screens samples.

Supplementary Data 3: Statistical analysis of GI scores calculated from Cas13d screens.

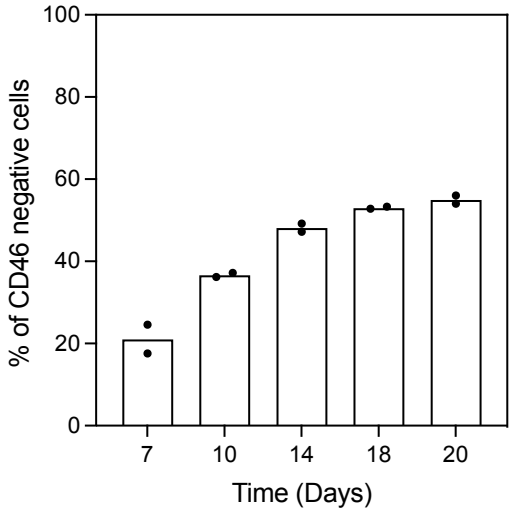
2.9 Appendix Chapter 2

2.9.1 Supplementary Figures



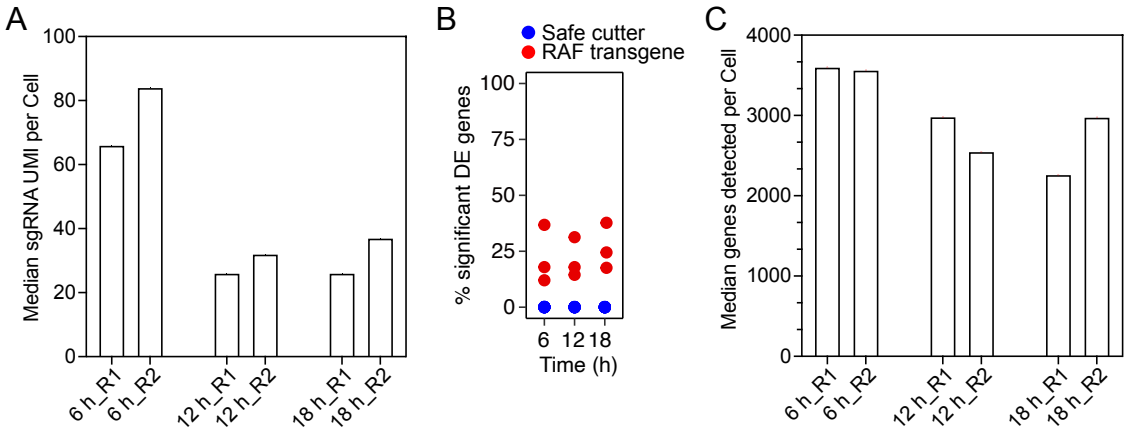
Supplementary Figure 8 - Example of the flow cytometry gating strategy used in the analysis of the CD46 knockout kinetics in HEK293ΔRAF1:ER cells at different time points after lentiviral infection.

***(A)** Untransduced control cells. **(B)** Non-target control cells. **(C)** CD46 knockout cells. Percentage of CD46 knockout cells was determined via flow cytometry analysis of >10,000 cells stained with CD46 antibodies (Miltenyi, 130-130-362).*



Supplementary Figure 9 - CD46 knockout kinetics in HEK293ΔRAF1:ER cells at different time points after lentiviral infection.

Percentage of CD46 knockout cells was determined via flow cytometry analysis of >10,000 cells stained with CD46 antibodies (Miltenyi, 130-130-362). Values represent the mean of biological replicates (n=2).



Supplementary Figure 10 - Quality assessment of Perturb-seq screen performance.

***(A)** Median sgRNA UMI counts per cell detected in the respective Perturb-seq samples. **(B)** Percentage of significantly differentially expressed genes in the RAF1-knockout cells from the total number of genes induced by RAF activation (adjusted p-value<0.05). Red = RAF1 sgRNAs, Blue = safe cutter sgRNAs. **(C)** Median number of genes detected per cell in the respective Perturb-seq samples.*

2.9.2 Supplementary Tables

Supplementary Table 11 - List of sgRNA sequences in Cas9 Perturb-seq library.

Gene	sgRNA	sgRNA sequence
ARID3B	ARID3B_1	AATTGATGGCAACCGCAGGG
ARID3B	ARID3B_2	GCTGTGGCGGAAAGGAGAGT
ARID3B	ARID3B_3	GTGGCCCAAGTGTTTGAACG
ARID3B	ARID3B_4	TTCGTCTTTATGCAGAAGAG
BHLHE40	BHLHE40_1	CAAGTGTACAAGTCAAGACG
BHLHE40	BHLHE40_2	GACTTTCGCTCACTCGAGTG
BHLHE40	BHLHE40_3	GCAGCTCCGAGACCACCCGG
BHLHE40	BHLHE40_4	TCAAAAGTGACCACGGACGC
CSRNP1	CSRNP1_1	AGAAGCGACGGCAAGCACTG
CSRNP1	CSRNP1_2	AGCCCGCTTCAGGATAGACA
CSRNP1	CSRNP1_3	GGCACTGCGCCAATCCCGGG
CSRNP1	CSRNP1_4	TAAATTCCACACGGCCCATG
EGR1	EGR1_1	AAGGCCTTAATAGTAGACAG
EGR1	EGR1_2	ACGCCCTTACGCTTGCCCAG
EGR1	EGR1_3	CGGCCAGTATAGGTGATGGG
EGR1	EGR1_4	GAAAATGTCAGTGTTCGGCG
EGR2	EGR2_1	CTCCGTTCATCTGGTCAAAG
EGR2	EGR2_2	CTTTGACCAGATGAACGGAG
EGR2	EGR2_3	GGGAAAGATGGTCACCGACG
EGR2	EGR2_4	GTCTGACAACATCTACCCGG
EGR3	EGR3_1	AAAGACAAGCAGATCCACCC
EGR3	EGR3_2	GAAGGCGAACTTTCCCAAGT
EGR3	EGR3_3	TACCACCACCCCAACGACAT
EGR3	EGR3_4	TGGCCGATTGGTAATCCTGG
EGR4	EGR4_1	CGGGTTCGGAAAACTCGCTA
EGR4	EGR4_2	GGCCTAAGATGCCCGACATG
EGR4	EGR4_3	GGGAAAACGAGCCTCCGGGG
EGR4	EGR4_4	TCCGGGGAGTAAAGGTCCGG
ELF4	ELF4_1	ACATGAACTATGAGACAATG
ELF4	ELF4_2	ATTGGGACCGTCGCTAGACG
ELF4	ELF4_3	CTCGCACACCATGTCAACCG
ELF4	ELF4_4	GCAGACGATCCCACTGACCA
EN2	EN2_1	AAGACGCTCTCGCTGCACGG
EN2	EN2_2	CCGAGTCCGAGCTCACCGAC
EN2	EN2_3	CGCGTACAGTAGACCCACGC
EN2	EN2_4	CGGTAGCAGCCCGGGCGAAG
ETV5	ETV5_1	ACCAGTATCCATCAGAACAG
ETV5	ETV5_2	GAATCGACGCAGTAATCCCG
ETV5	ETV5_3	GTGTCGTCAAAGTATAATCG

ETV5	ETV5_4	TAGGTTCTGACATTTGCCGA
FOS	FOS_1	GCTGACTGATACACTCCAAG
FOS	FOS_2	GGAAAAACTAGAGTTCATCC
FOS	FOS_3	GTAGTAAGAGAGGCTATCCC
FOS	FOS_4	GTCGAGATGGCAGTGACCGT
FOSB	FOSB_1	CGAAGACAGATATTGAGACT
FOSB	FOSB_2	GCTGACCGACCGACTCCAGG
FOSB	FOSB_3	GTCACCCCGTAGGAGTGCGC
FOSB	FOSB_4	TCGTAGGGGTCGACGACCGG
FOSL1	FOSL1_1	AGTACAGCCCCCCACAACCC
FOSL1	FOSL1_2	CCAAGCATCAACACCATGAG
FOSL1	FOSL1_3	TATTCCTTAGAAGTTCCACC
FOSL1	FOSL1_4	TGGTACAGCCTCATTTCCTG
FOSL2	FOSL2_1	AGGAGAAGCGTCGCATCCGG
FOSL2	FOSL2_2	AGTGTGCAAGATTAGCCCCG
FOSL2	FOSL2_3	CTGAGCCAGGCATATCTACC
FOSL2	FOSL2_4	GATCACGCCAGGTCTTGGGA
ID4	ID4_1	ACGACTGCTATAGCCGCCTG
ID4	ID4_2	CAGGCACAGCGCCGGCTCGT
ID4	ID4_3	CGCAGCCCGACGGCGCCTTG
ID4	ID4_4	GCACGTTATCGACTACATCC
JUNB	JUNB_1	CACAGCTACGGGATACGGCC
JUNB	JUNB_2	CCGGAGTCTCAAAGCGCCTG
JUNB	JUNB_3	CTGAGGTTGGTGTAAACGGG
JUNB	JUNB_4	GGGTAAAAGTACTGTCCCGG
KLF10	KLF10_1	AAGAACCCACCTAAATGTTG
KLF10	KLF10_2	ACCAGTATCTGATTTGTCAG
KLF10	KLF10_3	ATGGGCACACAAGTCCCCAA
KLF10	KLF10_4	GAACGACTGTTGTCACAACA
MXD1	MXD1_1	ACAGAGATGCCTTAAAACGG
MXD1	MXD1_2	AGGCGGCCGACTATCTGGAG
MXD1	MXD1_3	CGACTTGATTCGGGTCCAAG
MXD1	MXD1_4	GGGCATTGAGAGGATCCGGA
NPAS2	NPAS2_1	ATTATCGCAGTGACAACAGA
NPAS2	NPAS2_2	GTACCTTTAGTGCTGAGGAG
NPAS2	NPAS2_3	GTGGTAGTAGTCATAGCCTG
NPAS2	NPAS2_4	TGGTGTTGCCAGACGAACGG
NR4A1	NR4A1_1	GGCTAACAAGGACTGCCCTG
NR4A1	NR4A1_2	GTCCAGGTGTGCACGGACCA
NR4A1	NR4A1_3	TACACCCGTGACCTCAACCA
NR4A1	NR4A1_4	TCGCCCAGCCAGACTTACGA
RAF-transgene	raf-transgene_1	CGTGCCAGCACAAAGAGAGC
RAF-transgene	raf-transgene_2	GTGATGCTGTCCACTCGGAT

RAF-transgene	raf-transgene_3	TGGCTGTTCTGCGCAAAACA
RAF-transgene	raf-transgene_4	GATCAGATCATCTTCATGGT
TCF7	TCF7_1	AAGGCCAATCAGCCCCCCCA
TCF7	TCF7_2	AAGGCCCCGGAGTGCACCAG
TCF7	TCF7_3	CAGAACCTAGCATCAAGGAT
TCF7	TCF7_4	TCGCTCGTGAACGAGTCCGA
ZNF26	ZNF26_1	CAGAGACTCTATAACACACG
ZNF26	ZNF26_2	CTATCATAACCTGATATCAG
ZNF26	ZNF26_3	GTTCTCTGGTGTACAACGAA
ZNF26	ZNF26_4	TATGCATTCGGATGTGAACA
NTC	NTC-1	AGTATGAGACTCATAGGGTG
NTC	NTC-2	ATGTAACGAGTTGTAAGTCA
NTC	NTC-3	AGAAGACCGAGGCGCTTCAA
NTC	NTC-4	CATTAGCAGCCCAGCGCCCA
NTC	NTC-5	CGGGATGGTCCCTGCCGAGA
NTC	NTC-6	GCCATTCTAGTCCCGGCATA
NTC	NTC-7	GGTGCGGTCTGCTTGTAGAT
NTC	NTC-8	TATAGACCTCCCAAATACAT
NTC	NTC-9	TTAGCCAGTAGTGCATATGA
NTC	NTC-10	GCTGTTCCGAAGTTGAGAAT
SAFE_CUTTER	SAFE_CUTTER-1	GCCATTAATCATGATCTGGA
SAFE_CUTTER	SAFE_CUTTER-2	GACAATAGAAATTCGAAATG
SAFE_CUTTER	SAFE_CUTTER-3	GAAATCTAAAGATGTACACT
SAFE_CUTTER	SAFE_CUTTER-4	GTCCATGTATTTGACCATCA
SAFE_CUTTER	SAFE_CUTTER-5	GCGTAGTACTGTGGGGTAAA
SAFE_CUTTER	SAFE_CUTTER-6	GGTTTGAGAAGCTCACTTAC
SAFE_CUTTER	SAFE_CUTTER-7	GTTCTGAGCATTCTTAGTCT
SAFE_CUTTER	SAFE_CUTTER-8	GGCCTCTGTAAGAGGGGAGT
SAFE_CUTTER	SAFE_CUTTER-9	GTGTGGTGGAGAGTGAGTCC
SAFE_CUTTER	SAFE_CUTTER-10	GCTGATGGAAGCTTTGCCAA

2.9.3 Supplementary Data

Supplementary Data 4: Proliferation screen read count table and MAGeCK MLE values.

Supplementary Data 5: Modified TAP-seq targets inner primers for gene expression library amplification.

Ghanem El Kassem

Education:

- **M. Sc. in Pharmaceutical Biotechnology** 10.2014 – 08.2017
  Martin-Luther-Universität Halle-Wittenberg**,** Halle (Saale), Germany
  Master thesis performed at Fraunhofer-IZI Institute, Halle (Saale)
- **Bachelor of Pharmacy** 09.2007 – 06.2012
  Beirut Arab University, Beirut, Lebanon
- **Lebanese Baccalaureate (Life Science Division)**
  06.2007
  English International School (EIS), Tyre, Lebanon

Work Experience:

- **Universitätsmedizin Halle (Saale)** 11.2019-Present
  Research Assistant
  -Designing and implementing pooled CRISPR screens, emphasizing combinatorial gene perturbation and single-cell CRISPR sequencing (scCRISPRseq) screens.
  -Development and optimization of tools for combinatorial gene perturbation screens.
  -Validation of pooled and single-cell CRISPR screening target hits.
  -Analysis of NGS sequencing data from CRISPR screens.
  -Supervision of students and lab organization.

- **BioNTech Delivery Technologies GmbH (Formerly Lipocalyx GmbH)** 11.2017 – 10.2019
  Halle (Saale), Germany
  Research Assistant
  -Planning, execution, and evaluation of experiments on the formulation of nanoparticles and complexes.
  -Design, development, documentation, and technical distribution of nucleic acid carrier reagents.
  -Preparation of reports and documentation, especially in the context of the QM system.
  -Products developed: Viromer pDNA, Viromer mRNA, Viromer CRISPR, Viromer *in vivo* HQ, and Viromer Cytostain.

- **American University of Beirut Medical Center (AUBMC)** 04.2013 – 09.2014
  Hospital Pharmacist

- **Wardieh Pharmacy, Beirut, Lebanon** 06.2012 – 03.2013
  Pharmacist

Skills:

- **Languages:** Arabic (Native), English (Fluent), German (Fluent), Russian (Basic).
- **Computer Skills:** Microsoft Office, scientific writing, utilizing scientific databases, R Programming Language, NGS Data Analysis, Command Line Tools, scCRISPRseq data analysis.
- **Laboratory methods:** CRISPR screening, tumor organoid culture, flow cytometry, nanoparticle formulation, biochemical and molecular biology techniques.

Interests:

- Technology and history

Ghanem El Kassem