

A NUCLEAR PHOSPHATASE-KINASE
SIGNALING COMPLEX THAT SUPPORTS
ACUTE MYELOID LEUKEMIA

A Dissertation Presented to
Cold Spring Harbor Laboratory School of Biological Sciences
at Cold Spring Harbor Laboratory

in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

by
Sofya Alekseyevna Polyanskaya

Cold Spring Harbor Laboratory
July 2021

*Dedicated to my beloved parents Irina and Aleksey Polyanskiy
and to my late grandfather Gennadiy Nikolaevich Polyanskiy.*

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to everyone without whom this dissertation would not have been possible. First and foremost, I wish to thank my research advisor, Christopher Vakoc, for his guidance throughout this stage of my career. Chris was an enthusiastic and caring mentor who showed great belief in my abilities and gave me support and encouragement at every step. His breadth of knowledge, scientific intuition, and impeccable work ethic have been tremendous help and inspiration.

Next, I thank my Thesis Committee members, Lloyd Trotman, David Tuveson, Alexander Krasnitz, and Nicholas Tonks. They provided me with invaluable feedback, stimulating suggestions, and insightful advice during these years, and their broad and deep scientific expertise has been instrumental for the project development.

I further thank the former and current administration of the Cold Spring Harbor Laboratory School of Biological Sciences: Alexander Gann, Alyson Kass-Eisler, Monn Monn Myat, Kimberly Creteur, Kimberley Graham, and Carie Cowan. You make the School an exceptional place, a safe harbor where young scientists can flourish, mature, and confidently set sail into the stormy seas of life ahead.

I am grateful to all the past and present members of the Vakoc laboratory I overlapped with. I was lucky to share the bench with some of the most passionate, driven, hard-working scientists there are. Almost every technique in this dissertation was taught to me by my colleagues, and my project would not be the same without the many discussions we had. Special thanks go to Joseph Milazzo, who is simply indispensable to the Vakoc laboratory's smooth and organized functioning. I want to additionally thank Bin Lu, Zhaolin Yang, Yiliang Wei, and Osama El Demerdash, who have directly contributed to this work. I thank Olaf Klingbeil for all the help with my project and fun conversations about science and life. Thanks to Martyna Sroka for all the banter, I hope to read your thesis soon!

ACKNOWLEDGMENTS

This work was enriched by our amazing collaborators from the laboratory of Yan Jessie Zhang at the University of Texas, Austin; and the laboratory of Mitchell Weiss at St. Jude Children's Research Hospital. I thank Rosamaria Moreno and Seema Irani for the help with the *in vitro* biochemistry and Ruopeng Feng and Yu Yao for the work we did together at St. Jude's. Thanks to Mitchell Weiss for making my short research trip to Memphis, Tennessee, an enjoyable and unforgettable experience.

I want to thank the people with whom I shared many wonderful memories during the past six years. To Alexandra Nowlan and Craig Jones, who have generously opened their homes and their hearts to me. To Matthew Lee and Benjamin Berube for helping me fall in love with New York City and inviting me to some of the best nights out. To Tobiloba Oni for all the interesting conversations and fun times. To Tumi Ngoc Tran for all the invaluable life lessons and joyful memories of foods, arts, and Acadia.

Big thanks go to my oldest friend Anna Bojko for all the support and care she bestowed on me. Thank you for staying in touch over the years and distance.

A special place in my acknowledgments is reserved for my partner Thomas Sparks for sharing this journey with me. You have enriched my life in so many ways. I could only get to where I am with your love and support.

Finally, I want to thank my family. My deepest gratitude goes to my parents Irina and Aleksey Polyanskiy, for being the most amazing parents. I love you dearly. I thank my aunt and uncle Elena and Stuart Henshaw, and my cousin Alissa Polanski for rooting for me all these years. Thank you to my late grandmother Lidiya Rybkina and my grandfather Evgeniy Rybkin who always encouraged me to follow my path. This dissertation is dedicated to my parents and to my late grandfather Gennadiy Polyanskiy, who have dreamt of the day when I get my Ph.D. degree.

SUMMARY

Acute myeloid leukemia (AML) cells rely on phospho-signaling pathways to gain unlimited proliferation potential. A tightly regulated balance between phosphorylation and dephosphorylation, sustained by coordinated and competing activities of phosphatases and kinases, lies at the basis of cellular signaling. Given the advances of the past two decades in designing potent and selective inhibitors against phosphatases and kinases, these phospho-signaling enzymes present some of the priority targets in AML.

Here, we used domain-focused CRISPR screening to identify the nuclear phosphatase SCP4 as a dependency in AML. We provide evidence that this enzyme is likely dispensable in normal hematopoietic progenitor cells and could constitute a novel therapeutic target in leukemia. Using CRISPR exon scanning and gene complementation assays, we showed that the catalytic function of SCP4 was essential in AML. Our work, for the first time, elucidates a link between SCP4 and human cancer and provides a context for a deeper understanding of the molecular functions of this poorly studied phosphatase.

Through mass spectrometry analysis of the SCP4 interactome, we identify the kinase paralogs STK35 and PDIK1L as binding partners and substrates of the SCP4 phosphatase domain. STK35/PDIK1L signaling roles and biochemical interactions are largely unknown, and this study sheds light on their involvement in leukemogenesis. We showed that STK35 and PDIK1L catalytic activity was required in AML. Moreover, these kinases function redundantly in the same pathway as SCP4 to maintain leukemia cell proliferation.

We found that SCP4 regulated STK35/PDIK1L through two distinct mechanisms: promoting kinase stability and kinase catalytic activity. Our study provides the first genetic evidence that the conserved serine residue at the DFG+2 position could serve as the site of inhibitory phosphorylation at the kinases activation loop. Overall, our findings reveal a novel phosphatase-kinase signaling complex that supports the pathogenesis of AML.

CONTENTS

Acknowledgments.....	i
Summary	iii
Contents	1
List of Figures	5
List of Tables & Appendices.....	7
Abbreviations	8
Author Contributions.....	11
1. Introduction.....	12
1.1. Acute Myeloid Leukemia.....	12
1.1.1. Epidemiology	12
1.1.2. Etiology	13
1.1.2.1. De novo AML	14
1.1.2.2. Inherited predisposition.....	14
1.1.2.3. t-AML.....	15
1.1.2.4. Secondary AML.....	15
1.1.3. Ontogeny	16
1.1.3.1. Cytogenetic abnormalities	16
1.1.3.2. Molecular abnormalities	17
1.1.4. The two-hit hypothesis of leukemogenesis	17
1.2. Kinases and phosphatases in AML.....	24
1.2.1. Brief introduction to kinases and phosphatases	24
1.2.1.1. Classification of kinases and phosphatases	26
1.2.1.2. Evolution of kinases and phosphatases.....	27
1.2.2. Notable kinases and phosphatases in AML	32
1.2.2.1. FLT3.....	32
1.2.2.2. c-KIT.....	34
1.2.2.3. BCR-ABL1	35
1.2.2.4. JAK2	36
1.2.2.5. CDK6.....	37
1.2.2.6. SIK2/3 and MARKs.....	39
1.2.2.7. SHP2.....	40
1.2.2.8. PP2A.....	42

CONTENTS

1.2.2.9. CD45.....	44
1.2.2.10. PTP1B.....	45
1.2.2.11. PTPRJ.....	46
1.2.2.12. PPM1D.....	47
1.3. Current and emerging agents and regimens for AML.....	49
1.3.1. Intensive therapy.....	49
1.3.2. Gemtuzumab ozogamicin (GO).....	50
1.3.3. FLT3 inhibitors.....	51
1.3.4. IDH1/2 inhibitors.....	51
1.3.5. Hypomethylating agents (HMAs).....	52
1.3.6. Hedgehog (HH) pathway inhibitor: glasdegib.....	53
1.3.7. Bcl-2 inhibitor: venetoclax.....	54
1.4. The search for non-oncogene addictions.....	56
1.4.1. Examples of mechanisms for non-oncogene addictions.....	56
1.4.1.1. Genetic streamlining.....	56
1.4.1.2. Altered metabolism.....	57
1.4.2. Genetic screens for non-oncogene addictions.....	60
1.4.2.1. Before CRISPR.....	60
1.4.2.2. CRISPR-Cas9.....	61
1.5. HAD phosphatases in cancer.....	63
1.5.1. EYA phosphatases.....	66
1.5.2. PDXP/chronophin.....	67
1.5.3. Other.....	68
1.6. Current state of research on the topic of this dissertation.....	70
1.6.1. SCP4.....	70
1.6.2. STK35/PDIK1L.....	73
2. Results.....	75
2.1. Phosphatase domain-focused CRISPR screening identifies context-specific dependencies in human cancer cell lines.....	75
2.2. SCP4 is an acquired dependency in human AML cells.....	81
2.3. SCP4 is dispensable in normal hematopoietic progenitor cells.....	89
2.4. The catalytic phosphatase activity of SCP4 is essential in AML.....	95
2.5. N-terminus of SCP4 promotes nuclear localization of the full-length protein.....	100
2.6. STK35 and PDIK1L bind selectively to the catalytically active form of SCP4.....	106

CONTENTS

2.7.	STK35 and PDIK1L function redundantly in the same genetic pathway as SCP4 ..	110
2.8.	Biochemical evidence that SCP4 functions upstream of STK35/PDIK1L.....	121
2.9.	SCP4-STK35/PDIK1L sustains the aberrant metabolic state of AML	130
2.10.	Results summary	136
3.	Discussion and Perspectives	139
3.1.	The novelty	139
3.2.	Dissecting the roles for the N-terminus in SCP4 regulation.....	141
3.3.	Phosphatase/kinase complexes.....	143
3.4.	Phospho-regulation of eukaryotic kinases	147
3.5.	AML metabolism.....	149
4.	Future directions.....	156
4.1.	Immediate degradation.....	156
4.2.	The role of SCP4 on chromatin.....	157
4.3.	Understanding the phenotype	158
4.4.	Identifying and validating SCP4 substrates.....	160
4.5.	Identifying kinases substrates	161
4.6.	Structural and biochemical studies of purified proteins.....	162
4.7.	Therapeutic implications	163
5.	Materials and Methods	165
5.1.	Cell lines	165
5.2.	Plasmid construction and sgRNA cloning	165
5.3.	Construction of sgRNA libraries.....	166
5.4.	Lentivirus production and infection	167
5.5.	Pooled negative-selection CRISPR screening and data analysis.....	168
5.6.	Competition-based cell proliferation assays.....	169
5.7.	Western Blot	170
5.8.	Cell cycle arrest and apoptosis analysis	171
5.9.	Editing and differentiation of human peripheral blood CD34+ cells.....	171
5.10.	In vivo transplantation of MOLM-13 cells into NSG mice	173
5.11.	Sub-cellular fractionation and Immunoprecipitation in MOLM-13 cells	174
5.12.	Protein identification by mass spectrometry (MS).....	175
5.13.	Co-IP in HEK293T cells	176
5.14.	RNA-Seq.....	177
5.15.	RNA-Seq data analysis.....	178

CONTENTS

5.16. Protein Expression and Purification.....	179
5.17. Phosphatase activity assay	180
5.18. Metabolomics analysis by liquid chromatography coupled to mass spectrometry (LC-MS)	181
5.19. Quantification and statistical analysis	182
5.20. Data and software availability	182
References	183
Appendices	211

LIST OF FIGURES

Figure A. Simplified hematopoiesis and AML origin.	20
Figure B. AML statistics.....	21
Figure C. History of AML therapies.	55
Figure D. Catalytic mechanism for HAD phosphatases.	65
Figure 1. Phosphatase domain-focused CRISPR screen library design and negative selection ‘dropout’ screen schematics.	77
Figure 2. The summary of dependencies discovered in the phosphatase domain-focused CRISPR screens.	78
Figure 3. The performance of spike-in negative and positive control sgRNAs.	79
Figure 4. Phosphatase domain-focused CRISPR screening identifies context-specific dependencies in human cancer cell lines.	80
Figure 5. SCP4 is an AML-biased dependency.	83
Figure 6. The differential requirement for SCP4 in leukemia cells.	84
Figure 7. Validation of the on-target effect of SCP4 knockout.....	85
Figure 8. SCP4 inactivation in MOLM-13 cells leads to a G1/G0-arrest and the induction of apoptosis.	86
Figure 9. Targeting SCP4 attenuates the growth of MOLM-13 cells in xenograft mouse model of AML.	87
Figure 10. AML-biased dependency on SCP4 is conserved across species.	88
Figure 11. SCP4-depleted HSPCs did not experience any fitness disadvantage during 16 days of culturing under conditions that promote myeloid, erythroid, or megakaryocytic differentiation.	92
Figure 12. SCP4-depleted HSPCs showed no defect in differentiation into myeloid, erythroid, or megakaryocytic lineages.	93
Figure 13. SCP4-depleted HSPCs showed no differences in relative representation of the colony-forming units.....	94
Figure 14. Defining functionally important regions of SCP4.....	97
Figure 15. SCP4236-466 is sufficient to support AML cell proliferation.....	98
Figure 16. SCP4 catalytic function is necessary to support AML cell proliferation.....	99
Figure 17. SCP4 is localized to the nucleus.	102
Figure 18. Optimization of SCP4 recovery from nuclear fractions.	103

LIST OF FIGURES

Figure 19. IP-MS identification of FLAG-SCP4 associated proteins.....	104
Figure 20. SCP4 N-terminus promotes nuclear localization of the full-length protein.	105
Figure 21. STK35 and PDIK1L bind selectively to the catalytically active form of SCP4.....	108
Figure 22. STK35 and PDIK1L exist in a stable complex with the endogenous SCP4.	109
Figure 23. STK35 and PDIK1L function redundantly to support MOLM-13 proliferation.	113
Figure 24. Either STK35 or PDIK1L is sufficient to fully rescue the endogenous STK35/ PDIK1L double knockout phenotype in MOLM-13 cells.....	114
Figure 25. STK35/PDIK1L double knockout in MOLM-13 cells leads to a G1/G0-arrest.	115
Figure 26. STK35/PDIK1L requirement is correlated with the SCP4 dependency in the context of AML.....	116
Figure 27. The kinase catalytic function of STK35/PDIK1L is essential in AML.....	117
Figure 28. STK35 and PDIK1L function redundantly in the same genetic pathway as SCP4.	118
Figure 29. STK35/PDIK1L dual perturbation significantly affects SCP4 signatures.	120
Figure 30. Biochemical evidence that SCP4 functions upstream of STK35/PDIK1L.....	126
Figure 31. Schematics of STK35/PDIK1L phosphorylation sites.....	127
Figure 32. S194 on PDIK1L and S385 on STK35 could serve as inhibitory phosphorylation residues.	128
Figure 33. SCP4 exhibits exquisite substrate specificity in vitro.	129
Figure 34. Nominating the downstream output of the SCP4-STK35/PDIK1L complex.	132
Figure 35. SCP4 knockout and PDIK1L-STK35 double knockout result in consistent changes in the leukemia cell metabolism.	133
Figure 36. SCP4-STK35/PDIK1L signaling supports the aberrant metabolic state in AML. ...	135
Figure 37. Model for SCP4-STK35/PDIK1L cooperativity in supporting the aberrant metabolic state in AML.	138
Figure S1. Transcriptional downregulation of the myeloid protease cathepsin G in response to SCP4 knockout and STK35/PDIK1L double knockout.	152
Figure S2. Catalytic mutants of SCP4, but not the wild-type SCP4, interact with FBXW11. ..	153
Figure S3. Predicted three-dimensional structures for the catalytic domains of PDIK1L and STK35.....	154
Figure S4. Gluconeogenesis genes upon SCP4 knockout and STK35/PDIK1L double knockout.....	155

LIST OF TABLES & APPENDICES

Table A. National Comprehensive Cancer Network (NCCN) cytogenetic/ molecular classification of AML	22
Table B. Functional Categories of Genes Commonly Affected in AML (according to the TCGA dataset)	23
Table C. Classification of human kinases	29
Table D. Classification of human phosphatases	30
Table 1. Kinetic constants of different phosphatases towards their substrates	125
Appendix 1. Categorization of human phosphatases targeted by CRISPR-Cas9 genetic screens in this study	212
Appendix 2. Average log ₂ FC in phosphatase domain-focused screen.....	213
Appendix 3. Average log ₂ FC in SCP4 CRISPR exon scanning screen in MOLM-13	220
Appendix 4. Genes downregulated upon SCP4 knockout	223
Appendix 5. Genes upregulated upon SCP4 knockout	224
Appendix 6. sgRNA sequences.....	225
Appendix 7. Primers for the generation of CRISPR-resistant constructs and introduction of the point mutations.....	226
Appendix 8. sgRNAs; media and cytokines; and antibodies used in flow cytometry panels in CD34+ experiments	228

ABBREVIATIONS

α KG	alpha-ketoglutarate
ACHD	age-related clonal hematopoiesis disorder
AHD	antecedent hematological disorder
AID	auxin-inducible degron
ALL	acute lymphoblastic leukemia
ALV	avian leukosis virus
AML	acute myeloid leukemia
APE	the conserved motif at the end of the activation segment of eukaryotic protein kinases comprising alanine (A), proline (P), and glutamate (E) residues
AS	analog-sensitive [kinase mutants]
ATAC-seq	assay for transposase-accessible chromatin coupled with DNA sequencing
ATP-BS	ATP [adenosine triphosphate] binding site
BCAA	branched-chain amino acid
Cas9	CRISPR associated protein 9
CBF-AML	core binding factor acute myeloid leukemia
CFC	colony-forming cell [assay]
ChIP-qPCR	chromatin immunoprecipitation coupled with qPCR
ChIP-seq	chromatin immunoprecipitation coupled with DNA sequencing
CLL	chronic lymphocytic leukemia
CML	chronic myelogenous leukemia
Co-IP	co-immunoprecipitation
CpG	regions of DNA where a cytosine nucleotide is followed by a guanine nucleotide in the linear sequence of bases along its 5' \rightarrow 3' direction.
CR	CRISPR resistant
CRISPR	clustered regularly interspaced short palindromic repeats
CTD	carboxy terminal domain
DFG	the conserved motif at the start of the activation segment of eukaryotic protein kinases comprising aspartate (D), phenylalanine (F), and glycine (G) residues
dgRNA	bi-cistronic vector that encodes two sgRNAs to be expressed in the cell simultaneously upon lentiviral infection
DNA	deoxyribonucleic acid
ED	extracellular ligand-binding domain
EV	empty vector
FC	fold change
FL	FLT3 ligand

ABBREVIATIONS

FLT3-ITD	FLT3 internal tandem duplication
FLT3-TKD	FLT3 tyrosine kinase domain [point mutation]
FWER	familywise error rate
GFP	green fluorescent protein
GO	gene ontology
GSEA	gene set enrichment analysis
HAD	haloacid dehalogenase[-like]
HH	Hedgehog
HMA	hypomethylating agent
HMMS	hereditary myeloid malignancy syndrome
HOX	homeobox [genes]
HSPC	hematopoietic stem/progenitor cell
indel	an insertion or deletion of bases into the genomic DNA
IP	immunoprecipitation
JMD	juxtamembrane domain
KDM	histone lysine demethylase
KMT	histone lysine methyltransferase
LC-MS	liquid chromatography-mass spectrometry
LSC	leukemia stem cell
MA9	MLL-AF9
mAb	monoclonal antibody
mCherry	red fluorescent protein
MDS	myelodysplastic syndrome
MPN	myeloproliferative neoplasm
MS	mass spectrometry
ND	not detected
NES	normalized enrichment score
NLS	nuclear localization signal
NT	non-targeting
OXPHOS	oxidative phosphorylation
PAD	PP2A-activating drug
PAM	protospacer adjacent motif
PDAC	pancreatic ductal adenocarcinoma
PDX	patient-derived xenograft
PLP	pyridoxal 5'-phosphate

ABBREVIATIONS

Pol II	RNA polymerase II
PTP	protein tyrosine phosphatase
qPCR	quantitative polymerase chain reaction
R-2-HG	R-2-hydroxyglutarate
R/R	relapsed or refractory
RMS	rhabdomyosarcoma
RNA	ribonucleic acid
RNA-Seq	RNA sequencing
RNAi	RNA interference
RNP	ribonucleoprotein
ROS	reactive oxygen species
RT-qPCR	real-time quantitative polymerase chain reaction
RTK	receptor tyrosine kinase
SCLC	small cell lung cancer
SEER	Surveillance, Epidemiology, and End Results [Program]
SFK	Src family protein tyrosine kinase
sgRNA	single guide RNA
t-AML	therapy-related acute myeloid leukemia
TCA	tricarboxylic acid
TKD	tyrosine kinase domain
TKI	tyrosine kinase inhibitor
TPM	transcripts per million
TSC	total spectrum counts
WT	wild type

Formatting of gene and protein names in this thesis

Avian gene names	lowercase, italicized (<i>scp4</i> , alias <i>ctdspl2</i>)
Avian protein names	lowercase (scp4)
Murine gene names	capitalized, italicized (<i>Scp4</i> , alias <i>Ctdspl2</i>)
Murine protein names	capitalized (Scp4)
Human gene names	uppercase, italicized (<i>SCP4</i> , alias <i>CTDSPL2</i>)
Human protein names	uppercase (SCP4)

AUTHOR CONTRIBUTIONS

I would like to acknowledge and once again thank everyone who contributed to this work. Rosamaria Moreno, Seema Irani, and Yan Jessie Zhang purified, aided characterization, and measured phosphatase activity of SCP4, presented in Fig. 18B and Fig. 33. Bin Lu performed and analyzed SCP4 CRISPR exon scan, presented in Fig. 14. The editing in human CD34⁺ cells was executed by me together with the help and guidance from Ruopeng Feng and Yao Yu at the laboratory of Mitchell Weiss. These data are presented in Fig. 11–13. Zhaolin Yang performed and analyzed cell cycle arrest and apoptosis experiments, presented in Fig. 8 and Fig. 25. Yiliang Wei conducted the mouse xenograft experiments, presented in Fig. 9A. I have additionally disclosed these contributions in the corresponding figure legends.

Lukas Benjamin performed a few competition assays replicates in Fig. 5 to aid statistical significance and reproducibility. Quantification of bioluminescence intensity in Fig. 9B was done by me with the help of Scott Lyons from the Animal Imaging & Tissue Imaging Shared Resource. Olaf Klingbeil and Osama El Demerdash assisted me with the RNA-Seq data analysis. Olaf Klingbeil provided me with the bi-cistronic vector for the double knockout. Michael Lukey gave me helpful feedback and comments that were reflected in Fig. 34 and Fig. 36.

The mass spectrometry protein identification from affinity purification experiments was performed by the Proteomics & Metabolomics Facility at the Center for Biotechnology/University of Nebraska-Lincoln, specifically by Mike Naldrett and Sophie Alvarez. The metabolites mass spectrometry measurements were done by the CSHL Mass Spectrometry Shared Resource supported by the Cancer Center Support Grant 5P30CA045508, specifically by Sofia Costa.

I thank CSHL Functional Genomics Core Facility, in particular Kenneth Chang, for helpful discussions during phosphatase library design and cloning; CSHL Next Generation Sequencing Core Facility for deep sequencing; CSHL Bioinformatics Shared Resource for providing the Galaxy data analysis platform; and Hillside shared resources ran by Amy Brady and Julie Cheong.

The data and corresponding text presented in this thesis are part of the following preprint:

Polyanskaya SA, Moreno RY, Lu B, Feng R, Yao Y, Irani S, Klingbeil O, Yang Z, Wei Y, Demerdash OE, Benjamin LA, Weiss MJ, Zhang YJ, Vakoc CR. 2021. The SCP4-STK35/PDIK1L complex is a dual phospho-catalytic signaling dependency in acute myeloid leukemia. *bioRxiv* doi: <https://doi.org/10.1101/2021.05.09.443327>.

1. INTRODUCTION

1.1. Acute Myeloid Leukemia

This dissertation has uncovered a novel dependency in a deadly cancer of the blood and bone marrow called Acute Myeloid Leukemia. The following chapter outlines our current understanding of this disease and the rationale for focusing our research efforts on the enzymes involved in cell signaling.

Acute Myeloid Leukemia (AML) is an aggressive hematologic malignancy caused by the impaired differentiation of myeloid progenitor cells, which results in the unrestrained proliferation of early immature clones, called myeloblasts, or leukemic blasts (Fig. A). 20% blasts in the complete blood count or bone marrow differential is generally required for AML diagnosis. The accumulation of aberrant myeloblasts in the bone marrow prevents its normal functions, leading to the associated clinical presentations, such as bleeding, anemia, and infections.

1.1.1. Epidemiology

There are about 20,000 new AML cases diagnosed in the United States every year – this is around 1% of all the estimated new cancer cases, according to the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) Program (Fig. B). AML has one of the lowest 5-year relative survival rates of 29.5% (2011–2017), fifth after the pancreas (10.8%), esophagus (19.9%), liver (20.3%), and lung (21.7%) cancers. AML is more common in older adults, with incidence rates dramatically increasing after the age of 45 and the median age at diagnosis of 68 years old. There are considerable differences in AML survival by age that, together with the prevalence of AML in older patients, contribute to the low

five-year overall survival rate. Additionally, AML incidence is modestly (1.2 – 1.6) higher in males than in females, with the difference first detectable after 50 and increasing with age (Shallis et al. 2019).

Most commonly used AML therapies, including intensive chemotherapy and stem cell transplant, are more likely to benefit a minority of younger and otherwise fit patients, whereas most older patients exhibit poor prognoses. Population-based statistics from the SEER program report 5-year relative survival of 63%, 39.6%, 15.5%, and 3.7% for patients aged below 39, 40 – 64, 65 – 74, and 75 or older, respectively (2011–2017). There are several reasons for this disparity in disease outcomes. First, older patients are more likely to have secondary AML (see below), linked with lower complete remission rates and decreased disease-free survival (Lazarevic 2021). Second, AML in older patients is commonly associated with unfavorable cytogenetic and molecular abnormalities that cause resistance to standard chemotherapy and poor recovery of normal hematopoiesis (Lazarevic 2021). Lastly, older patients are generally presented with more comorbidities and organ dysfunctions that could increase treatment-related mortality and non-receipt of any anti-leukemic therapy (Medeiros et al. 2015). The differences between AML biology and pathophysiology in older and younger patients illustrate the notion of AML as an umbrella diagnosis comprising different subtypes with their respective prognostic features and treatment options, which are being continuously optimized as we enter the era of precision medicine (Kantarjian et al. 2021).

1.1.2. Etiology

AML can be classified into *de novo* AML and secondary AML. Secondary AML can evolve from untreated or treated antecedent hematological disorders (AHDs) that are often undiagnosed, or be therapy-related (t-AML), accounting for 7–15% of all AML cases.

1.1.2.1. *De novo* AML

De novo AML is referred to any case of AML not arising from germline predisposition, prior treatment, or prior myeloid malignancy. While as many as two-thirds of all diagnosed AML cases are described as *de novo* AML, this concept fails to account for poor diagnosis for AHDs and non-therapeutic exposures such as tobacco smoking and ambient presence of carcinogens, including benzene, formaldehyde, and pesticides (Shallis et al. 2020). Noteworthy, a gene mutations profiling study showed differences between *de novo* AML and t-AML. *De novo* AML was characterized by higher frequencies of mutations in *FLT3* and *NPM1* and lower in *TP53* and *PTPN11*, compared to t-AML (Ok et al. 2015). Although some of these observations did not reach statistical significance, this and similar studies provide the rationale for further research concerned with differences in clonal representation between *de novo* and secondary AML.

1.1.2.2. Inherited predisposition

The incidence of AML is increased in patients with specific genetic syndromes such as Fanconi's anemia, Down syndrome, Bloom syndrome, and others (Alter 2014; Caldwell et al. 2014; Bochtler et al. 2015). Physicians also distinguish genetically defined hereditary myeloid malignancy syndromes (HMMSs). One of the first HMMSs was discovered in 1999 by the presence of inherited mutations in *RUNX1* (Song et al. 1999). It was closely followed by the report on the incidence of familial AML with *CEBPA* mutation in 2004 (Smith et al. 2004). The accessibility of next-generation sequencing led to the discoveries of other AML heritable predispositions, including mutations in *GATA2* (Ostergaard et al. 2011) and *ETV6* (Topka et al. 2015). Early identification of genetic predisposition to AML could benefit patients via implementation of preventive therapies aimed at the currently present abnormal cells before cancer evolution results in clones conferring therapy resistance.

1.1.2.3. t-AML

The first reports of the increased leukemogenesis risk following primary cancer regimens appeared in the literature in the 1960s for patients treated with alkylating agents (such as nitrogen mustard). Old combination chemotherapies generally posed a higher risk for t-AML development, and this risk sometimes declined with the introduction of new regimens (Schonfeld et al. 2006).

t-AML is about twice as likely to be associated with poor-risk cytogenetics than *de novo* AML (Ok et al. 2015). While initially it was thought that t-AML was rooted in DNA damage induced by prior exposure to chemotherapy or radiotherapy, another possibility is that the exposure to these regimens promotes the selection of pre-existing mutant clones rather than leading to new mutations (McNerney et al. 2017). One argument in favor of this hypothesis would be the statistically significant increase in the rate of clonal hematopoiesis in solid tumor patients post-radiotherapy as compared to the general population (Coombs et al. 2017). Additionally, immunosuppressive therapies for organ transplant recipients were shown to allow t-AML development in the absence of direct DNA damage, supporting the model in which clonal expansion may arise due to abnormal cells escaping weakened immune surveillance (Offman et al. 2004).

1.1.2.4. Secondary AML

Secondary AML is diagnosed in conjunction with an AHD, either myelodysplastic syndrome (MDS) or myeloproliferative neoplasm (MPN) (Genovese et al. 2014; Jaiswal et al. 2014). AHDs can arise from age-related clonal hematopoiesis disorders (ACHD). More than 75% of ACHDs are characterized by recurrent somatic mutations in one of the three genes, *DNMT3a*, *ASXL1*, or *TET2* (Genovese et al. 2014). Additionally, mutations in *TP53*, *JAK2*, or genes of the spliceosome machinery (*SF3B1*, *SRSF2*, and *U2AF1*) were also linked to an

increased risk of developing AML (Jaiswal and Ebert 2019). The abnormal clones still have the differentiation capacity but acquire a proliferative advantage leading to their expansion.

The persistence of mutant clones may be conducive to accumulating additional mutations over time, explaining why clonal hematopoiesis is associated with an elevated risk of developing hematologic malignancies. The prevalence of clonal hematopoiesis conditions increases with age, and about 10% of the population above 70 years old is estimated to carry leukemia-related somatic mutations (Jaiswal et al. 2014). It is important to note that most people with clonal hematopoiesis never develop hematologic cancer (Steensma 2018).

1.1.3. Ontogeny

1.1.3.1. Cytogenetic abnormalities

Clinical practice guidelines in oncology usually involve cytogenetic/molecular classification of AML (Table A) and its associated risks. The karyotype¹ and cytogenetic² abnormalities are present in 50–60% of AML cases and strongly correlate with response to therapies and overall survival. For example, t(15;17) (*PML-RARA*) is considered better risk cytogenetics due to the discovery of effective targeted therapy for it.

Patients above 60 years old are reported to have a higher proportion of poor-risk associated cytogenetics such as most *MLL* translocations (translocations involving 11q23) and inv(3) (*EVII* induction), and adverse karyotypes including deletions of chromosomes 5q, 7q, and 17p (*TP53* deletion) (Lazarevic 2021). Low-risk cytogenetic abnormalities such as t(8;21) (*RUNX1T1-RUNX1*) or inv(16) (*CBFB-MYH11*) are more common in younger patients (Tsai et al. 2016). Notably, the prognostic value of these genetic events depends on both age

¹The number and visual appearance of the chromosomes in the cell nuclei of an organism or species.

²Here: changes in chromosomes, including deletions, duplications, insertions, inversions, and translocations.

and etiology. For example, single translocation (9;11)(p22;q23) (*KMT2A-MLL3*) may pose an intermediate risk in younger patients with *de novo* AML (Bill et al. 2020).

1.1.3.2. Molecular abnormalities

While AML prognosis is defined mainly by the cytogenetic abnormalities among the favorable and unfavorable karyotype subsets, in the normal karyotype (i.e., diploid), somatic mutations are considered prognostic (Pastore et al. 2014). At a minimum, recent experts' recommendations include testing for *FLT3-ITD*, *FLT3-TKD*, *TP53*, *NPM1*, *RUNX1*, *ASXL1*, and *CEBPA* mutations (Dohner et al. 2017). Among frequently mutated genes (>5%) in AML, mutations in *FLT3*, *TP53*, *RUNX1*, and *ASXL1* are associated with poor outcomes identifying a molecular subgroup of high-risk AML. AML with *NPM1* or biallelic *CEBPA* mutations and normal karyotype are considered a better risk (Table A).

Similarly to *PML-RARA*, the predictive value of mutations in AML is altered with the advances in diagnosis and treatment. The normal karyotype AML with mutations in *NPM1* (50%) generally responds well to cytarabine-based regimens and to regimens containing the hypomethylating agents (HMAs) and venetoclax (Kantarjian et al. 2021). *FLT3* inhibitors midostaurin and gilteritinib have been approved by FDA and now became standard practice in *FLT3*-mutated AML. Since FDA approval of *IDH1/2* inhibitors, screening for *IDH1/2* mutations is now being considered in clinical practice, particularly at the time of relapse.

1.1.4. The two-hit hypothesis of leukemogenesis

Although AML etiological and cytogenetic/molecular classification is broadly used in clinical practice to inform treatment strategies and prognoses as well as case studies, considerable heterogeneity exists within every recognized subgroup. AML genetic alterations exhibit complex co-occurrence patterns or mutual exclusivity that result in

enormous combinatorial diversity in phenotypes and clinical outcomes. The continuous accumulation of mutations and subsequent clonal expansions dramatically increase AML clonal representation. The stepwise acquisition of mutations and association with clonal disorders consolidated the notion of AML as an excellent model for studying the principles of cancer evolution, especially with the single-cell sequencing technologies becoming more widely available (Grove and Vassiliou 2014; van Galen et al. 2019; Potter et al. 2019). Understanding AML pathways of clonal dominance and shifts will guide precision medicine and diagnosis, for example, at the level of the minimal residual disease (Moffitt et al. 2020). There is also a considerable need for identifying general principles utilized by AML cells for proliferation and targeted therapies resistance.

In 2002, with several co-authors, Gilliland proposed a so-called “two-hit” model of leukemogenesis, which still serves as a useful conceptual framework (Gilliland and Griffin 2002; Kelly and Gilliland 2002). They noticed that while the commonly mutated in AML constitutively activated kinases such as *FLT3*, *BCR-ABL1*, and *JAK2* are necessary and sufficient to cause the myeloproliferative clonal expansion phenotype, they are not enough for the differentiation block characteristic for AML. Additional mutations in transcription factors and epigenetic regulators are required for full-blown AML development. Gilliland and co-authors suggested the term “class I” mutations for the mutations conferring proliferative advantage to cells and “class II” for the mutations rewiring myeloid differentiation program. One of the arguments in favor of this hypothesis was the relatively low likelihood of two mutations from the same class co-existing together in the same cell, such as *FLT3* and *RAS* (class I) or *KMT2A (MLL)* and *DNMT3A* (class II) (Gilliland and Griffin 2002; Wang et al. 2017). Other recurrent mutation groups in AML discovered later (spliceosome genes, epigenetic modifiers, cohesins) are likely to synergistically produce similar results (Table B) (Grove and Vassiliou 2014).

Aberrant activation of cellular signaling pathways commonly results in increased proliferation; hence these mutations are frequently considered class I. Out of AML patients in the Cancer Genome Atlas (TCGA) cohort (median age, 55 years), 59% carry mutations in the genes involved in signal transduction pathways, which makes them the most commonly mutated group in AML (Table B) (Bullinger et al. 2017). Moreover, acquisition or enrichment of clones with activation of specific signaling pathways can result in the primary and adaptive resistance to the existing therapies (DiNardo et al. 2020b). A tightly regulated balance between phosphorylation and dephosphorylation, sustained by coordinated and competing activities of kinases and phosphatases, lies at the basis of cellular signaling. Given the advances of the past two decades in designing potent and selective inhibitors against kinases and phosphatases, these phospho-signaling enzymes present some of the priority targets in AML.

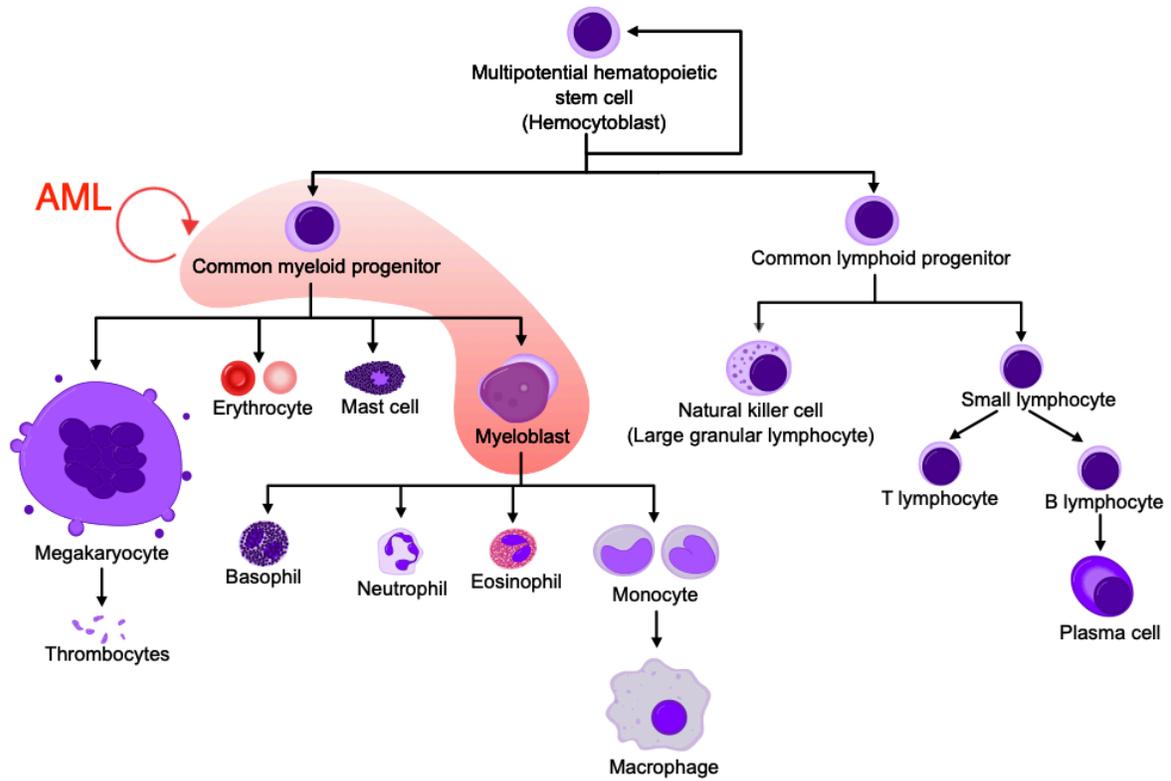


Figure A. Simplified hematopoiesis and AML origin.

Adapted from Rad and Häggström (CC-BY-SA 3.0 license).

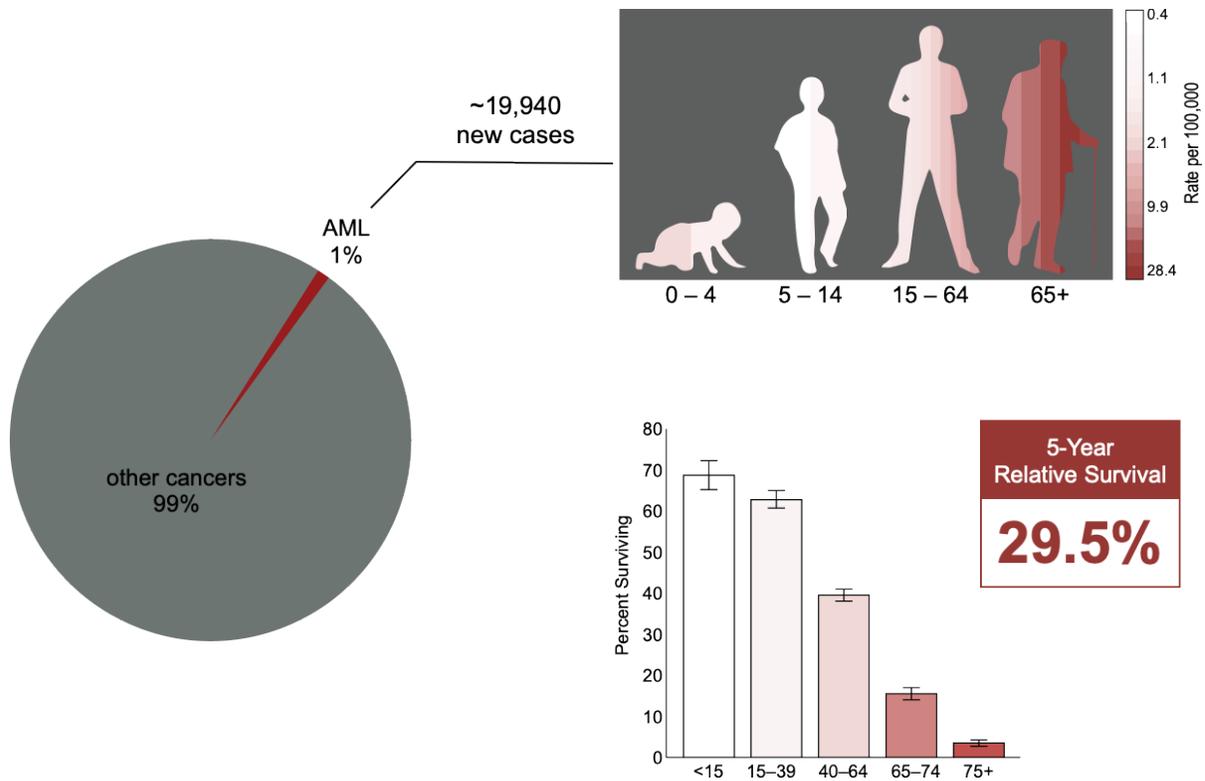


Figure B. AML statistics.

(A) The proportion of AML among all cancer cases diagnosed in the US in 2019. Data from the American Cancer Society, 2020. (B) Schematic representation of the incidence rate of AML by age groups at diagnosis. Data from the SEER, 2014–2018. (C) 5-Year relative survival rates by age groups at diagnosis. Data from the SEER, 2011–2017.

**Table A. National Comprehensive Cancer Network (NCCN)
cytogenetic/molecular classification of AML**

NCCN Classification	Cytogenetics	Gene fusion	Molecular Abnormalities
Better risk	t(15;17)(q24;q21) t(8;21)(q22;q22) inv(16)(p13;q22) t(16;16)(p13;q22)	<i>PML-RARα</i> <i>AML1-ETO</i> <i>CBFα-MYH11</i> <i>CBFα-MYH11</i>	Normal cytogenetics; <i>NPM1</i> mutation without <i>FLT3-ITD</i> ; isolated biallelic <i>CEBPα</i> mutation
Intermediate risk	t(9;11)(p22;q23) or normal cytogenetics; +8 alone; other undefined	<i>MLLT3-MLL</i>	t(8;21), inv(16), t(16;16) with <i>c-KIT</i> mutations; <i>NPM1</i> mutation with <i>FLT3-ITD</i> ^{high} ; <i>NPM1</i> wt without <i>FLT3-ITD</i> ; <i>NPM1</i> wt and <i>FLT3-ITD</i> ^{low}
Poor risk	t(9;22)(q34.1;q11.2) t(v;11q23.3) non-t(9;11) inv(3)(q21;q26.2) t(3;3)(q21;q26.2) t(6;9)(p23;q34) or complex karyotype (≥ 3 clonal chromosomal abnormalities); monosomal karyotype; -5 or del(5q) -7 or del(7q) abn(17p)	<i>BCR-ABL1</i> <i>MLL</i> -rearranged <i>GATA2, EVI1</i> <i>GATA2, EVI1</i> <i>DEK/NUP214</i>	<i>TP53</i> mutation; <i>RUNX1</i> mutation; <i>ASXL1</i> mutation; <i>NPM1</i> wt and <i>FLT3-ITD</i> ^{high}

inv, inversion; t, translocation; wt, wild-type; ^{low}, low allelic ratio (<0.5); ^{high}, high allelic ratio (≥ 0.5).

Table adapted from Kantarjian et al. 2021.

**Table B. Functional Categories of Genes Commonly Affected in AML
(according to the TCGA dataset)**

Functional Category	Selected Gene Members	Role in Leukemogenesis	TCGA Cohort (%)*
Signaling genes	Kinases (e.g., <i>FLT3</i> , <i>c-KIT</i>), phosphatases (e.g., <i>PTPN11</i>), or RAS family members (e.g., <i>KRAS</i> , <i>NRAS</i>)	Activated signaling confers a proliferative advantage through RAS/RAF, JAK/STAT, and PI3K/AKT signaling pathways.	59
DNA methylation-associated genes	<i>DNMT3A</i> , <i>TET2</i> , <i>IDH1/2</i>	Deregulated DNA methylation patterns lead to transcriptional deregulation of leukemia-relevant genes.	44
Myeloid TF gene fusions or mutations	TF fusions	Aberrant TF function results in transcriptional deregulation and impaired hematopoietic differentiation.	25
	TF mutations (<i>RUNX1</i> , <i>CEBPA</i>)		22
Chromatin-modifying genes	Mutations (e.g., <i>ASXL1</i> , <i>EZH2</i>) or <i>KMT2A</i> fusions	Deregulation of histone marks leads to transcriptional deregulation.	30
Nucleophosmin (<i>NPM1</i>) gene	<i>NPM1</i>	Aberrant cytoplasmic localization of <i>NPM1</i> and <i>NPM1</i> -interacting proteins.	27
Tumor suppressor genes	<i>TP53</i> , <i>WT1</i> , <i>PHF6</i>	Loss of tumor suppressors	16
Spliceosome-complex genes	<i>SRSF2</i> , <i>SF3B1</i> , <i>U2AF1</i> , <i>ZRSR2</i>	Impaired spliceosome functions and deregulated RNA processing result in aberrant splicing patterns.	14
Cohesin-complex genes	<i>STAG2</i> , <i>RAD21</i>	Mutations may lead to impaired accurate chromosome segregation and impact transcriptional regulation.	13

TF, transcription factor.

*; a dataset from TCGA was selected for younger patients (median age, 55 years); thus, the respective data might be biased, and incidences could vary if additional older patients were included.

Table adapted from Bullinger et al. 2017 with modifications based on Grove and Vassiliou 2014.

1.2. Kinases and phosphatases in AML

Aberrant signaling in AML is dependent on net substrates phosphorylation levels regulated by kinases and phosphatases. This chapter introduces the reader to these enzymes, commonly hijacked by AML cells, and describes several prominent examples of molecular mechanisms employing kinases or phosphatases to support leukemia cell proliferation.

1.2.1. Brief introduction to kinases and phosphatases

Phosphorylation is an addition of a phosphate group to a given substrate by a kinase, whereas dephosphorylation is a removal of a phosphate group from a substrate by a phosphatase. Phosphorylation and dephosphorylation are some of the most prominent substrate modifications due to the chemical properties of the phosphate group, the chemical stability of phosphate esters in aqueous solutions at physiological pH, and the favorable kinetics of these reactions in the presence of ATP (for phosphorylation) and appropriate enzyme catalysts. The unique interactions of the phosphate group with the solvent (ionic shell) and its negative charge allow new physicochemical properties for the biomolecular surfaces gaining or losing the covalently bound phosphate groups (Hunter 2012). The predominant protein residues undergoing phosphorylation/dephosphorylation in eukaryotes are serine (86.4% of proteins), threonine (11.8%), and tyrosine (1.8%) (Olsen et al. 2006), with many of these sites playing critical regulatory roles.

In the field of tyrosine phosphorylation, a preconception was born where tyrosine kinases were thought of as “switch on,” or “writers,” with predominantly active roles in cell signaling, allowing for the “reader” function by proteins recognizing the phosphorylation at the specific target tyrosine residues. Protein tyrosine phosphatases in this “three-part toolkit” were reduced to the role of “erasers,” merely counteracting the function of kinases by “switching off” their substrates (Lim and Pawson 2010). However, both phosphorylation

and dephosphorylation can activate or deactivate proteins, and the whole signaling cascades through the changes in protein conformation and the creation or ablation of binding surfaces, affecting protein functions, protein-protein interactions, localization, and stability (Nishi et al. 2014). Moreover, the activity of some phosphatases can result in an increase in phosphorylation of downstream targets (see Introduction §1.2.2.9). Along these lines, the preconception that protein kinases act as an “on switch” in cancer, whereas phosphatases only serve as tumor suppressors, is now refuted by a growing body of evidence demonstrating that some phosphatases are dominant oncogenes, or “on switches,” in a number of contexts with several remarkable examples in leukemia discussed below (SHP2, PPM1D, and others) (Elson 2018).

The potency of phosphorylation/dephosphorylation in regulating the functions of their substrates made these processes central in cell signaling. Phosphorylation signaling networks are cellular information processing systems. Signal recipients at the cell surface, such as receptor tyrosine kinases and phosphatases, detect an extracellular stimulus and pass it to the downstream network of signal transducers (RAS, Src, SHP2, and many others). The signal gets amplified or dampened via numerous positive and negative feedback loops in the complex signaling networks. Multiple upstream signal transduction pathways can converge on signal integrators and processors (such as mTOR or p53). At the final step of the simplified signal transduction pathway, phosphorylation regulates signal effectors, such as transcription factors, mediating the cell response to the stimulus. At every level of signal transduction, the signaling output defined as net protein phosphorylation is controlled by a dynamic equilibrium between kinases and phosphatases. The cooperation between these classes of enzymes is especially crucial in the space and time-sensitive processes that require highly coordinated activities of all players, such as cell division (Szomolay and Shahrezaei 2012; Nasa and Kettenbach 2018; Gelens et al. 2018).

1.2.1.1. Classification of kinases and phosphatases

By latest estimates, there are ~768 protein-coding genes with known and predicted kinase domains (616 protein kinases, ~152 non-protein and pseudokinases; Table C) and 264 genes with known and predicted phosphatase domains (~128 protein phosphatases, ~136 non-protein and pseudophosphatases; Table D) (Klingbeil, unpublished; Tweedie et al. 2021; Chen et al. 2017). This imbalance in the absolute numbers of protein kinases and phosphatases, as well as the seeming lack of specificity exhibited by some serine/threonine phosphatases *in vitro*, – even though for the tyrosine phosphatases both observations were simply not true (Tonks 2013), – served some of the arguments for the initial dismissal of the phosphatases as “housekeeping enzymes,” maintaining the basal unphosphorylated state rather than regulating signaling. The latter notion was at least in part responsible for the lack of interest in the careful studies of phosphatases and the development of specific inhibitors. However, by the end of the 1990s, powerful models emerged from the field of tyrosine phosphatases, proving that together with the substrate specificity *in vitro*, phosphatase deletion *in vivo* could drive a “fabulously specific phenotype” (Elchebly et al. 1999; Klamann et al. 2000; Cheng et al. 2002). Moreover, further studies of serine/threonine phosphatases, such as Protein Phosphatase 1 (PP1) and 2A (PP2A), uncovered the combinatorial complexity driven by these enzymes forming multimeric complexes (holoenzymes) with non-catalytic subunits that determine their distinct cellular localization, substrate specificity, and precise regulation (Brautigan 2013).

In terms of classification, kinases and phosphatases can be divided on the basis of acting on protein versus non-protein substrates (Table C; Table D). With the availability of the human genome sequence, researchers could easily identify a large number of protein kinases (581) that belong to a single superfamily containing a eukaryotic protein kinase catalytic domain (Manning et al. 2002). However, the discovery of kinases lacking this

sequence similarity (atypical) has been lagging. There are now 43 atypical protein kinases and ~144 genes with predicted kinase activity against non-protein and yet unknown substrates in the Human Kinome project.

In contrast to kinases, even protein phosphatases are a highly diverse group of enzymes comprising many different protein structural folds and catalytic mechanisms (see Table D; curated from Chen et al. 2017). Moreover, phosphatases with the same fold can act on protein and non-protein substrates, and a given function may not be limited to a single fold. Around 40% of all the 264 human protein phosphatases belong to the CC1 (Cys-based class I) fold superfamily. The next largest group (16%) is HAD phosphatases (discussed in more detail below). Phosphatases with the PPM (metal-dependent protein phosphatase) and PPPL (phosphoprotein phosphatase-like) folds are responsible for the majority of phosphoserine/phosphothreonine dephosphorylation and account for ~14% of all phosphatases. The remaining folds are less common.

1.2.1.2. Evolution of kinases and phosphatases

The reason for a striking overtake of the protein kinases by a single protein structural fold is that the majority of them originate from a common ancestor through genetic duplications. The characteristic kinase fold is shared between eukaryotes and bacteria, archaea, and viruses (Jin and Pawson 2012). The expansion and functional differentiation of the kinome went hand in hand with the development of multicellularity and cellular specialization. For example, the number of kinase subfamilies more than doubled in basal metazoa. The increase in the numbers and specialization of RTKs in humans, with some families unique to this clade, has been associated with the advancement of multiple organ systems and processes (Manning et al. 2002).

True to the complexity of their classification, phosphatases are thought to have

distinct ancestry with complex evolution, including both divergent and convergent pathways. Moreover, some motifs and catalytic mechanisms are thought to arise independently several times during evolution, as exemplified by the three classes of dual-specificity and tyrosine phosphatases containing the same catalytic core motif (CX5R). Similar to kinases, big radiation of the phosphatases subfamilies started at the transition from unicellular to multicellular organisms. This included the divergent spread of the CC1 and PPM folds (Chen et al. 2017). The sequences for the characteristic folds of serine/threonine phosphatases (PPM and PPPL) are unrelated, yet, at the structure level, they converged to the high similarity in their catalytic centers (Moorhead et al. 2009). Many phosphatases co-evolved with their regulatory subunits that guide their substrate specificity and subcellular localization. The latter is true both for the complex network of docking interactions observed for the PPP members and the genetic fusions of catalytic domains to their regulatory sequences (PTP, PPM, HAD), with PTP phosphatases serving notable exception due to their encoded substrate specificity (Moorhead et al. 2009).

Table C. Classification of human kinases

#	Human kinome #	By domain structure #	Substrate #	Major groups #	Notable families #	
768	Protein kinases 624	Atypical 43				
		Eukaryotic 581	Ser/Thr 478	AGC 69	PKA 6 PKG 2 PKC 10 other families 51	
				Casein kinase 1 17		
				STE 53		
				Tyrosine kinase-like 49	RAF 8 other families 41	
				CMGC 73	cyclin-dependent kinase (CDK) 24	
					mitogen-activated protein kinase (MAPK) 18	
					glycogen synthase kinase (GSK) 2	
					CDC-like kinase (CLK) 6	
					dual-specificity tyrosine-regulated kinases (DYRK) 11	
				Ca ²⁺ /calmodulin-dependent protein kinase 113	other families 12	
			Other 104	NKF4 2 other families 102		
			Tyr 95	Tyrosine kinases 95	Receptor Tyrosine Kinases 60	Alk 2
						TAM 4
						EGFR 4
						FGFR 4
						Met 2
						PDGFR 5
		VEGFR 4				
		other families 35				
Non-Receptor Tyrosine Kinases Src 21	Abl 2 Src 12 other families 7					
Other Cytoplasmic Tyrosine Kinases 13	JAK 4 other families 9					
TK-Unique 1						
Pseudokinase 8	Receptor Guanylate Cyclases 8					
Non-protein kinases 144	Lipid 20	Phosphatidylinositol kinases				
		Sphingosine kinases				
		other families?				
	Sugar and other substrates 121					
	Riboflavin 1	Riboflavin kinase				
Thymidine 2	Thymidine kinases					

Table D. Classification of human phosphatases (1/2)

Human phosphatases	#	Fold	#	Substrate	Family	#	Members				
Protein phosphatases	93	CC1	52	Tyr	PTP	32	<i>PTPN1, PTPN11, PTPN12, PTPN13, PTPN14, PTPN18, PTPN2, PTPN22, PTPN23, PTPN3, PTPN4, PTPN5, PTPN6, PTPN7, PTPN9, PTPRA, PTPRB, PTPRC, PTPRD, PTPRE, PTPRF, PTPRG, PTPRH, PTPRJ, PTPRK, PTPRM, PTPRO, PTPRR, PTPRS, PTPRT, PTPRU, PTPRZ1</i>				
							Dual specificity	DSP	20	<i>CDC14A, CDC14B, CDKN3, DUSP1, DUSP2, DUSP4, DUSP5, DUSP10, DUSP14, DUSP18, DUSP15, DUSP22, DUSP26, DUSP3, DUSP6, DUSP7, DUSP9, SSH1, SSH2, SSH3</i>	
										SSU72	1
		CC2	1					<i>CDC25A, CDC25AB, CDC25C</i>			
		CC3	3					<i>PGAM5, UBASH3A, UBASH3B</i>			
		HP						<i>PDP1, PDP2, PHLPP1, PHLPP2, PPM1A, PPM1B, PPM1D, PPM1E, PPM1F, PPM1G, PPM1H, PPM1M, PPM1K, PPM1L</i>			
		PPM	28	Ser/Thr	PPM	14	10	<i>PPP1CA, PPP1CB, PPP1CC, PPP2CA, PPP2CB, PPP3CA, PPP3CB, PPP4C, PPP5C, PPP6C</i>			
		PPP						1	<i>RPAP2</i>		
		RTR1							<i>FCP1 (CTDP1), SCP1 (CTDSP1), SCP2 (CTDSP2), SCP3 (CTDSPL)</i>		
		HAD	8	Ser/Thr	FCP	4	3	<i>EYA1, EYA2, EYA3</i>			
								Tyr	Other	1	<i>MDP1</i>
PHP	1					1	<i>PHP (PHPT1)</i>				
CC2	1					1	<i>LMWPTP</i>				
Protein and non-protein substrates	7	CC1	4	Dual specificity	DSP	2	<i>Laforin (EPM2A), PTP4A3</i>				
							Lipid	Myotubularin	1	1	<i>PTEN</i>
		HAD	2					2	<i>NagD</i>		
								2	<i>CIN (PDXP), PGP</i>		
								2	<i>PTPRQ, PTPRN2</i>		
		CC1	19	Lipid	Myotubularin	7	5	7	<i>MTM1, MTMR1, MTMR2, MTMR14, MTMR3, MTMR6, MTMR7</i>		
	Sac	5							<i>FIG4, INPP5F, SACM1L, SYNJ1, SYNJ2</i>		
Non-protein phosphatases	89	HAD	24	5'-nucleotidases	12	9	<i>NT5C, NT5C1A, NT5C1B, NT5C2, NT5C3A, NT5C3B, NT5DC1, NT5DC2, NT5DC3, NT5DC4, NT5E, NT5M</i>				
							Lipin	3	3	<i>ENOPH1, EPHX2, NANP, PHOSPHO1, PHOSPHO2, PMM1, PNKP, PSPH, PUDP</i>	
											<i>LPIN1, LPIN2, LPIN3</i>
		Chloroperoxidase	14	Lipid	Phosphatidic acid phosphatase	9	2	9	<i>LPPR2, LPPR3, LPPR4, PPAP2A (PLPP1), PPAP2B (PLPP3), PPAP2C (PLPP2), PPAPDC1A (PLPP4), PPAPDC1B (PLPP5), PPAPDC2 (PLPP6)</i>		
										<i>T2LPP</i>	2
		HP	9	Sugar	Glc-6-Pase	3	8	1	<i>G6PC (G6PC1), G6PC2, G6PC3</i>		
									HP1	8	<i>PFKFB1, PFKFB2, PFKFB4, BPGM, PGAM1, PGAM2, PGAM4, TIGAR</i>
									HP2	1	<i>MINPP1</i>
									IPP5	8	<i>INPP5A, INPP5B, INPP5D, INPP5E, INPP5J, INPP5K, INPPL1, OCRL</i>
									IMFB	7	<i>FBP1, FBP2, BPNT1, IMPA1, IMPA2, IMPAD1, INPP1</i>
									INPP4	2	<i>INPP4A, INPP4B</i>
		Nudix	5				5	<i>NUDT10, NUDT11, NUDT14, NUDT3, NUDT4</i>			
		ABD	1					1	<i>ITPA</i>		

Table D. Classification of human phosphatases (2/2)

Human phosphatases	#	Fold	#	Substrate	Family	#	Members	
Undetermined	48	CC1	20	Tyr?	PTP	3	<i>PTPN20, PTPN21, PTPRN</i>	
					Paladin	1	<i>PALD1</i>	
				Dual specificity?	DSP	12	<i>DUPD1, DUSP12, DUSP13, DUSP16, DUSP19, DUSP21, DUSP23, DUSP28, DUSP8, PTP4A1, PTP4A2, PTPDC1</i>	
					PTEN	2	<i>Tensin, VSP</i>	
		Lipid?	Myotubularin	2	<i>MTMR8, MTMR9</i>			
		PPM	10	Ser/Thr?		4	<i>ILKAP, PPM1J, PPM1N, PPTC7</i>	
		PPPL				4	<i>PPEF1, PPEF2, PPP3CC, VPS29</i>	
		OB				2	<i>PPA1, PPA2</i>	
		HAD	10				4	<i>CECR5, CTDSPL2, Dullard, EYA4, HDHD2, HDHD3, LHPP, PMM2, TIMM50, UBLCP1</i>
		Chloroperoxidase	4	Lipid?		4	<i>DOLPP1, LPPR1, LPPR5, PPAPDC3 (PLPP7)</i>	
TMEM55	2			2	<i>TMEM55A (PIP4P2), TMEM55B (PIP4P1)</i>			
HD	2				<i>HDDC2, SAMHD1</i>			
Alkaline phosphatases	4					<i>ALPG, ALPI, ALPL, ALPP</i>		
Acid phosphatases	7					<i>ACP1, ACP2, ACP3, ACP4, ACP5, ACP6, PXYLP1</i>		
Verified or predicted pseudo-phosphatases	16	CC1	6	DSP	3	<i>DUSP27 (DUSP29), STYX, STYXL1,</i>		
				PTEN	3	<i>DNAJC6, GAK, TNS1</i>		
		Various	10	Other	2	<i>PP2D1, TAB1</i>		
				Other	3	<i>PFKFB3, PPIP5K1, PPIP5K2</i>		
		Other	5	<i>MTMR10, MTMR11, MTMR12, SBF1 (MTMR5), SBF2 (MTMR13)</i>				
TOTAL	264		264			264		

1.2.2. Notable kinases and phosphatases in AML

1.2.2.1. FLT3

Activating mutations in the FMS-related tyrosine kinase 3 (*FLT3*) are present in one-third of adult AML patients (Dohner et al. 2010). Most health organizations, including the World Health Organization, NCCN, and European LeukemiaNet, consider mutations in *FLT3* poorly affecting the clinical outcomes (Table A). Commonly observed *FLT3* mutations include *FLT3* internal tandem duplications (*FLT3*-ITD) and tyrosine kinase domain (TKD) point mutations (Grafone et al. 2012).

FLT3 is a cell membrane protein that belongs to the class III receptor tyrosine kinases, characterized by an extracellular ligand-binding domain (ED) with five immunoglobulin-like modules, a transmembrane domain, and a cytoplasmic region consisting of a juxtamembrane domain (JMD) and two highly conserved TKDs (Matsumura et al. 2008). In the inactive state, *FLT3* is auto-inhibited by the JMD. Upon cytokine binding to the ED, *FLT3* receptors oligomerize and undergo activating transphosphorylation on their JMD and TKD domains. Activated *FLT3* phosphorylates downstream signaling effectors in the RAS/RAF, JAK/STAT, and PI3K/AKT pathways, regulating cellular proliferation, inhibition of apoptosis, and differentiation. ITD mutations are in-frame mutations in the JMD region that result in the multiplication of its coding sequence and the subsequent loss of the JMD auto-inhibitory activity. Point mutations in the TKD stabilize the active kinase conformation and also result in the aberrant activation of downstream targets and the cytokine-independent growth of hematopoietic cells, though their role in leukemogenesis and prognostic values are less defined (Schnittger et al. 2012).

FLT3 is specifically expressed in early CD34⁺ hematopoietic stem/progenitor cells (HSPCs) and absent from more mature blood cell populations (Grafone et al. 2012). Early knockout mouse models showed that *Flt3*^{-/-} mice developed normally with the only

recorded deficiencies in early B cell progenitors and reduced repopulation capacity upon bone marrow transplantation (Mackarehtschian et al. 1995). However, it is important to note that murine *Flt3* expression is restricted to common lymphoid progenitors and leads to the loss of self-renewal ability (Mackarehtschian et al. 1995). Flt3 ligand (FL) also fails to promote the survival of mouse hematopoietic stem cells and is dispensable for myeloid development (Sitnicka et al. 2002). In contrast, human CD34⁺FLT3⁺ cells can efficiently and continuously reconstitute both lymphoid and myeloid lineages, may have erythroid differentiation potential, and FL promotes their viability, underlining the differences in mouse and human roles for FLT3 (Kikushige et al. 2008). Additionally, PI3K activation by human FLT3 is different to mice since human FLT3 does not directly bind the p85 α subunit of PI3K (Zhang and Broxmeyer 1999).

FLT3-ITD is a potent driver of leukemogenesis in humans, as shown by its frequent presence at baseline and relapse (Kronke et al. 2013) and the emergence of on-target *FLT3* mutations upon the use of potent inhibitors (Smith et al. 2012). *Flt3*-ITD knock-in mice develop myeloproliferative and lymphoid disorders and overt AML in synergy with other common AML oncogenes (reviewed in Muller and Schmidt-Arras 2020).

FLT3^{wt} in human hematopoietic stem cells was reported to support cell survival through multiple mechanisms, including upregulation of MCL-1 – an anti-apoptotic protein (Kikushige et al. 2008). This and other functions of FLT3^{wt} could be exploited by its constitutive activation in cancer cells (Yoshimoto et al. 2009). Among FLT3-ITD mutant specific effectors are transcription factor STAT5 (Choudhary et al. 2007) and protein tyrosine phosphatase SHP2 (Chen et al. 2016). Apart from aberrant proliferation, FLT3-ITD was shown to induce differentiation block through inhibiting several myeloid transcription factors, such as PU.1 and RUNX1 (Takahashi 2011).

1.2.2.2. c-KIT

c-KIT is another class III receptor tyrosine kinase essential for the early hematopoietic stem cells self-renewal and differentiation and having a recognized role in leukemogenesis. It has a similar overall structure to FLT3 (see above). While mutations in the *c-KIT* gene in AML happen in ~5% of cases (Badr et al. 2018), it is often aberrantly expressed in AML blasts. In a subset of AML characterized by t(8:21) and inv(16)/t(16;16) translocations (core-binding factor, or CBF-AML) frequency of c-KIT mutations reaches 30% and is associated with worse prognoses in this otherwise better risk group (Table A) (Ayatollahi et al. 2017). However, c-KIT mutations prognostic values vary between reports depending on the type of the translocation and the specific *c-KIT* variant, age of patients, and mutation allele burden (Tarlock et al. 2019).

c-KIT^{wt} activates similar downstream pathways to FLT3: PI3K/AKT (through p85 α PI3K subunit), RAS/RAF, JAK/STAT, and other. c-KIT mutations are gain-of-function and lead to ligand-independent constitutive kinase activity.

c-KIT mutations promote cell growth and radiation resistance through aberrant signaling (Malaise et al. 2009). Among mutant-specific effectors of c-KIT were reported the increased PI3K and STAT3 activation (Chian et al. 2001; Ning et al. 2001), possibly due to c-KIT^{mut} mislocalization to the intracellular membranes (Schmidt-Arras and Bohmer 2020). Other interesting mechanisms reported for c-KIT^{D816V} were circumventing the requirement for phosphorylation by the Src family of protein tyrosine kinases (SFKs) for downstream signaling (Sun et al. 2009) and degradation of its negative regulator src homology 2 domain (SH2)-containing protein tyrosine phosphatase SHP1 (Piao et al. 1996).

Although there are no selective c-KIT inhibitors available yet, the addition of multikinase inhibitors, such as midostaurin or dasatinib, to chemotherapy has shown clinical benefits for the c-KIT^{mut} AML patients (Marcucci et al. 2020).

1.2.2.3. BCR-ABL1

The translocation between chromosomes 9 and 22 [t(9;22)(q34.1;q11.2)] was the first identified cytogenetic abnormality in leukemia. It was discovered in 1960 in two chronic myelogenous leukemia (CML) patient samples and named after the place of its discovery, the Philadelphia chromosome (Nowell 1962). Later studies clarified that the translocation involved the protein tyrosine kinase *ABL1* (Abelson Murine Leukemia Viral Oncogene Homolog 1) and the gene breakpoint on the breakpoint cluster region (*BCR*) of chromosome 22. The fusion transcript renders a constitutively activated non-receptor tyrosine kinase (de Klein et al. 1982; Shtivelman et al. 1985). Like the tyrosine kinases above, it controls many cellular processes associated with cell proliferation, differentiation, and stress response.

Constitutive activation of BCR-ABL1 is achieved through several mechanisms, including loss of auto-inhibitory myristoylation at the N-terminus of ABL1 (Hantschel 2012). It is a powerful driver of leukemogenesis shown to cause myeloproliferative disorders in syngeneic mouse models. One of the critical effectors of BCR-ABL1 is the adapter GAB2, required for its downstream signaling to PI3K and SHP2-mediated signaling pathways (Gu et al. 2016).

Various *BCR-ABL1* isoforms are present in most CML patients (0.5% of all yearly cancer cases), ~25% of adult B-cell ALL patients, and rarely in AML, lymphoma, and myeloma (Zhou et al. 2018). Imatinib mesylate (Gleevec) was one of the first rationally designed selective anticancer drugs and received the FDA approval for CML patients in 2001, just two and a half years after its development (Iqbal and Iqbal 2014), and many more effective tyrosine kinase inhibitors (TKIs) followed.

The introduction of imatinib and second- and third-generation TKIs dramatically improved survival outcomes in CML patients, with their life expectancy in the chronic phase approaching 98% of that in the general population (Bower et al. 2016). Most patients with

CML take TKIs for life, even though there have been cases of successful discontinuation (Hughes and Ross 2016). TKIs immediate and long-term side effects are some of the standing challenges in BCR-ABL1-positive patients, adversely affecting their quality of life and treatment adherence. There are also multiple mechanisms of drug resistance and the lack of options for the patients with advanced disease in accelerated and blast phases (Shah 2008).

One potential explanation for the high likelihood of relapse upon TKI withdrawal is the presence of the minimal residual disease presented by leukemia stem cells (LSCs) that are intrinsically resistant to BCR-ABL1 inactivation (Morotti et al. 2014). Understanding the biology of LSCs is essential to eradicate this population and prevent future disease reoccurrences. Another explanation for the high relapse rate – not mutually exclusive with the first one – is that the BCR-ABL1 protein itself has a structural role in supporting leukemia cell and LSC proliferation regardless of its kinase activity (whereas TKIs selectively block only its kinase function) (Neviani et al. 2013).

1.2.2.4. JAK2

Approximately 90% of patients with BCR-ABL1-negative MPNs bear mutations in the Janus-Associated Kinase/Signal Transducers and Activators of Transcription (JAK/STAT) signaling pathway (Venugopal and Mascarenhas 2020). Between 5 and 10% of these patients develop AML with poor prognoses (Kennedy et al. 2013).

JAKs have two kinase domains: catalytically active C-terminal tyrosine kinase domain and N-terminal pseudokinase domain. Almost 70% of patients with MPNs harbor V617F mutation in the JAK2 pseudokinase domain, which results in its cytokine-independent activation, potentially through disrupting its inhibitory cellular interactions (Liau et al. 2019).

JAKs regulate survival, proliferation, and differentiation of hematopoietic cells in response to cytokines, serving as essential catalytic counterparts of the cytokine receptors. Upon activation, JAK2 initiates downstream signaling through STATs nuclear localization, RAS/RAF, and PI3K/AKT signaling pathways.

STAT5 is one of the main effectors of JAK2^{V617F}, and *Stat5* deletion impaired *Jak2*^{V617F} driven pathogenesis in mice (Walz et al. 2012). Besides, JAK2^{V617F} catalytic function has been implicated in epigenetic regulation. Dawson et al. showed that JAK2^{V617F} is present in the nucleus and phosphorylates histone H3Y41, protecting the associated chromatin from epigenetic silencing (Dawson et al. 2009). Jeong et al. reported that JAK2^{V617F} activates Ten-Eleven-Translocase 2 (*TET2*) – an enzyme that removes cytosine methyl marks from DNA – leading to upregulation of HOXA9 and MEIS1 (Jeong et al. 2019).

Data from our group points to potential cooperativity between JAK2^{V617F} and a dual-specificity tyrosine-regulated kinase 1A (DYRK1A) (Tarumoto et al. 2018). This phenomenon could be mediated by the converging signaling of JAK2 and DYRK1A on STAT3, constitutively activated in hematologic malignancies. Maximum transcriptional activity of STAT3 is achieved through JAK2-mediated phosphorylation of its Y705 residue and DYRK1A mediated phosphorylation of S727 residue (Wen et al. 1995; Matsuo et al. 2001). However, inhibition of DYRK1A has not yet been carefully explored in JAK2^{V617F} AML, to our knowledge.

1.2.2.5. CDK6

Characteristic AML translocations involving chromosome 11q23, the locus for *MLL1*, are associated with poor prognosis (Table A) and account for ~10% of adult leukemias and ~60% of infant leukemias (Grigsby et al. 2021). *MLL1* encodes histone 3 lysine N-methyltransferase 2A (KMT2A). KMT2A is an epigenetic regulator controlling the

transcription of many genes important in normal development and hematopoiesis (Hsieh et al. 2003). *MLL1* locus contains a breakpoint region after the DNA-binding portion of *KMT2A*, which is involved in recurrent chromosomal translocations generating *KMT2A*-fusion proteins with over 80 different partners (Meyer et al. 2009). A characteristic transcriptional signature in *MLL*-fusion AML is the upregulation of *HOX* genes (Armstrong et al. 2002). *MLL*-fusions exploit various epigenetic regulators for the aberrant activation of their targets, such as Disrupter of Telomeric silencing 1-Like (DOT1L) (Okada et al. 2005), alone or in a complex with the nuclear scaffold protein menin (*MEN1*) (Jin et al. 2010). Both of these *MLL*-fusion partners have been explored as therapeutic avenues for *MLL*-fusion-positive leukemias in pre-clinical and clinical studies (Grieselhuber and Mims 2021).

Functional genomics studies from Fröhling group (RNAi) and later ours (CRISPR-Cas9) provided evidence for the cell division protein kinase 6 (CDK6) being specifically essential for the survival of *MLL*-fusion-positive human leukemia cell lines (Placke et al. 2014; Tarumoto et al. 2018). CDK6 role in these cells was independent of its control of cell-cycle G1/S progression but instead supported the myeloid differentiation arrest through its catalytic activity. CDK6 was shown to be a direct transcriptional target of *MLL*-AF9, and CDK6 inhibition impaired colony formation and promoted differentiation of *MLL*-fusion-positive cell lines in a dose-dependent manner, and conferred significant survival benefits in mouse models.

Later reports suggested a more broad role for CDK6 in leukemogenesis, including supporting transformation by *JAK2*^{V617F} and *FLT3*-ITD. *CDK6* gene is frequently upregulated in human leukemias. Among CDK6 potential downstream effectors were named transcription factors STATs and activator protein 1 (AP-1), and negative regulation of p53. CDK6 inhibitors clinical trials in AML are currently ongoing (Uras et al. 2020).

1.2.2.6. SIK2/3 and MARKs

Another transcription factor exploited by MLL-fusions is the myocyte enhancer factor 2C (MEF2C), critical for AML proliferation (Krivtsov et al. 2006; Schwieger et al. 2009) and dispensable for the establishment or maintenance of the myeloid lineage *in vivo* albeit slight monocyte deficiency (Schuler et al. 2008). One way to target this transcription factor through a phospho-signaling pathway emerged in our laboratory from an unbiased exploration of the specific AML-promoting roles of all human kinases (Tarumoto et al. 2018).

Human leukemia cell lines with *MLL*-fusion showed increased sensitivity to LKB1 knockout and SIK3/SIK2 double knockout. In a series of genetic and biochemical experiments, Tarumoto et al. showed that this effect, at least in part, was explained by LKB1 activating phosphorylation on SIK3, which in turn phosphorylated HDAC4. HDAC4 phosphorylation by SIK3 results in its nuclear exclusion (Walkinshaw et al. 2013), therefore inhibiting its function. Thus, LKB1-SIK2/3 signaling was essential in the AML context to prevent HDAC4 from counteracting MEF2C functions on chromatin. SIK3 early stage development inhibitors are available and showed promise *in vivo* in two different mouse models of MLL-fusion driven leukemias while somewhat lacking in efficacy as compared with the direct gene inactivation (Tarumoto et al. 2020).

MEF2C can also be directly phosphorylated and activated by MARK kinases, which results in increased AML pathogenesis and worse prognoses, at least in part due to the MEF2C aberrant phosphorylation being specifically associated with LSCs maintenance (Brown et al. 2018). In a series of elegant experiments, Brown et al. showed that mice with homozygous S222A MEF2C mutation were resistant to MLL-AF9-induced leukemogenesis due to significantly reduced MEF2C transcriptional activity. A dual selective MARK/SIK inhibitor MRT199665 effectively induced apoptosis in AML cells and could potentially exhibit synergistic effects in patients with aberrant MEF2C activation.

1.2.2.7. SHP2

SH2 domain-containing phosphatase 2 (SHP2) is a non-receptor protein tyrosine phosphatase encoded by the *PTPN11* gene. It is a major signaling node transmitting the information from upstream tyrosine kinases to the central cellular signaling highways, such as RAS/RAF, JAK/STAT, and PI3K/AKT pathways (Frankson et al. 2017). SHP2 contains two substrate-binding SH2 domains at the N-terminus, a protein tyrosine phosphatase (PTP) domain, and a C-terminal domain with two tyrosine phosphorylation sites. Upon substrate binding, SHP2 undergoes a conformational change that results in its activation through enabling access to the PTP domain catalytic site (Barford and Neel 1998). This mechanism became the basis for the later discovery of potent and specific SHP2 allosteric inhibitors locking it in its inactive conformation (Chen et al. 2016).

SHP2 is necessary for normal hematopoiesis (Qu et al. 1998). Activating mutations in *PTPN11* were initially described for the RASopathies, rare congenital conditions leading to multiple developmental abnormalities and an increased risk of leukemias (Rauen 2013). Approximately 35% of patients with juvenile myelomonocytic leukemia carry somatic mutations in *PTPN11* associated with worse outcomes and increased risk for relapse (Niemeyer and Flotho 2019). Large cohorts genomic studies reported *PTPN11* mutations in 4–8% of adult AML patients, also associated with worse prognoses (Alfayez et al. 2021). The common SHP2 mutations include D61V and E76K at the interface between the N-terminal SH2 domain and the PTP domain, which weaken SHP2 auto-inhibitory conformation (Padua et al. 2018).

One of the major effectors of SHP2 is RAS, another well-established oncogene in AML (Chan et al. 2004). SHP2 was reported to interact with directly and dephosphorylate RAS, promoting the RAS GTPase cycle (Bunda et al. 2015). Besides, SHP2 downstream targets (Sprouty, GRB2/SOS, and RAS-GAP) also regulate RAS activity. Many pre-clinical

and clinical studies evaluated SHP2 inhibitors alone or in combination with MEK inhibitors, and several clinical trials are still ongoing.

Mutations in *RAS* and *PTPN11* were shown to be largely mutually exclusive (Paulsson et al. 2008), suggesting their combination disadvantageous to the cell or simply redundant for providing any additional selective benefit. Notably, several research groups reported that tumors with *RAS* mutations were resistant to SHP2 inhibition, consistent with the hypothesis that these mutations confer biochemical properties to *RAS*, enabling its function in the absence of upstream RTK/SHP2 signaling (Chen et al. 2016; Ahmed et al. 2019; Valencia-Sama et al. 2020). This was challenged in a 2019 study providing evidence for *KRAS*^{mut} cells being sensitive to the SHP2 inhibitor SHP099 in 3D multicellular spheroids but not in a 2D monolayer (Hao et al. 2019). Of note, most leukemia cell lines grow in suspension. Further dissection of SHP2 dependent and independent *RAS* signaling in different cancer contexts is necessary for the clinical development of SHP2 inhibitors as a reasonable treatment strategy. One potential resistance prevention approach could be to combine SHP2 inhibitors with *RAS*^{mut} specific inhibitors (Ryan et al. 2020).

SHP2 inhibitor SHP099 screening showed a correlation between human cell lines' sensitivity to SHP2 inactivation and the dependence on RTK signaling, such as activating *FLT3* mutations and *BCR-ABL1*. Wild type *FLT3* engages SHP2 through *GRB2* and *GAB1/2* (Zhang and Broxmeyer 2000). Knockdown of *Shp2* was shown to significantly reduce the proliferation of the murine pro-B cell line Ba/F3 constitutively expressing human *FLT3-ITD* whereas having little effect on the cells expressing wild-type *FLT3* (Nabinger et al. 2013). SHP099 eliminated tumor burden in the primary-tumor-derived *FLT3-ITD* AML xenograft setting (Chen et al. 2016). It was also shown to be one of the critical downstream effectors of *BCR-ABL1*, allowing its leukemogenic functions. A 2018 study from Gu et al. reported that SHP099 suppressed *BCR-ABL1*-positive pre-B cell proliferation (Gu et al. 2018).

1.2.2.8. PP2A

Protein phosphatase 2A (PP2A) is a serine/threonine phosphatase with numerous functions in the cell possible by forming up to 60 distinct holoenzymes with diverse substrate specificities and subcellular localizations (Fowle et al. 2019). PP2A activity is tightly regulated at every level, including post-transcriptional via interactions with various endogenous inhibitors of its functions (such as SET and cancerous inhibitor of PP2A, CIP2A). Endogenous PP2A inhibitors were shown to be overexpressed in many leukemias and other cancer contexts, whereas the mutations in PP2A itself are relatively rare (Kauko and Westermarck 2018). This allowed the development of P2A-activating drugs (PADs) – small molecules disrupting the interaction between PP2A and its endogenous inhibitors (Vainonen et al. 2021).

In 2016 Smith et al. showed that *FLT3*-ITD expression led to *PP2A* transcriptional downregulation in cell lines and patient samples (Smith et al. 2016). The use of PADs FTY720 and AAL(S) resulted in proliferation arrest and cell death specific for *FLT3*-ITD expression.

The same group earlier discovered a similar relationship between PP2A and c-KIT. Mutant *c-KIT* myeloid cells were shown to downregulate PP2A, and overexpression of PP2A inhibited the proliferation of *c-KIT* mutant cells (Roberts et al. 2010). FTY720 exhibited preferential cytotoxicity to *c-KIT* mutant leukemia models and was suggested as a therapeutic strategy for c-KIT-positive cancers.

Similar to the other tyrosine kinases, BCR-ABL1 was shown to rely on PP2A inhibition to drive leukemia. Importantly, BCR-ABL1 induced SET expression in the dose- and kinase activity-dependent manner upon CML progression from the chronic phase to the blastic phase (Neviani et al. 2005). The protein levels of CIP2A were also significantly higher at the time of diagnosis in CML patients who later progressed to blast crisis (Lucas et al.

2011). These studies suggested that that PP2A inactivation is essential for CML blastic transformation – one of the main therapeutic challenges in the BCR-ABL1-positive population. FTY720 showed great promise in impairing the proliferation of BCR-ABL1-positive myeloid and lymphoid cell lines and progenitors (Neviani et al. 2007). Overall, PP2A inactivation seems to be a common tumorigenesis mechanism in cancers driven by oncogenic tyrosine kinases.

PP2A was also proposed as a critical regulator of LSCs survival. The 2013 study from Perrotti's group reported that while BCR-ABL1 activity in TKI-resistant CML quiescent LSCs was low, its expression was essential for the JAK2/ β -catenin mediated inhibition of PP2A. In the context of CML LSCs, PP2A activation was shown to decrease survival and self-renewal of quiescent CML LSCs in serial transplantation assays (Neviani et al. 2013).

One of the best characterized oncogenic targets of PP2A is MYC. High MYC expression in AML was associated with poor prognosis and the increased risk of relapse (Ohanian et al. 2019; Yun et al. 2019). MYC acts as a signaling crosspoint of the cell with multiple growth-regulatory pathways converging with its upregulation, including MLL-fusion-driven oncogenesis, RAS, Hedgehog, WNT, NOTCH, and JAK/STAT pathways – all reported to be deregulated in AML (Kress et al. 2015). c-MYC is activated by phosphorylation at serine 62 (S62) that stabilizes its protein levels (Pulverer et al. 1994), and subsequent phosphorylation on threonine 58 (T58) is reported to target it to degradation (Gregory et al. 2003; Yeh et al. 2004). Leukemia and lymphoma cell lines are reported to stabilize MYC protein levels via mutations at or near the T58 site, impaired activity of glycogen synthase kinase3 β (GSK3 β) (responsible for T58 phosphorylation) (Malempati et al. 2006), mutations in the MYC ubiquitin ligase, FBW7 (O'Neil and Look 2007), and downregulation of PP2A.

PADs effectively reduced MYC activity in other types of cancers (Farrell et al. 2014;

Farrington et al. 2020). Controversially, in human T-cell ALL cell lines, PAD iHAP1 induced apoptosis through dephosphorylation of the transcription factor MYBL2, but not c-MYC, suggesting that different PP2A substrates are kept in check by diverse mechanisms (Morita et al. 2020). Of note, other phosphatases have also been reported to support MYC stability by dephosphorylating its T58 residue, such as PHLPP2 in the context of a mouse model of metastatic prostate cancer (Nowak et al. 2019), Eya3 (Zhang et al. 2018), or SCP1 (Wang et al. 2016a), and their role in MYC regulation could be of interest in leukemias.

1.2.2.9. CD45

CD45 is also known as a protein tyrosine phosphatase receptor type C (PTPRC). It comprises a glycosylated extracellular segment, a transmembrane domain, and a conserved intracellular fragment with two tandemly repeated domains (Thomas et al. 1985). The sequence similarity of the intracellular domains of CD45 to PTP1B led Tonks et al. to demonstrate that this membrane protein indeed exhibited a protein tyrosine phosphatase activity (Tonks et al. 1988a).

Historically, CD45 has been widely used in flow-cytometry as a marker of all nucleated hematopoietic cells. In the absence of CD45 or its phosphatase catalytic activity, T lymphocytes fail to become activated through the T-cell receptors, leading to severe combined immunodeficiency (Pingel and Thomas 1989; Koretzky et al. 1990; Kung et al. 2000). Thomas et al. and others showed that CD45 dephosphorylated the auto-inhibitory tyrosines of SFKs, resulting in their conformational change that activated downstream signaling (Ostergaard et al. 1989; Thomas and Brown 1999). The expression of a constitutively active SFK, Lck^{Y505F}, in CD45-deficient mice rescued CD45-loss associated block in T-cell development (Seavitt et al. 1999).

The discovery of the phosphatase activity of CD45 and its biological role in

promoting T lymphocyte proliferation served as an important conceptual advance in the field of phosphatases as a demonstration of their ability to *actively* regulate cell signaling rather than merely antagonize RTKs functions (Tonks 2013). For one, the example of CD45 invalidates the proposed designation of phosphatases as “erasers” since its activity leads to an increase in phosphorylation of a number of downstream targets.

CD45 knockdown impaired AML outgrowth upon transplantation (Saint-Paul et al. 2016). On the other hand, CD45 was also shown to be a negative regulator of JAKs (Irie-Sasaki et al. 2001), thus AML cells have to strike a balance between CD45 cellular functions that are likely context-specific.

While looking for novel compounds with preferential toxicity towards HOXA9-MEIS1-transformed cells, Saint-Paul et al. showed that Pyrido[4,3-b]quinoxaline (PyQ) exhibited anti-leukemic properties both *in vitro* and *in vivo* (Saint-Paul et al. 2016). They reported that CD45 in AML cells was associated with lipid rafts – cholesterol- and glycosphingolipid-enriched patches located in the plasma membrane, promoting Src signaling. PyQ treatment resulted in CD45 dissociation from the lipid rafts and inhibitory phosphorylation on Src kinase Lyn.

Of note, CD45 CRISPR-mediated knockout in the standardized culture conditions did not cause proliferation defects in any human leukemia cell lines, according to DepMap (Meyers et al. 2017; Dempster et al. 2019). Since CD45 is also reported to suppress JAKs, more studies are needed to delineate the roles of CD45 in different contexts.

1.2.2.10. PTP1B

PTP1B was the first isolated protein-tyrosine phosphatase (Tonks et al. 1988b). Early on, its activity has been associated with type II diabetes and obesity, and it has numerous other functions in the cell (Tonks 2003).

Tonks' group was the first to notice in 1998 that PTP1B was specifically enhanced at the transcriptional level in the human myeloid cells expressing BCR-ABL1 and in CML patient samples (LaMontagne et al. 1998). They showed that this overexpression was mediated through BCR-ABL1 kinase activity. Using PTP1B substrate-trapping mutant – an elegant phosphatase substrates discovery method also invented in Tonks' laboratory – and measuring BCR-ABL1 phospho-tyrosine (pTyr) levels upon PTP1B overexpression in the cells, the study conclusively showed that PTP1B dephosphorylated BCR-ABL1. Later, the group moved on to structural studies of PTP1B that allowed to identify trodusquemine (MSI-1436) – a potent allosteric inhibitor of PTP1B (Krishnan et al. 2014).

In a 2011 study, Cotter's laboratory using PTP1B knockdown and previous generation PTP1B inhibitors showed that inhibition of PTP1B phosphatase activity resulted in BCR-ABL1 degradation in human CML cell lines through the polyubiquitination-mediated autophagy-lysosome pathway (Alvira et al. 2011). A 2016 study by Elgehama et al. corroborated these findings by using a small molecule that blocked the interaction between BCR-ABL1 and PTP1B and exhibited preferential cytotoxicity toward human BCR-ABL1-positive cell lines (Elgehama et al. 2016). Neviani et al. reported the importance of the structural role of BCR-ABL1 oncogene in leukemogenesis, particularly in maintaining LSC fate (Neviani et al. 2013). In this regard, PTP1B might be an attractive therapeutic target in BCR-ABL1-positive malignancies.

1.2.2.11. PTPRJ

The receptor-type protein tyrosine phosphatase PTPRJ (alias DEP-1) was first discovered in adherent cell cultures as associated with the high cell density (Ostman et al. 1994) and later named one of the critical phosphatases in hematopoietic cells regulating SFKs, similarly to CD45 (Hermiston et al. 2009). It is often mutated in different cancers

suggesting its role as a tumor suppressor (Julien et al. 2011).

PTPRJ takes off the inhibitory tyrosine phosphorylation on SFKs, enhancing their activity. On the other hand, PTPRJ dephosphorylates and negatively regulated downstream signaling by receptor tyrosine kinases, including FLT3 (Arora et al. 2011; Godfrey et al. 2012; Bohmer et al. 2013; Kresinsky et al. 2018). PTPRJ inactivation in AML cells was shown to happen via FLT3-ITD-induced reactive oxygen species (ROS) levels. In line with that, inhibition of ROS formation restored PtpRJ activity and impaired colony formation by murine lymphoblasts stably expressing FLT3-ITD (Godfrey et al. 2012). Of note, Tonks' group was the first to definitively show that oxidation and inactivation of PTPs by ROS was an important regulatory mechanism *in vivo* (Meng et al. 2002; Tonks 2005; Boivin et al. 2008).

1.2.2.12. PPM1D

TP53 is one of the most commonly mutated tumor suppressors in human cancers. It is a critical nod in cellular signaling networks accumulating information from multiple sources and processing go/no-go decisions concerning every aspect of cell viability and cell division. *TP53* mutations in AML are considered poor prognosis (Table A) (Herold et al. 2020). Only ~8% of *de novo* AML cases harbor *TP53* mutations, compared to 30% of t-AML patients, and up to 70% of AML cases in elderly patients with complex karyotype AML (Molica et al. 2020).

Zuber et al. noted the apparent absence of cooperativity between MLL-fusions and p53 in promoting leukemia, whereas p53-loss accelerated leukemias induced by most other common AML oncogenes (Zuber et al. 2009). The authors hypothesized that some leukemia cells develop mechanisms to evade the p53 network during leukemogenesis. For example, AML cells upregulate murine double minute 2 (MDM2), a negative regulator of p53 (Khurana and Shafer 2019). Several potent and selective small-molecule MDM2 inhibitors

are currently explored in clinical studies with mixed reports on their benefits (Grieselhuber and Mims 2021).

Protein phosphatase Mg^{2+}/Mn^{2+} Dependent 1D (PPM1D) is an important negative regulator of p53 (Lu et al. 2008). It is also called wild-type p53-induced phosphatase 1, WIP1, since p53 normally induces its expression as part of a regulatory feedback loop negatively regulating apoptosis. PPM1D was found to be overexpressed in several cancers at baseline or in response to chemotherapy (Hsu et al. 2018). Hsu et al. showed that PPM1D mutations conferred resistance to apoptosis to cells, specifically under the selective pressure of cytotoxic therapies. Recently, PPM1D gain-of-function mutations were reported to be common in patients with t-MDS and t-AML (Lindsley et al. 2017). This makes PPM1D a potential therapeutic target in the combination therapies against *TP53* wild-type leukemias.

1.3. Current and emerging agents and regimens for AML

This chapter describes currently approved therapies for AML and underlines the importance of identifying new targets and regimens, specifically for the patients ineligible for the intensive therapy and patients presented with relapsed AML or AML refractory to all other treatments (R/R).

1.3.1. Intensive therapy

The standard first-line therapy for AML remained largely unchanged since 1973, with few modifications made in the past two decades, discussed below. It consists of induction therapy and consolidation therapy (Carter et al. 2020). The goal of induction therapy is complete remission defined by the presence of less than 5% of leukemic blasts in the post-induction bone marrow differential. It is chemotherapy with a common “7+3” regimen of 7 days of cytosine-arabioside (cytarabine, or ara-C) infusions and co-administration of an anthracyclin (daunorubicin or idarubicin) in the first 3 days. Cytarabine mimics deoxycytosine and gets incorporated into human DNA in the S phase of mitosis, preventing DNA synthesis in rapidly dividing cells. Anthracyclines are DNA intercalating agents inhibiting topoisomerase II functions, which results in DNA damage and apoptosis induction. Besides severe bone marrow suppression and corresponding complications (anemia, need for blood transfusions, increased risk of infections, and sepsis), each drug has its own toxicity profile. For example, anthracyclines exhibit cardiotoxicity, limiting the most intensive induction therapy to people under 60 or otherwise in good health. The recent improvement in the induction therapy administration is CPX-351 – a liposomal formulation of cytarabine and daunorubicin with enhanced pharmacodynamics and pharmacokinetics, demonstrating better outcomes in high-risk patients 60–75 years old (Lancet et al. 2018). Following successful clinical trials, FDA approved CPX-351 for newly diagnosed AML with poor prognosis and t-AML (Alfayez et al. 2020), and the improved

overall survival was confirmed in a recent 5 years follow-up study (Lancet et al. 2020).

Even if a patient achieves a complete remission upon the conventional induction therapy, they will likely relapse in the following year without consolidation therapy. Again, more options are available for younger patients, such as high-dose ara-C (HiDAC) and stem cell transplant. Older patients ineligible for these treatments have benefited the most from the recent approvals of novel targeted therapies for AML discussed below.

Conventional chemotherapies cannot completely eliminate leukemic clones as they lack to address specific genetics and epigenetics alterations in the AML cells and distinct cellular states, making them more resistant to cytotoxic agents (such as LSCs). Most recently approved AML therapies target molecular pathways uniquely required for AML cell proliferation and survival (Fig. C).

1.3.2. Gemtuzumab ozogamicin (GO)

Gemtuzumab ozogamicin was the first monoclonal antibody (mAb) therapy developed against AML. It is an antibody-drug conjugate directed at the CD33 antigen highly expressed on the surface of AML cells. The CD33 mAb is covalently linked with a cytotoxic drug N-acetyl gamma calicheamicin. GO is usually considered in CD33^{high} and CBF-AML patients at both induction and consolidation phases. Despite low levels of CD33 in CBF-AML patients (Appelbaum and Bernstein 2017), the addition of GO increased their estimated long-term survival rate from 50% to 75%, making t(8:21) and inv(16)/t(16;16) CBF-characteristic translocations better risk in the absence of most other abnormalities (Table A) (Kantarjian et al. 2021). The proposed theories for GO efficacy in CD33^{low} CBF-AML patients cohort are the enhanced sensitivity of CBF-AML blasts to cytotoxic agents and the presence of CD33 at the surface of their LSCs (Walter et al. 2012).

1.3.3. FLT3 inhibitors

The prevalence of FLT3 mutations in the AML population, their established role in leukemogenesis, and association with poor prognoses made this kinase a desirable target for drug development. Midostaurin (April 2017) and gilteritinib (November 2018) were the first FLT3 inhibitors approved by the FDA for therapy in adult patients with relapsed or refractory (R/R) AML positive for *FLT3* mutations (Kim 2017; Dhillon 2019). They are now becoming a standard practice in *FLT3^{mut}* AML considered both at induction and consolidation phases (Kantarjian et al. 2021). Other highly selective FLT3 inhibitors such as quizartinib and crenolanib are currently evaluated in clinical trials both in the upfront and R/R settings.

Although FLT3 targeting showed great initial promise in the clinic, and FLT3 inhibitors achieved impressive potency and specificity, multiple resistance mechanisms have already emerged. The most common include on-target mutations, upregulation of alternative signaling pathways (e.g., mutations in RAS pathway and kinases *AXL* and *PIM1*), elevated paracrine signaling of FL in the bone marrow AML niche through the wild-type allele of *FLT3*, and drug clearance in the liver and bone marrow microenvironment (Daver et al. 2019; Eguchi et al. 2020).

1.3.4. IDH1/2 inhibitors

The inhibitors for isocitrate dehydrogenases *IDH1* and *IDH2* represent another set of FDA-approved drugs targeting AML-specific molecular alterations. Mutations in isocitrate dehydrogenases *IDH1* and *IDH2* are reported for up to 20% of AML patients (DiNardo et al. 2015). *IDH1/2* are metabolic enzymes that catalyze the conversion of isocitrate to α -ketoglutarate (α KG) accompanied by NADPH production (Martelli et al. 2020). The mutations result in the neomorphic activity of the enzymes: they consume NADPH to

convert α KG into R-2-hydroxyglutarate (R-2-HG). R-2-HG is an oncometabolite that accumulates in *IDH1/2^{mut}* AML cells at high levels and poisons the activity of enzymes, relying on α KG for their function. In particular, epigenetic modifiers, such as TET1/2 DNA demethylases and histone lysine demethylases (KDMs), are affected by the *IDH1/2* mutants, which results in DNA and histone hypermethylation. Other *IDH1/2^{mut}* specific changes include dysregulation of glutamine and tricarboxylic acid (TCA) cycle metabolism and multifaceted metabolic reprogramming (Carter et al. 2020).

IDH1^{R132H} knock-in allele was shown to cause myeloproliferative phenotype in mice, accompanied by the characteristic leukemic “hypermethylation signature” (Sasaki et al. 2012). The oncogenic role of *IDH2* mutation was directly demonstrated in mouse models first by acute genetic inactivation (Kats et al. 2014) and later with the development of *IDH2* specific inhibitors (Kats et al. 2017). Both *IDH1* and *IDH2* inhibitors lowered the levels of R-2-HG and induced terminal differentiation of cells, retaining the mutations. The success of pre-clinical trials and the ease of patient stratification studies prompted the rapid development of *IDH1/2* mutant specific inhibitors that soon passed the clinical trials for FDA approval (Martelli et al. 2020).

The activating mutations in *NRAS*, *KRAS*, and *PTPN11* are predictive of a decreased clinical response to *IDH1/2* inhibitors. Secondary resistance mechanisms also included phospho-signaling pathways and on-target mutations.

1.3.5. Hypomethylating agents (HMAs)

AML is characterized by dysregulated DNA methylation. Spencer et al. in 2017 attributed at least one such change, CpG island hypermethylation, to a byproduct of the rapid cellular proliferation, not inherently leukemogenic (Spencer et al. 2017). They argued that AML samples with the dominant-negative mutations in DNA methyltransferase 3A

(DNMT3A) exhibited dramatically reduced patterns of DNMT3A-driven CpG island hypermethylation, despite these mutations present in ~22% of *de novo* AML and significantly associated with poor prognoses (Ley et al. 2010). Nevertheless, DNA hypermethylation signature is characteristic for other common mutations in AML such as *IDH1/2* and *TET2*. The use of DNA HMAs, such as azacitidine or decitabine, has shown modest clinical benefits and has been approved by FDA for the first-line treatment of AML in patients unfit for conventional chemotherapies. In September 2020, azacitidine was also approved as first-of-kind maintenance therapy in AML patients in first remission after standard induction and consolidation regimens (Roboz et al. 2016).

1.3.6. Hedgehog (HH) pathway inhibitor: glasdegib

The non-canonical Hedgehog (HH) pathway has emerged as a target in AML only in the past decade. It is a well-conserved pathway in the embryogenesis regulating maintenance, fate determination, and regeneration of adult tissues (Ingham et al. 2011). In 2014 Zahreddine et al. showed that HH glioma zinc finger transcription factor *GLI1* mRNA levels were elevated at relapse and conferred drug resistance to AML cells (Zahreddine et al. 2014). The HH positive regulator Smoothed (SMO) and *GLI1* were also reported to be upregulated in radiation-resistant AML (Li et al. 2016). Later evidence suggested HH signaling to be implicated in LSCs maintenance (Zhao et al. 2009).

The SMO inhibitor glasdegib (PF-04449913) decreased leukemia initiation in serial transplantation assays in PDX mouse models (Fukushima et al. 2016). In clinical trials, glasdegib showed low toxicity and significant improvement in the overall survival of older patients with AML, resulting in its FDA approval in November 2018 for patients ≥ 75 years old or ineligible for intensive therapy (Hoy 2019). It is continued to be used in the clinic with moderate clinical benefits (Cortes et al. 2020).

1.3.7. Bcl-2 inhibitor: venetoclax

AML cytogenetic and molecular abnormalities often include activation of oncogenes (*FLT3*, *BCR-ABL1*, etc.) and inactivation of tumor suppressor genes (*TP53*, etc.). In addition to these driver mutations, cancer cells exhibit a variety of cellular stress phenotypes associated with the selective pressure for unlimited proliferation. These stress phenotypes need to be kept in check by non-mutated genes that cancer cells become dependent on. The arising context-specific cancer vulnerabilities are termed non-oncogene addictions (Luo et al. 2009). For example, cancer cells have to maintain a precarious balance between pro-apoptotic and anti-apoptotic machinery, paradoxically leaving them more susceptible to apoptosis (Sarosiak and Letai 2016).

Since the early 1990s, researchers reported that AML cells had increased BCL-2 expression associated with increased growth *in vitro* and poor prognoses. BCL-2 and MCL-1 (mentioned above as upregulated by the FLT3 kinase) are frequently exploited by AML cells to prevent the mitochondrial pathway of apoptosis (Carter et al. 2020). BCL-2 specific inhibitor venetoclax was designed to fit its BH3-binding site, mimicking the inactivating action of BH3-only proteins on BCL-2 and showed impressive selectivity against its target (Merino et al. 2018). FDA first approved it in 2016 as a monotherapy for patients with R/R chronic lymphocytic leukemia (CLL) with del(17p) (*TP53* deletion) (Deeks 2016). Venetoclax induced apoptosis in AML cell lines and primary patient samples in a BCL-2-dependent manner (Pan et al. 2014). While not as effective as monotherapy, venetoclax showed clinical benefits in combination with low-dose chemotherapy and HMAs and was approved in November 2018 for patients ≥ 75 years old or unfit for conventional chemotherapies (Apel et al. 2021). However, it is important to note that resistance to venetoclax regimens develops in less than 2 years (DiNardo et al. 2020a).

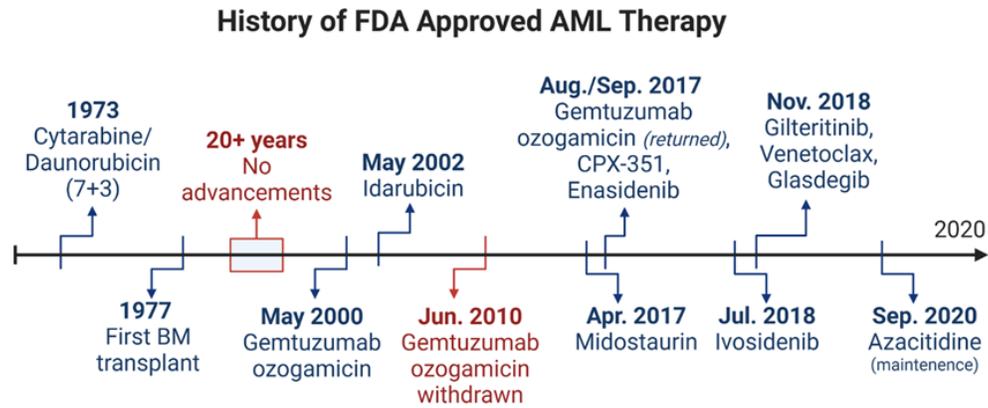


Figure C. History of AML therapies.

Timeline of approved clinical therapies in the United States for the treatment of AML. Figure from Carter et al. 2020, Fig. 1.

1.4. The search for non-oncogene addictions

This chapter introduces the reader to the high-throughput approach adopted by the Vakoc laboratory to search for novel genetic dependencies in cancer. The latter can include oncogene or non-oncogene addictions, and two examples of mechanisms for non-oncogene addiction are provided.

Many of the currently approved precision medicines against AML discussed in the previous section are based on the concept of oncogene addiction (FLT3, IDH1/2). However, tumor heterogeneity and secondary resistance challenge their long-term success in the clinic. The emerging strategy is identifying non-oncogene addictions contributing to a more extensive repertoire of cancer drivers, as exemplified by venetoclax. While not immune to resistance, venetoclax showed broad activity against AML with different drivers.

Some of these non-oncogene addictions arise as collateral vulnerabilities. For example, cancer cells may lose gene A with the deletion of a tumor suppressor gene or as a part of a genetic streamlining program. If gene A controlled one of the two redundant pathways important for cell fitness, the cells would become addicted to the genes supporting the remaining pathway. Other non-oncogene addictions are essential counterparts of cellular stress phenotypes constituting the cancer condition, such as the increased anti-apoptotic signaling discussed above or altered metabolic state of cancer cells (Luo et al. 2009).

1.4.1. Examples of mechanisms for non-oncogene addictions

1.4.1.1. Genetic streamlining

The genetic streamlining hypothesis postulates that cancer cells suffer from a burden of non-adaptive alterations while ensuring unlimited proliferation. For example, the increased hypermethylation observed in AML cells was recently implicated by our group in

two synthetic lethal non-oncogene addictions: the dependency on SLC5A3 myo-inositol transporter in the AML cells with silencing of *ISYNA1*, the rate-limiting myo-inositol biosynthesis enzyme; and the dependency on Fanconi anemia proteins in the AML cells that epigenetically silence the expression of *ALDH2* (Wei et al. 2020; Yang et al. 2021).

Another example is the hypermethylation of the CpG island at the promoter region of asparagine synthetase (*ASNS*) in ALL cells that prevents ATF4 recruitment and its subsequent transcription (Jiang et al. 2019). Asparagine depletion is an important therapy option in ALL patients, made possible via treatment with a purified bacterial enzyme L-asparaginase, depleting circulating asparagine (Avramis 2012).

AML cell lines that lack the expression of argininosuccinate synthetase 1 (*ASS1*) depend on extracellular arginine concentrations and were sensitive to arginine deiminases (Miraki-Moud et al. 2015; Mussai et al. 2015).

1.4.1.2. Altered metabolism

Cancer cells are characterized by metabolic reprogramming that serves their increased need for protein synthesis and energy production. One example of this metabolic reprogramming is the Warburg effect, whereas cancer cells exhibit high aerobic glycolysis using up pyruvate instead of a more efficient oxidative phosphorylation (OXPHOS) coupled ATP synthesis in the mitochondria (Warburg 1956). This apparent inefficiency was later appreciated as a necessary anabolic adaptation that allows cells to produce enough nutrients for proliferation via glycolysis interconnected pathways such as the pentose-phosphate pathway supplying cells with nucleotides (Hay 2016). Similar to other cancers, high glucose consumption and a characteristic glucose metabolic signature were observed in AML patients (Chen et al. 2014; Cunningham and Kohno 2016). AML oncogene *FLT3-ITD* was specifically shown to confer dependence on glycolysis and sensitivity to its pharmacologic

inhibition (Ju et al. 2017). Ye et al. proposed that leukemia cells could cause systemic glucose metabolism changes in patients through driving insulin resistance (Ye et al. 2018). Contrary to the original Warburg's hypothesis, both OXPHOS and TCA cycles are intact in cancer cells and also feed multiple biosynthetic pathways.

A few recent papers focused on the differences in metabolism between LSCs and bulk AML cells. First, LSCs, similar to HSCs, have been shown to contain less ROS levels as compared to the bulk AML blasts (Lagadinou et al. 2013). Second, these "ROS-low" cells with functional characteristics of LSCs preferentially relied on OXPHOS for ATP production as opposed to the bulk AML blasts that required glucose for their proliferation. OXPHOS is coupled with the TCA cycle, which, in *de novo* AML patients, is fueled by the constant supply and catabolism of amino acids, making these LSCs hypersensitive to the amino acid levels in contrast to bulk AML blasts or normal HSPCs (Jones et al. 2018). LSCs from R/R AML patients were demonstrated to develop metabolic plasticity allowing them to utilize other fuels and drivers for their energy metabolism, such as fatty acids and nicotinamide (Jones et al. 2018; Jones et al. 2020). The distinct metabolic states of different populations of AML cells led to context-specific dependencies, such as the novel requirement of R/R LSCs for nicotinamide phosphoribosyltransferase (NAMPT) – the rate-limiting enzyme in nicotinamide metabolism (Jones et al. 2020). Interestingly, Pereira et al. showed that different AML cell lines maintain distinct metabolic profiles, with some cell lines more dependent on OXPHOS, whereas others exhibited reliance on glycolysis. Thus, different AML cell lines could also potentially exhibit unique vulnerabilities that correlate with their energy metabolism (Pereira et al. 2018).

In the interest of time, we will only briefly outline some other examples of metabolic rewiring in AML with a specific focus on the alterations in amino acid metabolism.

Cancer cells were shown to become addicted to glutamine, which could be utilized

in addition to pyruvate to fuel the TCA cycle (Hensley et al. 2013). In AML cells, inhibition of glutaminase (GLS) catalyzing glutaminolysis, the first step of glutamine conversion into the critical TCA metabolite α KG, was reported to induce apoptosis and synergized with the BCL-2 inhibition (Jacque et al. 2015). Besides TCA, α KG is used by the cell for fatty acid synthesis, NADPH production, and protection from ROS via glutathione (Kreitz et al. 2019). Glutaminolysis was shown to act as a metabolic adaptation in response to TKIs in FLT3-ITD-driven leukemia and a metabolic dependency in IDH1/2^{mut} AML (Gallipoli et al. 2018; Matre et al. 2016).

Glutamine also serves as one of the central nodes in anabolism (Nicklin et al. 2009). The other glutaminolysis product, glutamate, acts as a nitrogen source for the biosynthesis of amino acids and nucleotides. Apart from amino acid biosynthesis, glutamine contributes to the transport of other amino acids into the cells, such as leucine, a major regulator of mammalian target of rapamycin mTORC1 nutrient sensor. One leucine molecule is transported into the cell via a heterodimeric bidirectional antiporter SLC7A5/SLC3A2 in exchange for the export of one glutamine molecule. The knockout of the high-affinity glutamine transporter ASCT2 (*SLC1A5*) was shown to impair leukemogenesis in human AML cell lines and in mouse AML models in contrast with minimal to moderate defects for the normal hematopoiesis (Willems et al. 2013; Ni et al. 2019). These effects were coupled with broad metabolic changes associated with the tight balance between glutamine levels in the cell and other amino acids biosynthesis and transport, as well as glycolytic and pentose phosphate pathways. In particular, permeable leucine analog rescued the *Slc1a5*-knockout phenotype through restoring mTORC1 signaling. Importantly, α KG failed to rescue *Slc1a5*-depleted leukemic cells whereas promoting survival of normal HSPCs with the *Slc1a5* knockout. This fact illustrates the AML-specific dependency on amino acid homeostasis independent of the other glutamine role in supporting aberrant glycolytic levels of AML.

Leucine and other branched-chain amino acids (BCAAs) can also be synthesized *de novo* via BCAA transaminases 1 (BCAT1) and 2 (BCAT2). Interestingly, BCAT1 was found to be upregulated in LSCs (Raffel et al. 2017). Raffel et al. found the increased BCAAs production to be associated with the decrease in α KG levels and the following downregulation of α KG-dependent epigenetic regulators resulting in DNA and histone hypermethylation similar to IDH1/2 mutations. BCAT1 overexpression and BCAAs were also implicated in CML progression from chronic to blast phase and *de novo* AML (Hattori et al. 2017).

FLT3-ITD was reported to drive *de novo* serine biosynthesis (increased expression of *PHGDH* and *PSAT1*) and uptake (increased expression of *SLC1A4* and *SLC1A5*) through regulating the mTORC1-ATF4 axis (Bjelosevic et al. 2021). *FLT3-ITD*-driven AML cells exhibited sensitivity to *PHGDH* inhibitors, potentially through the restriction of serine contribution to nucleotide biosynthesis.

1.4.2. Genetic screens for non-oncogene addictions

The cornerstone of modern cancer target discovery is functional genomics made possible by the Human Genome Project and the advances in the precise gene perturbation techniques. Simultaneous inactivation of multiple genes in isogenic backgrounds led to the rapid expansion of the universe of putative cancer targets and provided insights into already known and newly discovered molecular mechanisms and cellular functions.

1.4.2.1. Before CRISPR

Before the discovery of RNA interference (RNAi), large-scale genetic perturbations in mammalian systems were attempted via random mutagenesis, for example, the retroviral insertional mutagenesis, primarily useful for the discovery of oncogenes (Uren et al. 2005).

Later, transposon insertional mutagenesis was introduced, first the Sleeping Beauty, then the PiggyBac. All these systems are limited to their integration site preferences, but to this day, they serve as a powerful tool for *in vivo* forward genetic screens (Takeda et al. 2021).

RNAi, a conserved cellular pathway for the targeted protein expression downregulation, allowed high-throughput screening of hundreds of genes in mouse and human cancer cell lines. RNAi screens uncovered many novel therapeutic targets (Zuber et al. 2011b; Zuber et al. 2011c) and therapy resistance mechanisms. However, these techniques have limitations such as low resolution due to variable gene inactivation, dependent on the baseline expression level of the target, and serious off-target effects (McDonald et al. 2017).

1.4.2.2. CRISPR-Cas9

CRISPR-Cas9 system has proven to be a breakthrough innovation in gene editing, driving the explosive growth of applications and bringing its founders Emmanuelle Charpentier and Jennifer Doudna the Nobel Prize in Chemistry less than a decade after the discovery of its last component – tracrRNA – in 2011 (Deltcheva et al. 2011; Knott and Doudna 2018). CRISPR-Cas9 allowed high-resolution genome-wide loss-of-function screens in hundreds of cancer cell lines. Soon after the technology became accessible to independent research groups, multi-institutional consortia, such as DepMap, started a major effort in mapping cancer dependencies both through massively parallel technology implementation and providing open access to the results (Meyers et al. 2017; Dempster et al. 2019).

Some of the important considerations in designing CRISPR-Cas9 loss-of-function screens include predicting sgRNAs off-target effects (Hsu et al. 2013) and ensuring that the sgRNAs target exons that encode the functional domains of the proteins (Shi et al. 2015). The latter approach, described in detail by our group, dramatically increased the signal-to-noise ratio of the screens and even allowed *de novo* identification of cancer-relevant protein

domains upon tiling the exons of the candidate genes with potent sgRNAs.

Other targets, elusive for the straightforward screening approaches, are genes that share a common role in supporting a cell fitness pathway. These types of targets require double sgRNA library systems for combinatorial perturbation that have been described in the literature and independently designed in our laboratory (Wong et al. 2016; Han et al. 2017; Klingbeil, unpublished).

Most of the high-throughput screens carried out in the literature and in the public databases were performed in the standardized 2D cell cultures. There have been multiple reports in the literature on the discrepancies between these systems and more physiological mediums of cell culturing (albeit arguably less adaptable for the high-throughput screening approaches) (Boj et al. 2015; Kota et al. 2018; Chow and Chen 2018; Rossiter et al. 2020). However, the data from our laboratory suggest that 2D conditions largely recapitulate cell-autonomous mechanisms of cell proliferation, at least for AML (Wei et al. 2020).

1.5. HAD phosphatases in cancer

Our study identified an HAD phosphatase required for AML cell proliferation. This chapter introduces the reader to HAD phosphatases, their catalytic mechanism, and the few known examples of HAD phosphatases playing a role in cancer.

Haloacid dehalogenase (HAD) hydrolases, named after the archetypal prokaryotic enzyme, are present in all three domains of life. The superfamily encompasses a broad range of enzymes that catalyze the hydrolysis of diverse substrates, using an active site aspartate in nucleophilic catalysis (Burroughs et al. 2006). The majority of members of this superfamily serve as phosphatases (phosphate monoester hydrolases). They dephosphorylate a wide range of substrates: from low molecular weight metabolites to protein residues with exquisite specificities (Seifried et al. 2013).

The mammalian HAD phosphatase family members can be identified by the highly conserved catalytic core residues embedded in signature motifs. HAD signature motif 1 has the consensus sequence hhhDxDx(T/V)(L/V)h, where h is a hydrophobic residue, and x – any aminoacid. The two aspartate residues (Asp, D) in the active site are essential for the characteristic two-step catalytic mechanism (Fig. D), and coordinate the Mg²⁺ in the active site. Motif 2 has the consensus sequence hhh(S/T), and the conserved serine/threonine (Ser/Thr, S/T) residue forms a hydrogen bond with the phosphoryl group of the substrate to correctly position it for the nucleophilic attack. Motif 3 contains a conserved lysine (Lys, K) residue that stabilizes the negative charge of the phosphoryl group and the phospho-aspartyl intermediate together with Ser/Thr of motif 2. Motif 4 contains at least two of the Asp (D)/Glu (E)/Asn (N)/Ser (S) residues and hydrophobic residues. The conserved residues of Motif 4 together with Motif 1 coordinate Mg²⁺ in the active site.

Motifs 1–4 are positioned in a Rossmann-like (Rossmannoid) fold, characterized by a

β -sheet sandwiched between α -helices. Other conserved structural elements of the HAD phosphatases, such as the squiggle, flap, and cap domains, allow the enzymes to alternate between closed and open states of the active site cavity, regulating solvent access during the phosphoaspartyl transfer. Cap modules ensure substrate selectivity (Seifried et al. 2013).

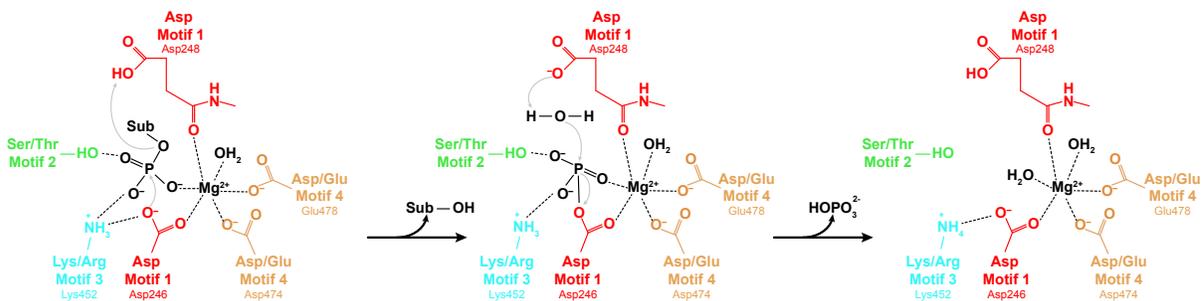


Figure D. Catalytic mechanism for HAD phosphatases.

Numbering indicates catalytic residues in Eya3, a murine HAD phosphatase. Sub – substrate.

Motif 1 residues are red, motif 2 residues are green, motif 3 residues are cyan, and motif 4 residues are orange.

The catalysis starts with binding the phosphorylated substrate mediated by Mg^{2+} ion and substrate specificity determinants outside of the active site (not shown). First, the catalytic Asp (Asp246 in Eya3) initiates a nucleophilic attack on the phosphoryl group of the substrate, forming a phosphoaspartyl intermediate stabilized by Mg^{2+} . Asp+2 (Asp248) serves as a general acid motif to protonate the leaving substrate. Second, the phosphoaspartyl intermediate is hydrolyzed by a nucleophilic attack of a water molecule, resulting in the release of the phosphate group and regeneration of the catalytic Asp. Asp+2, this time serves as a general base motif by deprotonating the water nucleophile.

Adapted from Jemc and Rebay 2007, Fig. 4C.

1.5.1. EYA phosphatases

The family of Eya proteins is one of the best studied in the HAD superfamily. Eya was originally identified in *Drosophila* as an essential transcriptional co-activator of the Six family of homeoproteins that regulated organ development (Pignoni et al. 1997; Chen et al. 1997). Soon followed the discovery that Eya belongs to the HAD superfamily and exhibits protein tyrosine phosphatase catalytic activity (Tootle et al. 2003), though it might not be required for its role in transcriptional activation after all (Jin et al. 2013). In mammals, Eya paralogs (Eya1-4) regulate the formation of many organs, and congenital mutations in Eya result in developmental abnormalities (Jemc and Rebay 2007).

The phosphatase activity of Eya proteins has been associated with multiple pathways important for cancer progression, such as promoting cancer cell proliferation and metastasis (Pandey et al. 2010), angiogenesis (Wang et al. 2016b), and activation of Sonic hedgehog (Shh) signaling in medulloblastoma (Eisner et al. 2015).

One of the best defined roles for Eya1-3 proteins is dephosphorylation of the C-terminal pY142 on histone H2AX, independently confirmed in 2009 by Tonks' and Rosenfeld's groups. H2AX phosphorylation drives DNA damage signaling priming cells for apoptosis. Dephosphorylation by Eya promotes DNA repair, making cells more resistant to DNA-damaging chemo- and radiotherapies (Krishnan et al. 2009; Cook et al. 2009). A few cancers are known to overexpress Eya, such as ovarian carcinomas and Ewing sarcomas (Zhang et al. 2005; Robin et al. 2012). Another potential substrate for Eya proteins is the estrogen receptor β (ER β) activating residue pY36. Active ER β acts as a tumor suppressor in cell lines and mouse xenograft models (Yuan et al. 2014), and high ER β ^{pY36} levels correlate with better prognoses in patients.

Abl kinase was shown to relocalize Eya to the cytoplasm by phosphorylating its PST-rich region (Xiong et al. 2009). Among the potential additional Eya functions in the

cytoplasm is the recently reported Eya3 interaction with the B55 α -subunit of PP2A, which results in c-Myc dephosphorylation at its pT58 residue, protecting it from degradation (Li et al. 2017; Zhang et al. 2018).

Given the diverse emerging roles of Eya phosphatases in human cancers, various research groups sought to develop inhibitors against these targets. The combination of high throughput screening and structural biology resulted in the discovery and design of the several lead compounds (Park et al. 2011; Tadjuidje et al. 2012; Pandey et al. 2013; Krueger et al. 2013) with some showing promise in cell line models (Wang et al. 2016b; Krueger et al. 2014).

1.5.2. PDXP/chronophin

The unique HAD phosphatases feature – cap domain – allows members of this phosphatase family to process a wide range of substrates. C0 caps are very small and do not restrict access to the open HAD active site. Therefore, C0 HAD members, such as CTD-phosphatases, are more likely to process macromolecular substrates that themselves would serve as a cap to exclude solvent during catalysis. Bulky C1/C2 cap modules are sufficient to restrict access to the catalytic cavity and allow HAD phosphatases to act on low molecular weight substrates trapped in the active site by cap closure (Seifried et al. 2013). Interestingly, for some capped phosphatases, such as PDXP (chronophin), protein phosphatase activity has also been reported despite the seemingly occluded active sites.

Pyridoxal phosphate (PDXP or chronophin) dephosphorylates pyridoxal 5'-phosphate (PLP), the active form of vitamin B6. It was initially purified from human erythrocytes in 1992 (Fonda 1992) and was later recognized as a C2-cap HAD phosphatase (Burroughs et al. 2006). PDXP regulates cellular PLP concentrations and consequently PLP-dependent enzymes. Increased PLP catabolism was associated with an increased risk of lung

cancer (Fanidi et al. 2018). Conversely, clear cell renal carcinomas and AML cells were shown to be dependent on the elevated PLP levels suggesting that in these contexts an increased PDXP function would be therapeutically beneficial (Ochocki et al. 2018; Chen et al. 2020). PDXP is often lost or hypermethylated in glioblastomas (Schulze et al. 2016), sometimes due to *IDH1/2*-mutations (Ceccarelli et al. 2016), also common in AML. It would be interesting to see whether similar PDXP promoter hypermethylation might be exploited by AML cells to sustain high PLP levels.

1.5.3. Other

There are few reliable reports on the roles of other HAD phosphatases in cancer.

CTDSP2 (encoding phosphatase SCP2) is often co-amplified in sarcomas with *CDK4* (Su et al. 1997), but its role as a driver mutation in this disease has not been definitively shown.

SCP3 was reported to activate the tumor suppressor retinoblastoma protein 1 (pRb1) in cells via dephosphorylation, thereby inhibiting cell cycle progression (Kashuba et al. 2004). Its gene, *CTDSPL*, is deleted or mutated in the majority of human carcinomas, including small cell lung cancer and renal cell carcinomas, and exhibited tumor suppressive phenotype in mice.

Gain-of-function mutations in 5'-nucleotidase, *NT5C2*, have been implicated in chemotherapy resistance in ALL since they were specifically acquired at relapse (Oshima et al. 2020). *NT5C2*^{mut} can inactivate nucleoside-analog chemotherapy drugs in the cytosol.

Polynucleotide 5'-kinase/3'-phosphatase (*PNKP*) is involved in DNA repair and confers cancer cells protection from ionizing radiation. This prompted the development of small-molecule inhibitors for PNKP for their potential utility in sensitizing human tumors to gamma-radiation (Freschauf et al. 2009).

L-3-Phosphoserine phosphatase (*PSPH*) upregulation was associated with squamous cell carcinoma in an insertion-mutagenesis transgenic mouse model and was shown to be a dependency in cell lines, albeit independent of its role in serine biosynthesis (Bachelor et al. 2011). The serine biosynthetic pathway is itself a vulnerability in many cancers, as shown by RNAi and CRISPR screens (Possemato et al. 2011; Meyers et al. 2017; Dempster et al. 2019). For example, serine biosynthesis was recently reported to be a dependency in FLT3-ITD-driven acute myeloid leukemia (Bjelosevic et al. 2021). A selective inhibitor of PSPH based on a structural analogy with L-3-Phosphoserine has been described and could be potentially explored in cancers dependent on the *de novo* serine biosynthesis (Hawkinson et al. 1996).

1.6. Current state of research on the topic of this dissertation

The main focus of this dissertation – phosphatase SCP4 – is a relatively poorly studied protein with no prior reports of SCP4 functions in the context of human AML. Similarly, two interacting partners of SCP4 – paralog kinases STK35 and PDIK1L – uncovered in the course of our study have no clearly defined roles in cell signaling. This chapter summarizes most of the published literature mentioning either one of these enzymes and emphasizes the novelty of our discovery.

1.6.1. SCP4

As of May 2021, a search of the scientific literature in PubMed using SCP4 aliases “HSPC129” OR “CTDSPL2” OR “SCP4” returns just 10 research papers after it was first cloned among other cDNAs in 2000, ironically, from CD34⁺ HSPCs (Zhang et al. 2000). Its length was correctly determined to be 466 amino acids. The next unbiased study exploring subcellular localizations of N-terminally FLAG-tagged cDNAs of unknown cellular functions showed striking localization for FLAG-SCP4 at the nuclear rim of human liver cancer cells. The latter led Kemmer et al. to propose its association with the nuclear pore complex (Kemmer et al. 2006).

The first study fully dedicated to SCP4 by Qian et al. appreciated the conservation of the catalytic residues between FCP1, SCP1, and SCP4 (Qian et al. 2007). All three enzymes belong to the HAD family of phosphatases. Although the previous chapter of this dissertation introduces the same conserved residues as largely invariant for the catalytic mechanism of most HAD phosphatases, Qian et al. proposed that SCP4 might perform a similar function to FCP1 and SCP1, which by that time was established as dephosphorylation of the C-terminal domain (CTD) tail of the largest subunit of RNA polymerase II (Pol II). Thus, the gene and protein names for the entry became CTD small

phosphatase-like 2 (CTDSPL2) or small CTD phosphatase 4 (SCP4). Qian et al. optimized the *E. coli* (BL21) purification protocol, including truncation of the first 156 N-terminal amino acids that interfered with its expression in bacteria and confirmed its strong phosphatase activity *in vitro* against p-nitrophenyl phosphate and both phosphoserine-5 (pSer5) and phosphoserine-2 (pSer2) of GST-CTD of Pol II.

The first paper implying a role for SCP4 in hematopoiesis came out in 2010. Ma et al. noticed that SCP4 was expressed at a higher level in erythroid cultures from umbilical cord blood than in the adult bone marrow blood, and its expression was upregulated upon hemin-induced erythroid differentiation of K562 (Ma et al. 2010). However, their SCP4 RNAi knockdown experiments showed that SCP4 was largely dispensable for that process. Later work, discussed in more detail below, reported relocalization of SCP4 from the nucleus into the cytoplasm in K562 cells treated with hemin for 48 hours (Wani et al. 2016).

In 2014 Zhao et al. presented data suggesting that SCP4 could dephosphorylate transcription factors SMAD1/5/8 in the canonical bone morphogenetic protein (BMP) signaling pathway *in vivo* and *in vitro*, but not Pol II *in vivo* (Shen et al. 2014; Zhao et al. 2014). Through dephosphorylation of the C-terminal SXS motif of SMADs and subsequent SMADs nuclear exclusion, SCP4 was reported to negatively regulate BMP signaling induced osteoblastic differentiation (Zhao et al. 2014). Notably, just the predicted phosphatase-domain fragment SCP4²⁶³⁻⁴⁶⁶ used in that study could not dephosphorylate SMAD1 *in vitro* in contrast to the full-length SCP4 while being expressed and retaining its phosphatase activity. The study also reported SMAD1/SCP4 direct binding positively correlating with the SMAD1 phosphorylation status.

The next relevant data on SCP4 came from a 2016 study by Wani et al., who first showed that besides preferential nuclear localization, also confirmed in most previous works, SCP4 exclusively localized to the chromatin fraction following histone acid

extraction protocol for fractionation (Wani et al. 2016). They reported ChIP-qPCR and immunofluorescence staining results suggesting the broad distribution of SCP4 over the chromatin with some preference towards regions with the enrichment of the H3K27me3 heterochromatin associated histone mark.

A 2017 study started a series in the literature on SCP4 potential involvement in FoxOs transcription factors regulation (Liu et al. 2017). In murine myoblasts cell line C2C12, Liu et al. showed evidence that SCP4 dephosphorylated FoxO1/3a transcription factors, promoting their retention in the nucleus. These findings were corroborated the following year *in vivo* and *in vitro* by Cao et al. (the same group that reported SMADs as potential SCP4 substrates) (Cao et al. 2018). Once again, this research group showed stable binding between SCP4 and its potential substrates FoxO1/3a in co-IP studies. Ectopic expression of SCP4 in a mouse immortalized fibroblasts cell line NIH3T3 resulted in the disruption of the dynamic FoxO1/3a nucleus-cytoplasm shuttling in serum starvation experiments dependent on SCP4 catalytic activity suggesting a physiological role for FoxO1/3a dephosphorylation by SCP4 in promoting nuclear FoxO1/3a retention. More importantly, Cao et al. generated *SCP4* knockout mice using an *SCP4* gene-trap construct. *SCP4*^{-/-} mice completed embryonic development but died within 24 hours after birth due to glucose deficiency. Cao et al. proposed that SCP4 was upregulated by glucose deprivation to positively regulate FoxO1/3a transcriptional activity in promoting gluconeogenesis under physiological conditions. Of note, the reported phenotype for *SCP4*-null mice is strikingly different to *FoxO1*-null and *FoxO3a*-null mice. *FoxO1*-null mice perish just before embryonic day 11 due to failures in vascular development, whereas *FoxO3a*-null mice are viable and look similar to their littermate controls (Hosaka et al. 2004).

The only study so far directly showing a potential role of SCP4 in cancer was done by Winans et al. in 2017. In an avian leukosis virus (ALV)-induced B-cell lymphoma model,

scp4 was shown earlier to be a common ALV-integration site (Justice et al. 2015). The viral promoter drove transcription of an almost intact *scp4* with only the first 63 amino acids missing (of note, the identity between human and chicken proteins is >93%). Winans et al. also reported that *scp4* ectopic expression confers some protection to chick embryo fibroblasts from apoptosis induced by oxidative stress.

Overall, our understanding of the cellular roles of SCP4 in various contexts is currently limited. First of all, there is no definitive evidence that SCP4 is indeed associated with Pol II *in vivo* during transcription. In this regard, the functional role of SCP4 could be elucidated through global ChIP-Seq experiments using well-established antibodies for Pol II total levels and Ser2/5/7 phosphorylation upon SCP4 knockout/overexpression. Similarly, there is still no ChIP-Seq data for SCP4, given its remarkable association with chromatin. No study has yet attempted unbiased RNA-Seq or proteomics analysis for SCP4 knockout or catalytically dead samples. 3/10 studies on SCP4 were performed by the same group that proposed new substrates for SCP4 in each new context, suggesting that SCP4 functions in the cell are largely context-dependent.

1.6.2. STK35/PDIK1L

STK35 and PDIK1L belong to a distinct New Kinase Family 4 with largely unknown functions. Interestingly, they are a relatively new evolutionary branch reliably detected (and relatively well conserved) only in vertebrates (Goyal et al. 2009). According to the Human Protein Atlas, STK35 and PDIK1L localize to the nucleoplasm, but the functions of these enzymes in this cellular compartment are unclear.

STK35 has been shown to confer a proliferative advantage in human colorectal cancer and osteosarcoma via suppressing apoptosis and promoting glycolysis (Wu et al. 2018; Yang et al. 2020). High expression of STK35 was associated with unfavorable prognosis

in the liver and colorectal cancers. In endothelial cells, STK35 supported cell migration in the wound healing assays (Goyal et al. 2011). In humans, STK35 was shown to have a higher expression in the retina and testis, and STK35 knockout mice exhibited a range of defects in spermatogenesis and eye development (Miyamoto et al. 2018). Interestingly, SCP4 was also shown to have a relatively higher expression in the testis in the single cell studies, specifically in spermatids (Human Protein Atlas: Qian et al. 2007; Uhlen et al. 2015). Serum PDIK1L levels have been implicated in endometriosis detection (Greenbaum et al. 2021).

Overall, the signaling function and biochemical interactions of STK35/PDIK1L remain unknown.

2. RESULTS

2.1. Phosphatase domain-focused CRISPR screening identifies context-specific dependencies in human cancer cell lines

A prior study from our laboratory described a domain-focused CRISPR-Cas9 genetic screening strategy that can nominate novel dependencies in cancer by targeting insertion/deletion (indel) Cas9-mediated mutagenesis to domain-encoding exons (Shi et al. 2015). Our group has previously applied this method to identify well-established and novel phosphocatalytic dependencies in AML among human kinases (Tarumoto et al. 2018). For this dissertation, we set out to discover AML-specific dependencies among human phosphatases.

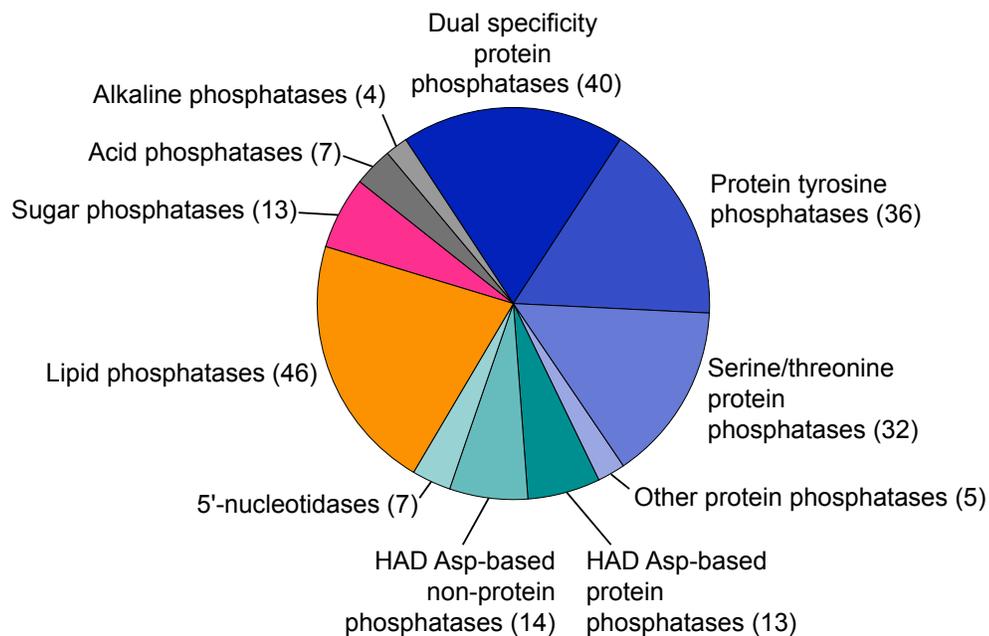
We have first selected all the genes comprised by the group “Phosphatases” according to the HUGO Gene Nomenclature Committee at the European Bioinformatics Institute. This database in 2016 included 217 genes that encoded proteins with phosphatase domains with protein and non-protein substrates (Fig. 1A; Appendix 1). Using the CSHL CRISPR sgRNA design tool (El Demerdash, unpublished), we manually curated 1,958 sgRNAs targeting exons that encoded phosphatase domains of the 217 genes. The library was supplemented with positive and negative control sgRNAs and cloned in a pooled manner. We used this bespoke library to perform negative selection ‘dropout’ screens in eight leukemia and four solid tumor human cancer cell lines (Fig. 1B; Fig. 2; Appendix 2). All the cell lines used in the screens have been previously derived in our laboratory from the original cell lines to stably express human codon-optimized *Streptococcus pyogenes* Cas9 protein. The performance of spike-in negative and positive control sgRNAs validated the overall quality of these screens (Fig. 3).

The screens revealed that sgRNAs targeting 187 of the phosphatase domains showed no significant impairment of cell fitness in any cell lines tested, similarly to the other screens

previously performed in our laboratory. 11/217 (5.1%) phosphatases were essential in ten or more cell lines (Fig. 2; Appendix 2) and are likely required for general cell survival. For example, FCP1 (*CTDP1*) and SSU72 phosphatases act on the C-terminal domain of the largest Pol II subunit and regulate transcription initiation, elongation, and termination (Cossa et al. 2021). The remainder of the targeted phosphatases showed varying degrees of specificity, which included the validation of known context-specific phosphatase dependencies. For example, the requirement for the phosphatase SHP2 (*PTPN11*) correlated with the RAS mutation status. KRAS G12D, G13D, and NRAS Q61L/H mutations conferred resistance to SHP2 CRISPR knockout (Fig. 4A), which corroborates and extends previous similar findings (see Introduction §1.2.2.7). Only the cell lines with the wild-type *TP53* showed dependence on WIP1 (*PPM1D*) (Fig. 4B). WIP1 is known to negatively regulate the p53 signaling pathway (see Introduction §1.2.2.12). We have also confirmed the non-oncogene addiction of the Philadelphia chromosome-positive CML cell line K562 to PTP1B (*PTPN1*) (Fig. 4C), consistent with the data suggesting that PTP1B supports the function of BCR-ABL1 protein (see Introduction §1.2.2.10).

Next, we ranked all of the phosphatases based on their relative essentiality in AML versus non-AML contexts, which nominated SCP4 (*CTDSPL2*) as the top-ranking AML-biased dependency (Fig. 2). Among the AML cell lines screened, five showed SCP4 dependency, with MOLM-13 being the most sensitive. In contrast, AML cell lines U937 and Kasumi-1 and the CML line K562 did not show significant cell fitness defects upon SCP4 targeting. None of the evolutionary close SCP subfamily members showed deleterious phenotype upon perturbation, except for the *CTDSP2* gene that is located in the 12q13-15 amplicon of the RMS cell line Rh30 (Fig. 4D) and thus constitutes a CRISPR-artifact due to an increased number of double-strand breaks. Since no prior study has implicated SCP4 in human cancer, we investigated further the essential function of this phosphatase in AML.

A



B

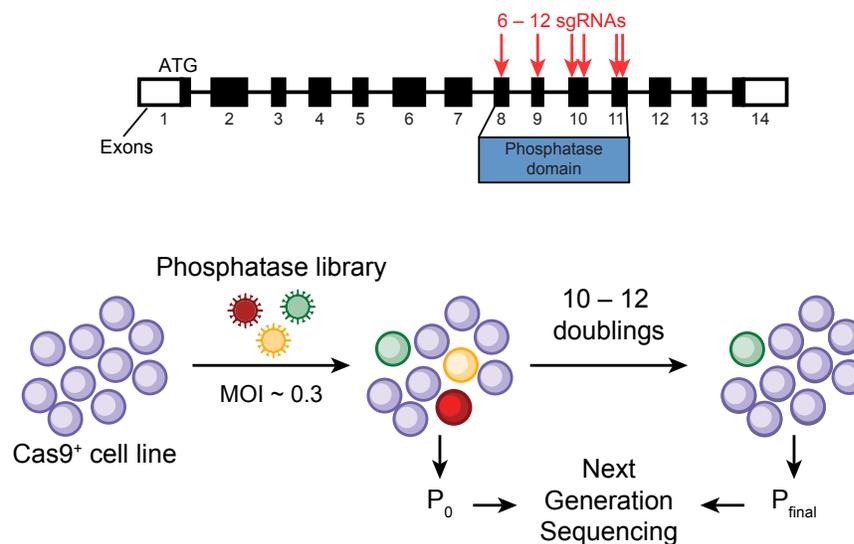


Figure 1. Phosphatase domain-focused CRISPR screen library design and negative selection 'dropout' screen schematics.

(A) Categorization of human phosphatases targeted by CRISPR-Cas9 genetic screens in this study.

(B) Overview of the genetic screening procedure using the phosphatase domain-focused sgRNA library.

RESULTS

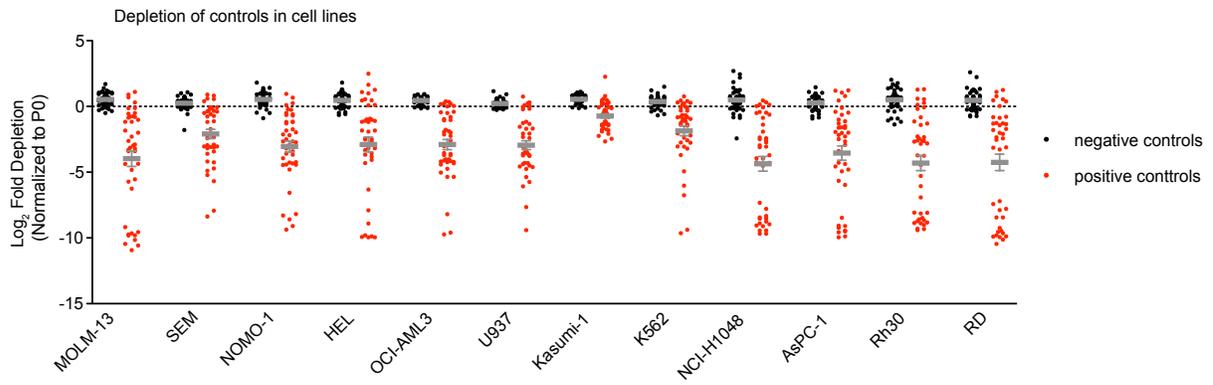


Figure 3. The performance of spike-in negative and positive control sgRNAs.

The log₂ fold-changes of individual sgRNAs serving as negative (black) or positive (red) controls in the screen.

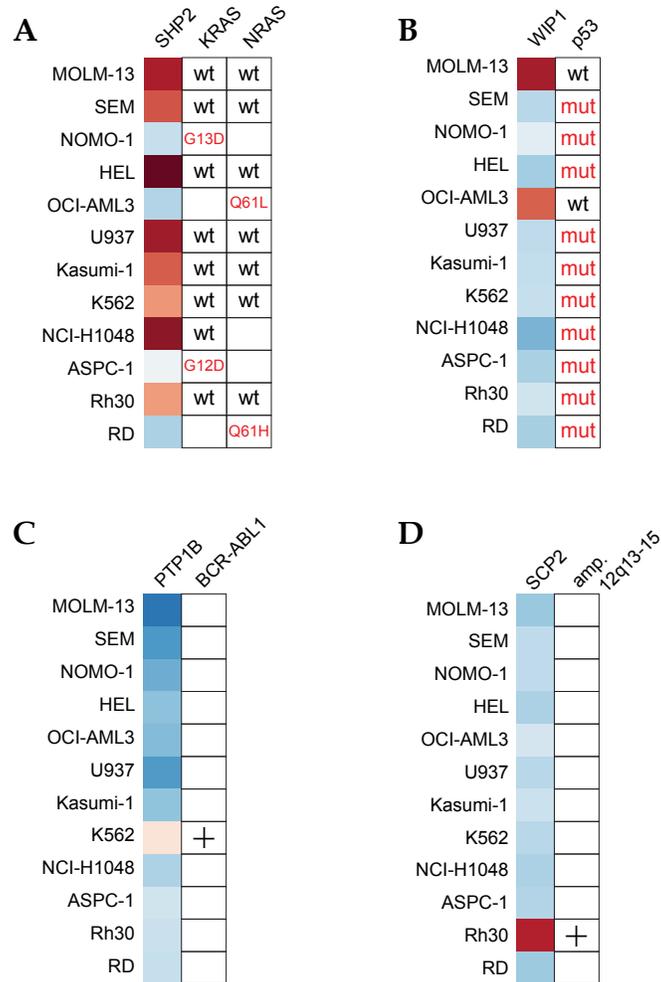


Figure 4. Phosphatase domain-focused CRISPR screening identifies context-specific dependencies in human cancer cell lines.

Extracted essentiality scores for selected phosphatases. Plotted is the \log_2 fold-change of sgRNA abundance during ~11 population doublings. The effects of individual sgRNAs targeting each domain were averaged. Also shown is the relevant oncogene/tumor suppressor status for each cell line.

2.2. SCP4 is an acquired dependency in human AML cells

To validate the AML-biased essentiality of SCP4, we performed competition-based proliferation assays comparing SCP4 sgRNAs to positive (ROSA) and negative (PCNA) control sgRNAs in a total of 23 human cancer cell lines: 14 leukemia, 4 pancreatic cancer, 3 rhabdomyosarcoma, and 2 lung cancer lines. All sgRNA experiments were carried out in Cas9-expressing cell lines. The indicated sgRNAs were linked to GFP, and the GFP⁺ population depletion indicated the loss of cell fitness caused by Cas9/sgRNA-mediated indel mutations in the endogenous gene sequence (Fig. 5A).

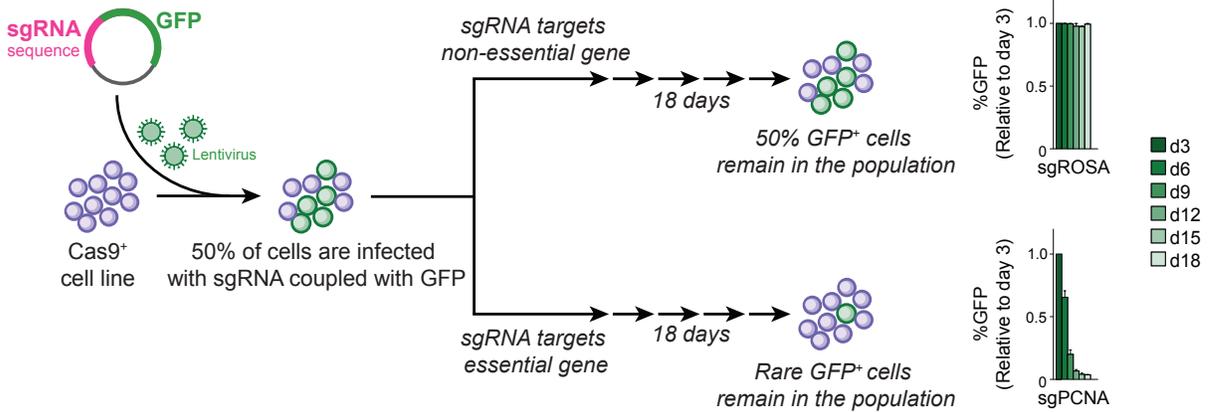
In these assays, ten leukemia cell lines demonstrated a clear dependency on SCP4, whereas four of the leukemia cell lines and all of the solid tumor cell lines were insensitive to SCP4 sgRNAs at the same timeframe (Fig. 5B). The sensitive cell lines included engineered human MLL-AF9/Nras^{G12D} and MLL-AF9/FLT3^{ITD} AML cell lines (Wei et al. 2008). In contrast, an sgRNA targeting the DNA replication protein PCNA suppressed the growth of all cell lines tested. We confirmed that the differential SCP4 requirement was not due to any differences in genome editing efficiency through Western blots (Fig. 6). We have also performed rescue experiments using a cDNA encoding N-terminally FLAG-tagged sgRNA-resistant SCP4 sequence to ensure that the growth arrest caused by SCP4 sgRNAs occurred via an on-target mechanism (Fig. 7).

In order to better characterize the AML cell phenotype following SCP4 knockout, we assessed cell cycle progression and apoptosis. Using flow cytometry-based BrdU incorporation measurements, we found that SCP4 inactivation in MOLM-13 cells by day 5 post-infection with sgSCP4 led to a G1/G0-arrest judged by the accumulation of the cells in the G1 phase of cell cycle and the depletion of cells progressing into the S phase (Fig. 8A,B). Concurrently, Annexin V staining revealed increases in the early and late apoptosis in the same population (Fig. 8C,D).

To extend our findings into a more physiological context, we validated that targeting SCP4 attenuated the growth of MOLM-13 cells when injected via tail vein and expanded in immune-deficient mice (Fig. 9). Therefore, the dependency on SCP4 is not altered by the *in vitro* versus *in vivo* growth conditions (the possibility discussed in Introduction §1.4.2.2).

We verified that Scp4 was essential in an engineered murine AML cell line MLL-AF9/Nras^{G12D} (RN2) (Zuber et al. 2011a), an effect that was rescued by the expression of the human SCP4 cDNA (Fig. 10A). However, immortalized NIH3T3 murine fibroblasts were unaffected by SCP4 sgRNAs (Fig. 10B), suggesting that the AML-biased dependency on SCP4 is conserved across species.

A



B

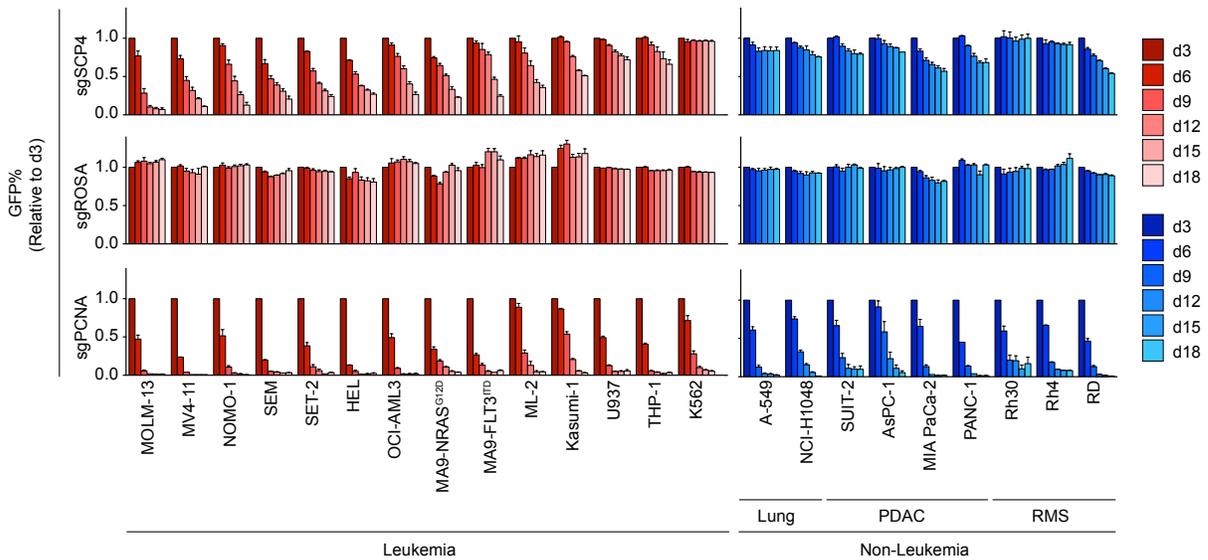


Figure 5. SCP4 is an AML-biased dependency.

(A) Scheme of a competition-based proliferation assay for a negative control (ROSA) and a positive control (PCNA). (B) Competition-based proliferation assays in 23 human cancer cell lines infected with the indicated sgRNAs. PDAC, Pancreatic Ductal Adenocarcinoma; RMS, rhabdomyosarcoma. n=3.

All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. ‘e’ refers to the exon number targeted by each sgRNA. The indicated sgRNAs were linked to GFP. GFP⁺ population depletion indicates loss of cell fitness caused by Cas9/sgRNA-mediated genetic mutations. ROSA, negative control; PCNA, positive control.

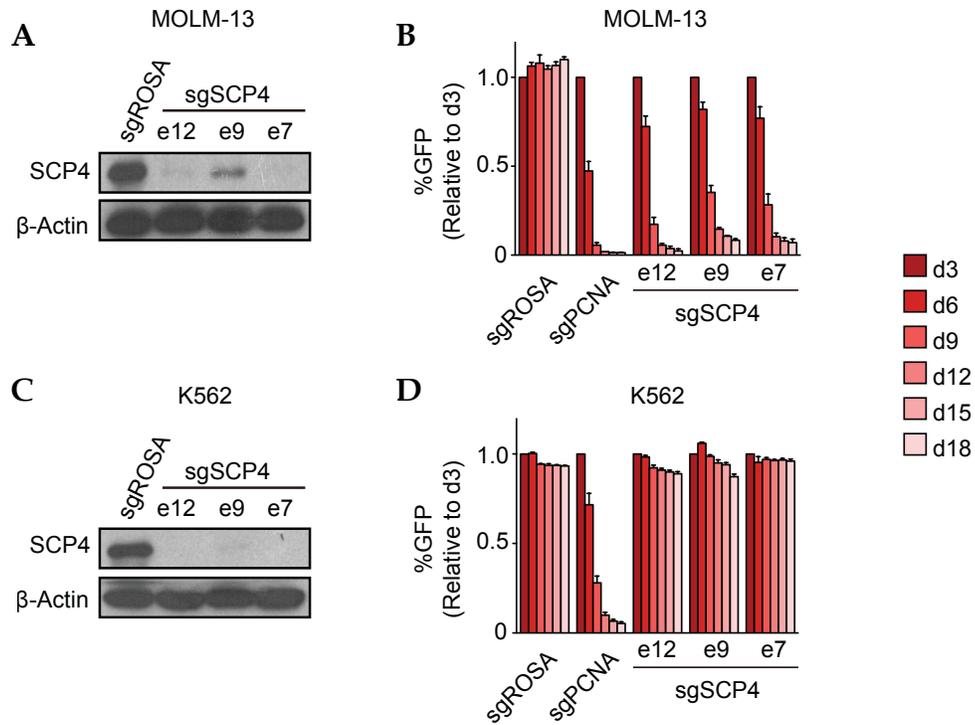


Figure 6. The differential requirement for SCP4 in leukemia cells.

(A) Western blot of whole-cell lysates from MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. **(B)** Competition-based proliferation assay in MOLM-13 cells infected with the indicated sgRNAs. n=3. **(C)** Western blot of whole-cell lysates from K562 cells on day 5 post-infection with the indicated sgRNAs. **(D)** Competition-based proliferation assay in K562 cells infected with the indicated sgRNAs. n=3.

All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. 'e' refers to the exon number targeted by each sgRNA. The indicated sgRNAs were linked to GFP. GFP⁺ population depletion indicates loss of cell fitness caused by Cas9/sgRNA-mediated genetic mutations. ROSA, negative control; PCNA, positive control.

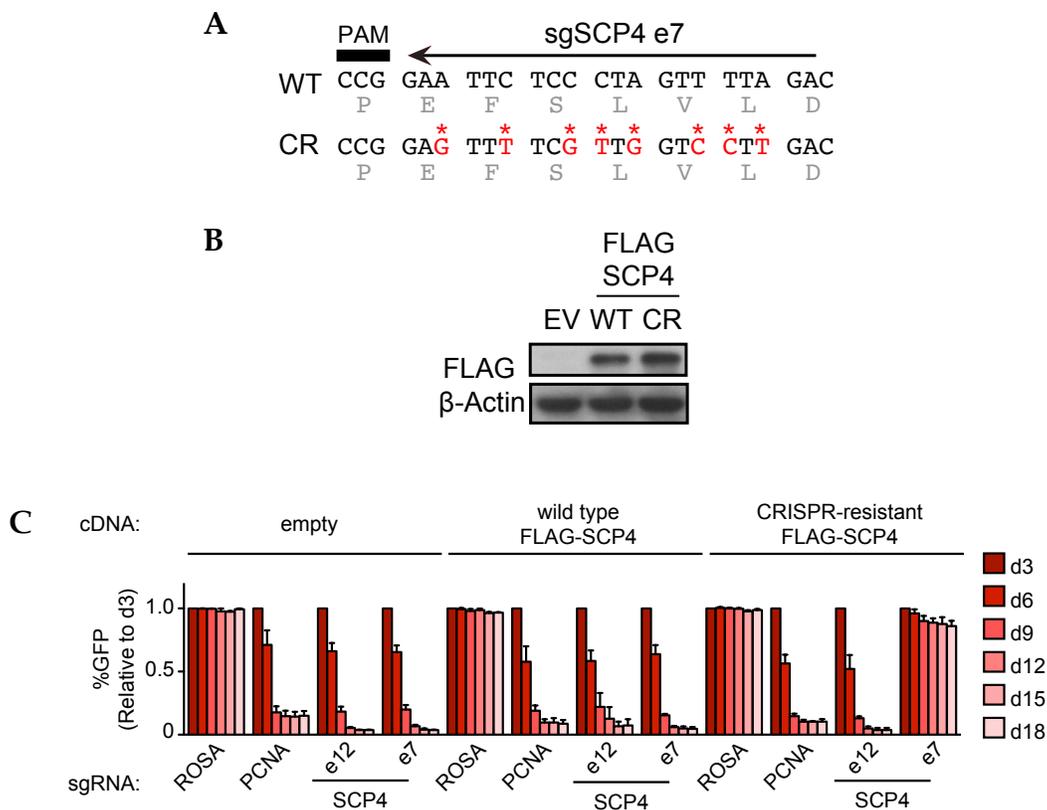


Figure 7. Validation of the on-target effect of SCP4 knockout.

(A) Design of CRISPR-resistant mutant of SCP4. **(B)** Western blot of whole-cell lysates from MOLM-13 cells stably expressing empty vector (EV), wild-type (WT), or CRISPR-resistant (CR) FLAG-SCP4. **(C)** Competition-based proliferation assay in MOLM-13 cells stably expressing empty vector, wild-type, or CRISPR-resistant FLAG-SCP4 infected with the indicated sgRNAs. $n = 3$.

All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. 'e' refers to the exon number targeted by each sgRNA. The indicated sgRNAs were linked to GFP. GFP⁺ population depletion indicates loss of cell fitness caused by Cas9/sgRNA-mediated genetic mutations. ROSA, negative control; PCNA, positive control.

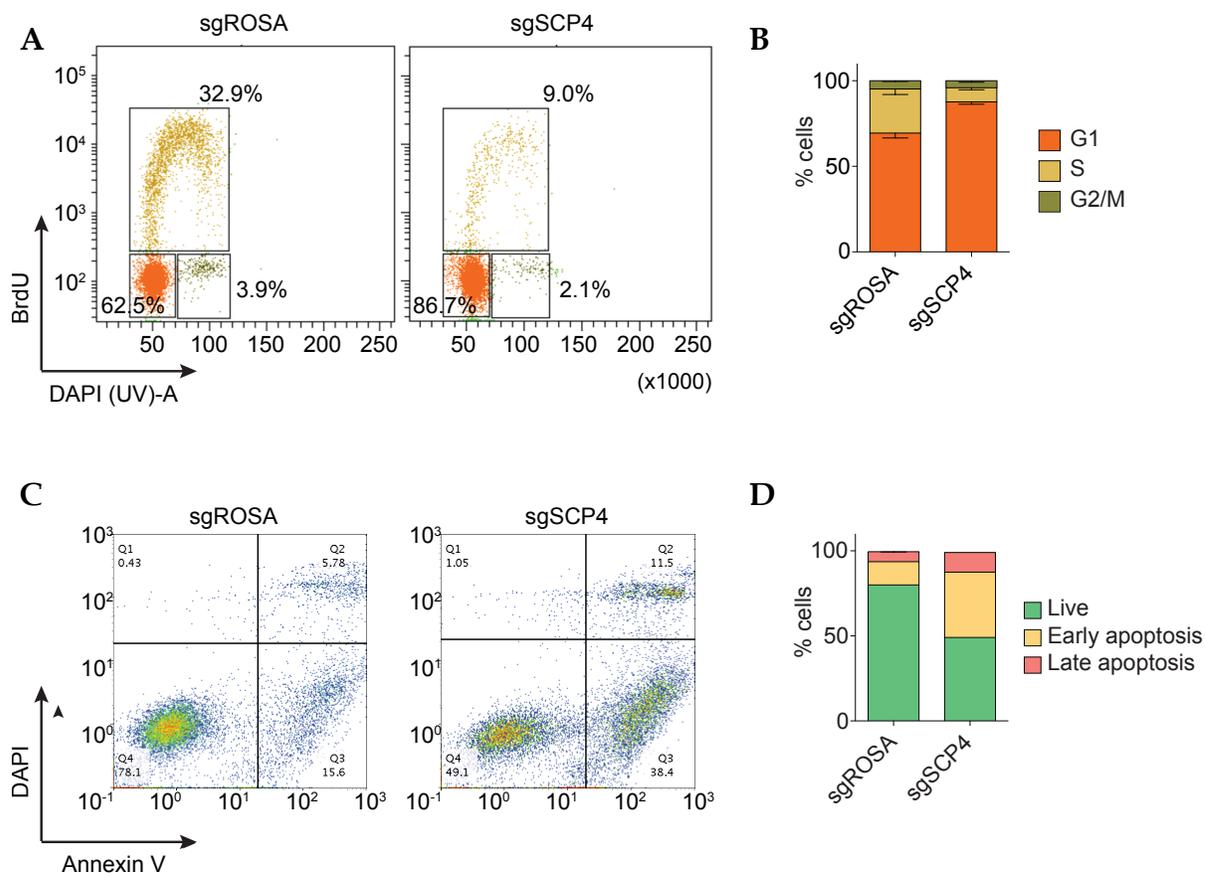


Figure 8. SCP4 inactivation in MOLM-13 cells leads to a G1/G0-arrest and the induction of apoptosis.

(A) Representative flow cytometry analysis of BrdU incorporation and DNA content to infer cell cycle stage in MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. (B) Quantification of the different cell cycle stages in MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. $n = 3$. (C) Representative flow cytometry analysis of DAPI (indicating permeable dead cells) and annexin-V staining (a pre-apoptotic cell marker) in MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. (D) Quantification of live and apoptotic cells.

All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. ROSA, negative control.

Data collected by Zhaolin Yang.

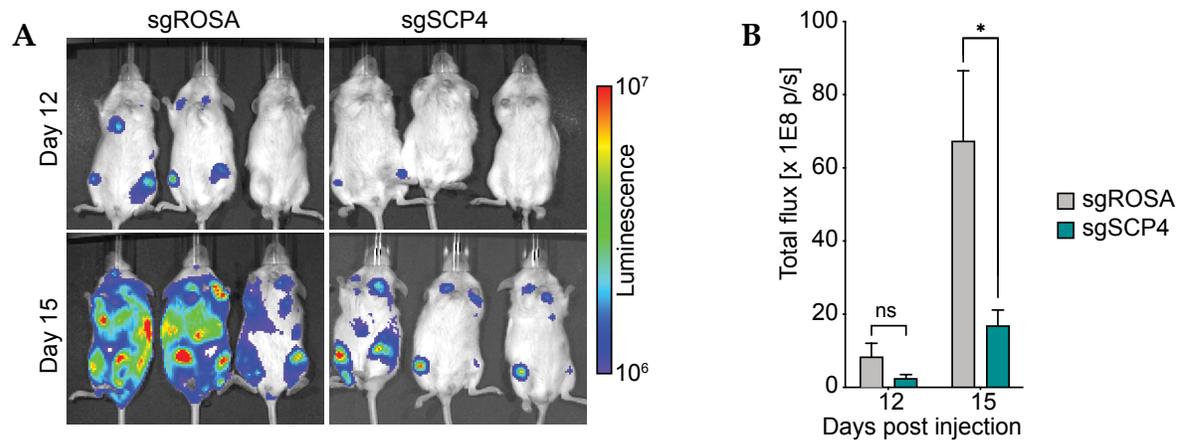


Figure 9. Targeting SCP4 attenuates the growth of MOLM-13 cells in xenograft mouse model of AML.

(A) Bioluminescence imaging of NSG mice transplanted with luciferase⁺/Cas9⁺ MOLM-13 cells infected with either sgROSA or sgSCP4. (B) Quantification of bioluminescence intensity. n=3. p-value was calculated by unpaired Student's t-test. *p < 0.05.

All bar graphs represent the mean \pm SEM. ROSA, negative control.

Data collected by Yiliang Wei.

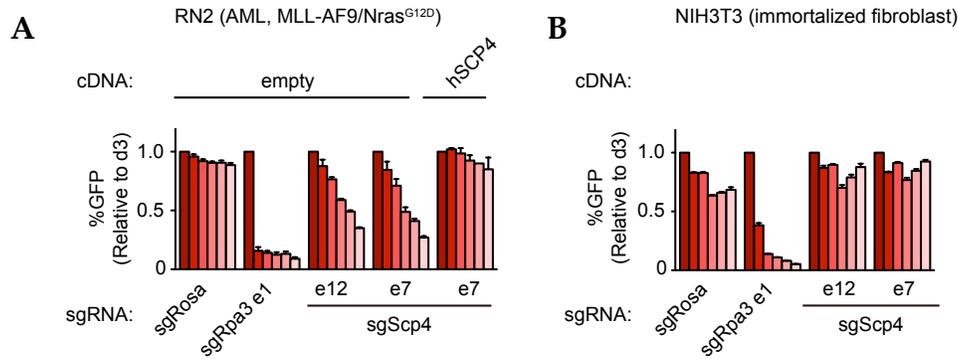


Figure 10. AML-biased dependency on SCP4 is conserved across species.

Competition-based proliferation assays in murine RN2 MLL-AF9/Nras^{G12D} AML cells (A) and NIH 3T3 immortalized fibroblasts (B) infected with the indicated sgRNAs. Rescue in the RN2 cells stably expressing human SCP4 is shown. $n=3$.

All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. ‘e’ refers to the exon number targeted by each sgRNA. The indicated sgRNAs were linked to GFP. GFP⁺ population depletion indicates loss of cell fitness caused by Cas9/sgRNA-mediated genetic mutations. Rosa, negative control; Rpa3, positive control; hSCP4, human SCP4^{wt}.

2.3. SCP4 is dispensable in normal hematopoietic progenitor cells

Blood is one of the few human organs that requires constant replenishing. Signaling alterations in AML can include genes and pathways essential for normal haematopoietic stem cell differentiation and proliferation of the early haematopoietic progenitors, thus posing significant challenges for the targeted therapies development. To evaluate the SCP4 requirement in non-transformed hematopoietic cells, we set up the collaboration with the group of Dr. Mitchell Weiss, the Chair of Hematology Department at St. Jude Children's Research Hospital (Memphis, Tennessee). Together with Weiss' group we performed electroporation of Cas9:sgRNA ribonucleoprotein (RNP) complexes into human CD34⁺ HSPCs (Metais et al. 2019). Since sgRNA presence in the cells in this experiment was not coupled with the GFP reporter expression, we have optimized an RT-qPCR assay to measure the presence of indels at the genomic loci targeted by each sgRNA. To this end, we designed pairs of RT-qPCR primers with one primer spanning the Cas9 endonuclease cleavage site and the second 100-150 bp away. We benchmarked RT-qPCR-based indel measurements against Western blot of MOLM-13/Cas9⁺ cells infected with our standard lentiviral vector encoding sgRNA linked to GFP at two different infection rates (Fig. 11A-C). RT-qPCR measurements of editing efficacy corresponded to the Western blot results of SCP4 depletion in the same populations, as well as their infection rates.

The Cas9:sgRNA RNP complexes electroporation into human CD34⁺ HSPCs resulted in a mixed population of cells with an average of 60% cells harboring on-target indel mutations on day 5 post-electroporation as estimated by RT-qPCR assay for each replicate. As a control, a Mock population (not electroporated) and the population electroplated with Cas9:sgRNA RNP harboring a non-targeting (NT) gRNA were used. Immediately after electroporation, the cells were divided into myeloid, erythroid, and megakaryocyte lineages or mixed with methylcellulose for the colony-forming cell (CFC) assay (Fig. 11D).

Each lineage was supplemented with the cytokines promoting its respective differentiation and propagated in culture for 16 days. Throughout differentiation into each lineage, the genomic DNA was collected from the cells for indel measurements, and Western blot samples were made on days 5 and 8 of differentiation. Unlike the effects of targeting the essential transcription factor MYC, we observed no evidence that SCP4-deficient cells experienced any fitness disadvantage during 16 days of culturing under conditions that promoted myeloid, erythroid, or megakaryocytic differentiation (Fig. 11E-G). For myeloid lineage, we additionally showed the presence of indels throughout the course of differentiation with an orthogonal method using Surveyor nuclease assay (Fig. 11K). More importantly, Western blots clearly showed decreased SCP4 protein levels on day 8 – the day by which the cell populations got committed to their respective lineages (Fig. 11H-J; Fig. 12).

Lineage-specific *in vitro* differentiation of HSPCs was measured by flow cytometry on days 5 and 8 with the cell surface markers corresponding to each respective lineage. Flow cytometry results demonstrated that SCP4-depleted HSPCs showed no defect in differentiation into these three lineages (Fig. 12). The CFC assay revealed that the absolute cell viability declined for each condition after RNP electroporation, including NT sgRNA, as compared to the Mock population (Fig. 13A). However, the HSPC population comprising SCP4 knockout clones gave rise to similar relative ratios of colonies as both control HSPCs populations (Fig. 13B).

These findings are consistent with the survival of SCP4^{-/-} mice to the neonatal stage of development (Cao et al. 2018) since a severe defect in hematopoiesis would have led to the embryonic lethality of these animals. Collectively, these data suggest that SCP4 is dispensable in normal hematopoietic progenitor cells and presents an acquired dependency in AML.

RESULTS

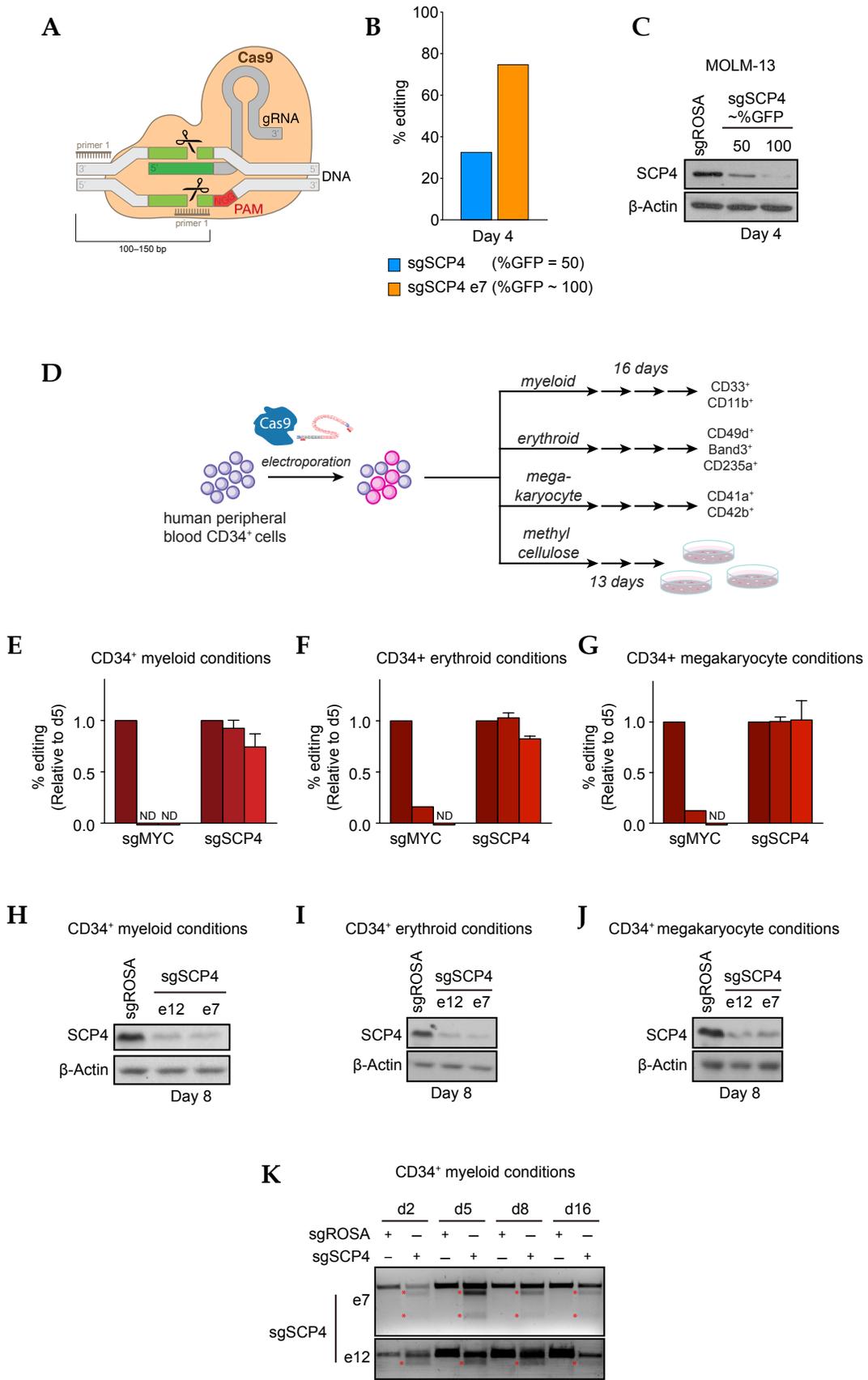


Figure 11. SCP4-depleted HSPCs did not experience any fitness disadvantage during 16 days of culturing under conditions that promote myeloid, erythroid, or megakaryocytic differentiation.

(A) Scheme of RT-qPCR assay design. Ribonuclease complex of CRISPR-Cas9 system showing sgRNA bound to Cas9 protein and PAM sequence downstream from cleavage site. Primers for RT-qPCR sequence are shown. Adapted from Marius Walter, CC BY-SA 4.0, via Wikimedia Commons. **(B)** RT-qPCR analysis of indels presence in MOLM-13 cells infected with sgSCP4 linked to GFP with %GFP measured by GUAVA as 50% (blue) or >90% (orange). **(C)** Western blot of whole-cell lysates from MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. **(D)** Scheme of the editing and differentiation experiments performed in human peripheral blood CD34⁺ cells. The cell surface markers, corresponding to each differentiated lineage, used in the flow cytometry panels in CD34⁺ experiments, are shown. **(E-G)** RT-qPCR analysis of indels presence in CD34⁺ cells electroporated with Cas9 loaded with the indicated sgRNAs over the course of culturing in myeloid (E), erythroid (F), and megakaryocyte (G) conditions. The effects of individual sgRNAs for SCP4 were averaged. n = 4. **(H-J)** Western blot of whole-cell lysates from CD34⁺ cells electroporated with Cas9 loaded with the indicated sgRNAs, day 8 post-electroporation, culturing in myeloid (H), erythroid (I), and megakaryocyte (J) conditions. **(K)** Surveyor Assay analysis of indels presence during in CD34⁺ cells electroporated with Cas9 loaded with the indicated sgRNAs over the course of culturing in myeloid conditions.

All bar graphs represent the mean \pm SEM. 'e' refers to the exon number targeted by each sgRNA. ND, not detected.

Experiment set up together with Ruopeng Feng and Yu Yao.

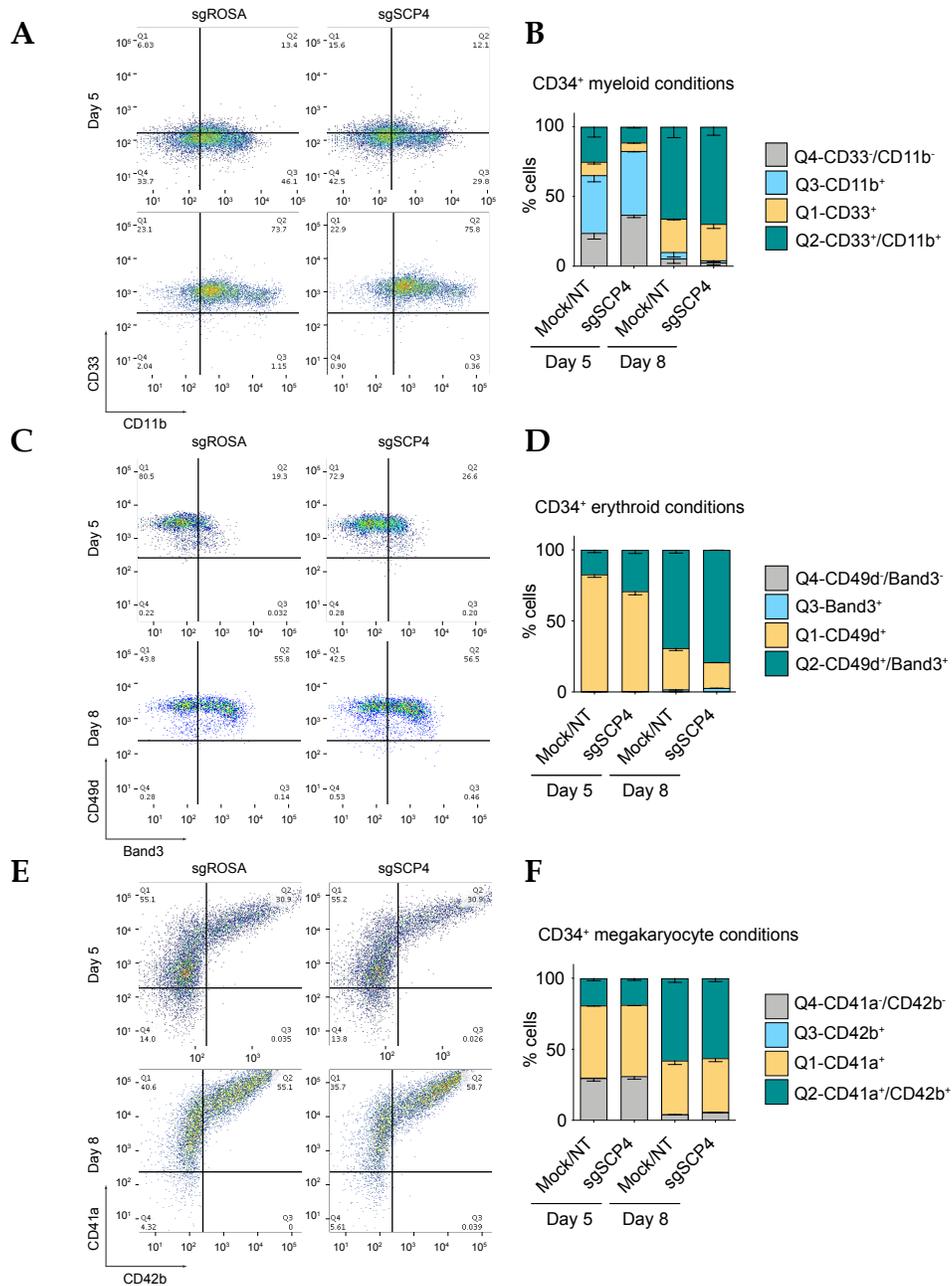


Figure 12. SCP4-depleted HSPCs showed no defect in differentiation into myeloid, erythroid, or megakaryocytic lineages.

(A, C, E) Representative flow cytometry analysis of CD34⁺ differentiation into myeloid (A), erythroid (B), and megakaryocyte (C) lineages. (B, D, F) Quantification of the flow cytometry analysis of myeloid (B), erythroid (D), and megakaryocyte (F) differentiation of CD34⁺ cells electroporated with Cas9 loaded with the indicated sgRNAs. The effects of individual negative controls and sgRNAs for SCP4 were averaged. n = 4.

Data collected with help from Ruopeng Feng and Yu Yao.

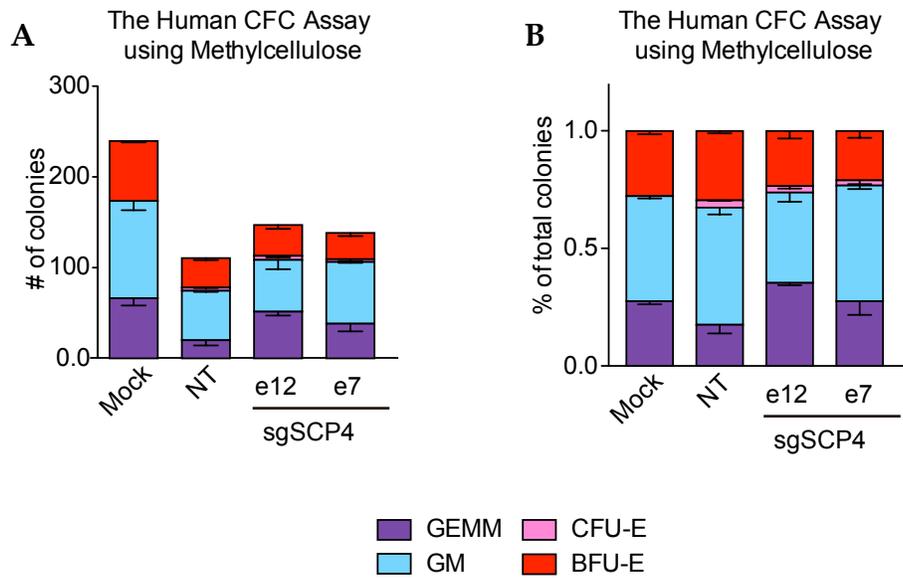


Figure 13. SCP4-depleted HSPCs showed no differences in relative representation of the colony-forming units.

The Human Colony Forming Cell (CFC) assay using methylcellulose. **(A)** Absolute numbers of colonies. **(B)** Relative numbers of colonies. GEMM, colony forming unit-granulocyte, erythrocyte, macrophage, megakaryocyte; GM, colony forming unit-granulocyte, macrophage; CFU-E, colony forming unit-erythroid; BFU-E, burst forming unit-erythroid.

Experiment set up together with Ruopeng Feng and Yu Yao.

2.4. The catalytic phosphatase activity of SCP4 is essential in AML

SCP4 (466 amino acids in *Homo sapiens*) is found in the majority of eukaryotes except for most fungi and *Drosophilids*. Its C-terminus (amino acids 236–466) containing the phosphatase domain is highly conserved among the vertebrates, whereas amino acids 1–236 are less conserved and are predicted to be intrinsically disordered (Fig. 14A,B) (Chang et al. 2018; Meszaros et al. 2018).

Our laboratory has previously used CRISPR exon scanning to map the regions of proteins functionally important for supporting cancer cell survival and proliferation (Shi et al. 2015; Shen et al. 2015). To determine the minimal fragment of SCP4 required in AML, we performed a negative selection CRISPR-Cas9 screening in MOLM-13 cells with a pooled library of 85 sgRNAs tiling along each of the SCP4 exons.

By quantifying sgRNA depletion along the length of the gene, we observed a pattern in which most sgRNAs (32/43, 74.4%) targeting exons encoding amino acids 1–235 did not result in any growth arrest phenotype, whereas most sgRNAs (32/42, 76.2%) targeting the 236–466 segment experienced negative selection (Fig. 14C,D; Appendix 3). These results suggest that indel mutations produce a higher proportion of null alleles when targeting the C-terminal segment of SCP4, which indicates that the phosphatase domain is critical for AML growth (Shi et al. 2015). Notably, SCP4/*CTDSPL2* sgRNAs used in Project Achilles/DepMap genome-wide screens (Meyers et al. 2017; Dempster et al. 2019) exclusively target the 1–235 segment of the protein (Fig. 14D), which explains why an AML dependency was not observed in these data.

To corroborate the CRISPR exon scanning results, we performed cDNA rescue experiments evaluating the ability of different SCP4 truncations and mutations to complement the inactivation of the endogenous SCP4. Similar to the results obtained with the wild-type protein, the SCP4^{236–466} truncation was sufficient to rescue the knockout of the

endogenous SCP4, which verified that the N-terminal half of SCP4 was dispensable in AML (Fig. 15).

SCP4 belongs to the HAD superfamily of phosphatases characterized by the conserved active site sequence DxDx(T/V) (see Introduction §1.5) (Fig. 16A,B). Both aspartates (D293 and D295 in SCP4) are required for the substrate dephosphorylation. To determine whether SCP4 catalytic function was required in AML, we cloned three different catalytic mutants of SCP4: SCP4^{D293A}, SCP4^{D295A}, and SCP4^{D293A/D295A}. Despite similar expression levels to the wild-type protein, all three mutants were unable to support MOLM-13 growth (Fig. 16C,D). Collectively, these genetic experiments suggest that the catalytic phosphatase function of SCP4 is essential in AML.

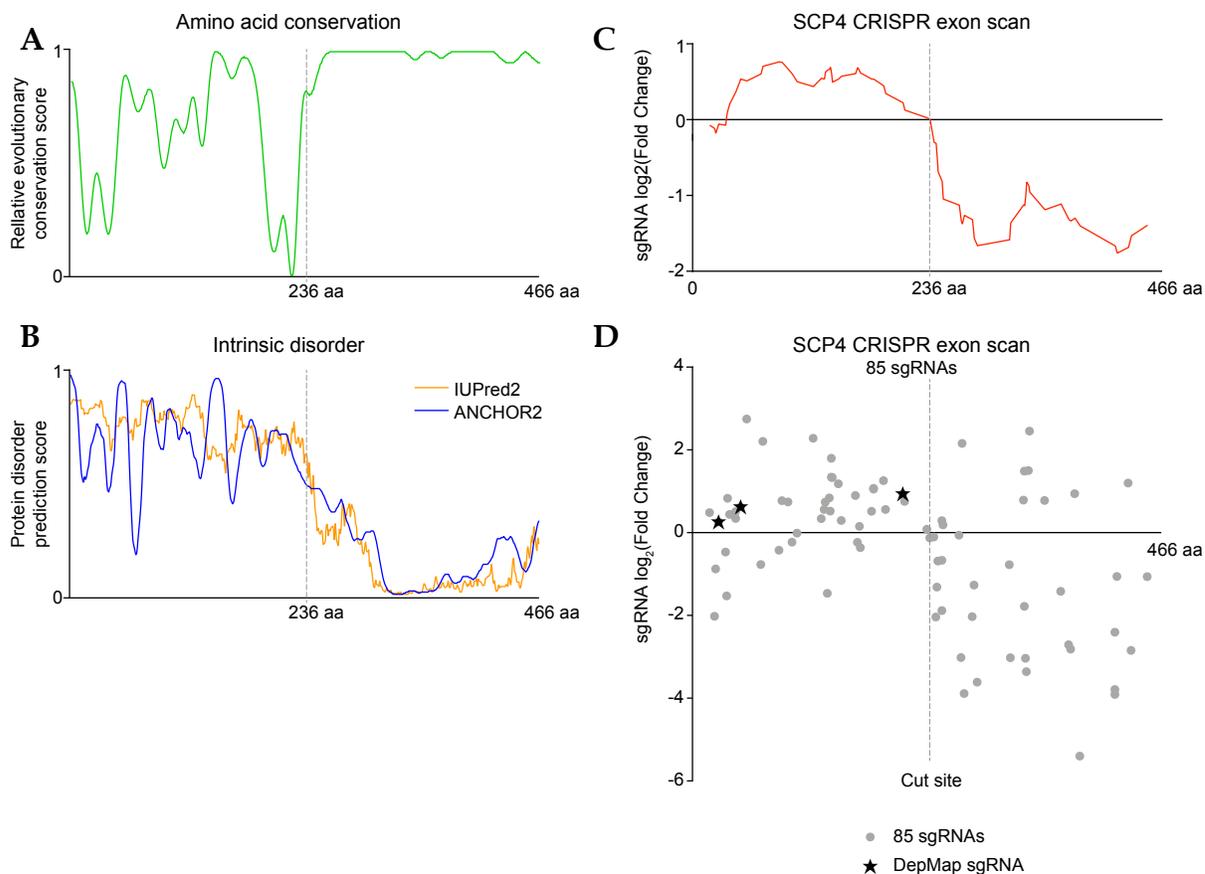


Figure 14. Defining functionally important regions of SCP4.

(A) Relative evolutionary conservation score for each residue, from 0 – least conserved to 1 – most conserved. Based on data from Chang et al. 2018. **(B)** Protein disorder prediction score for each residue, from 0 – least disordered to 1 – most disordered. IUPred2 (orange) predictions of the global structural disorder that encompasses at least 30 consecutive residues of SCP4. ANCHOR2 (blue) probabilities of each residue being part of a binding region. Based on data from Meszaros et al. 2018. **(C)** Running average of log₂ fold changes of the CRISPR-scan of SCP4 with all the possible sgRNAs. SCP4 protein amino acid numbers are indicated along the x axis. **(D)** The CRISPR-scan of SCP4 with all the possible sgRNAs. Deep sequencing-based measurement of the impact of 85 SCP4 sgRNAs on the proliferation of MOLM-13 cells.

Data in (C) and (D) collected by Bin Lu.

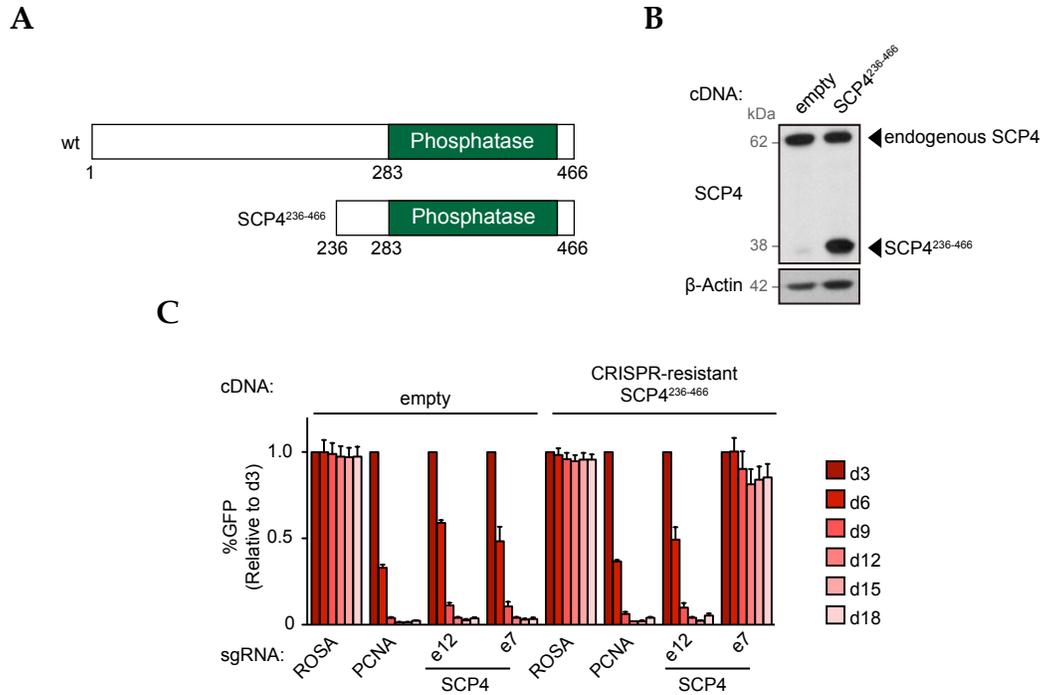


Figure 15. SCP4²³⁶⁻⁴⁶⁶ is sufficient to support AML cell proliferation.

(A) Domain architectures of human SCP4 wild-type and the truncated version SCP4²³⁶⁻⁴⁶⁶ used in this study. (B) Western blot of whole-cell lysates from MOLM-13 cells stably expressing empty vector or CRISPR-resistant SCP4²³⁶⁻⁴⁶⁶. (C) Competition-based proliferation assay in MOLM-13 cells stably expressing empty vector or CRISPR-resistant SCP4²³⁶⁻⁴⁶⁶ infected with the indicated sgRNAs. $n = 3$. All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. 'e' refers to the exon number targeted by each sgRNA. The indicated sgRNAs were linked to GFP. GFP⁺ population depletion indicates loss of cell fitness caused by Cas9/sgrNA-mediated genetic mutations. ROSA, negative control; PCNA, positive control.

2.5. N-terminus of SCP4 promotes nuclear localization of the full-length protein

To narrow down the potential mechanisms by which SCP4 supports AML cell proliferation, we have first decided to confirm its intracellular localization. Consistent with the previous findings (discussed in Introduction §1.6.1), both endogenous and FLAG-tagged SCP4 were only detected in the nuclear fraction (Fig. 17A,B) with the endogenous protein exclusively localized to the chromatin fraction showed by micrococcal nuclease digestion. However, our efforts at performing SCP4 ChIP-Seq in MOLM-13 cells failed to identify evidence of sequence-specific DNA occupancy both with the highly specific monoclonal SCP4 antibody and FLAG antibody against the ectopically expressed N-terminally tagged protein (data not shown).

Therefore, we turned to biochemical approaches to further understand the mechanism of SCP4 dependency. We decided to focus on the protein-protein interactions of SCP4 in its native cellular context in the nucleus with the hope that this approach could provide us with some insights into its biological functions. While trouble-shooting the scale-up fractionation protocol, we noticed that SCP4 was degraded upon the use of any detergents during fractionation, despite supplementing all the buffers with the Protease Inhibitor Cocktail. This degradation was specific to AML cell line MOLM-13 fractions and, in contrast, was not observed in a pancreas cancer cell line SUIT-2 fractions under identical conditions (Fig. 18A). This reminded us of another publication that reported a similar degradation pattern for HOXA9 in MOLM-13 cells and attributed it to the action of myeloid cells specific proteases neutrophil elastase (*ELANE*), proteinase 3 (*PRTN3*), and cathepsin G (*CTSG*) (Zhong et al. 2018). Our collaborators from University of Texas performed GST pulldown assay from MOLM-13 cytoplasmic fraction using a GST-tagged SCP4¹⁵⁶⁻⁴⁶⁶ as bait and indeed detected all three myeloid proteases as its partners (Fig. 18B). On the other hand, FLAG-SCP4²³⁶⁻⁴⁶⁶ was present in the cytoplasm, suggesting that the protease recognition

sites could be located somewhere at the N-terminus (Fig. 17C). Therefore, the nuclear localization of the full-length SCP4 in myeloid cells may be partially supported by its compartmentalized degradation. Thus, we optimized a fractionation protocol without the use of detergents and with stringent nuclei washes from cytoplasmic contamination that efficiently preserved full-length SCP4 (Fig. 18C).

Our FLAG immunoprecipitation (FLAG-IP) procedure efficiently recovered FLAG-SCP4 complexes and removed endogenous SCP4 (Fig. 19A). For the immunoprecipitation coupled with mass spectrometry (IP-MS) experiments, we compared wild-type SCP4 to the catalytic mutant SCP4^{D293A}, with the intent that any protein-protein interaction abrogated by this loss-of-function mutation would have a higher likelihood of functional significance.

A comparison of two independent IP-MS biological replicates revealed 59 protein candidates (Fig. 19B) shared between FLAG-SCP4 samples and absent in negative controls. Among the most abundant SCP4-associated proteins, based on total spectral counts, were several members of the importin protein family, which function to shuttle protein cargo into the nucleus (Cagatay and Chook 2018) (Fig. 19C). To discriminate between the functional role of importins in SCP4 nuclear localization and other hypotheses, we repeated FLAG-IP for all the FLAG-SCP4 constructs with the addition of FLAG-SCP4²³⁶⁻⁴⁶⁶. The latter protein is smaller than 40 kDa and thus would be expected to bypass the requirement for importins in its nuclear import (Weis 2003), but can complement the knockout of endogenous SCP4 in supporting AML cell proliferation (Fig. 15; Fig. 17C). We have shown that all the full-length SCP4 constructs interact with importin α , KPNA2, unlike FLAG-SCP4²³⁶⁻⁴⁶⁶ (Fig. 20A). This is coherent with the nuclear localization signal (NLS) predictions that show multiple potential NLSs for SCP4 among its first 100 amino acids (Fig. 20B).

Our results suggest that at least one function of the SCP4 N-terminal segment is to promote nuclear localization of the full-length protein (Fig. 20C).

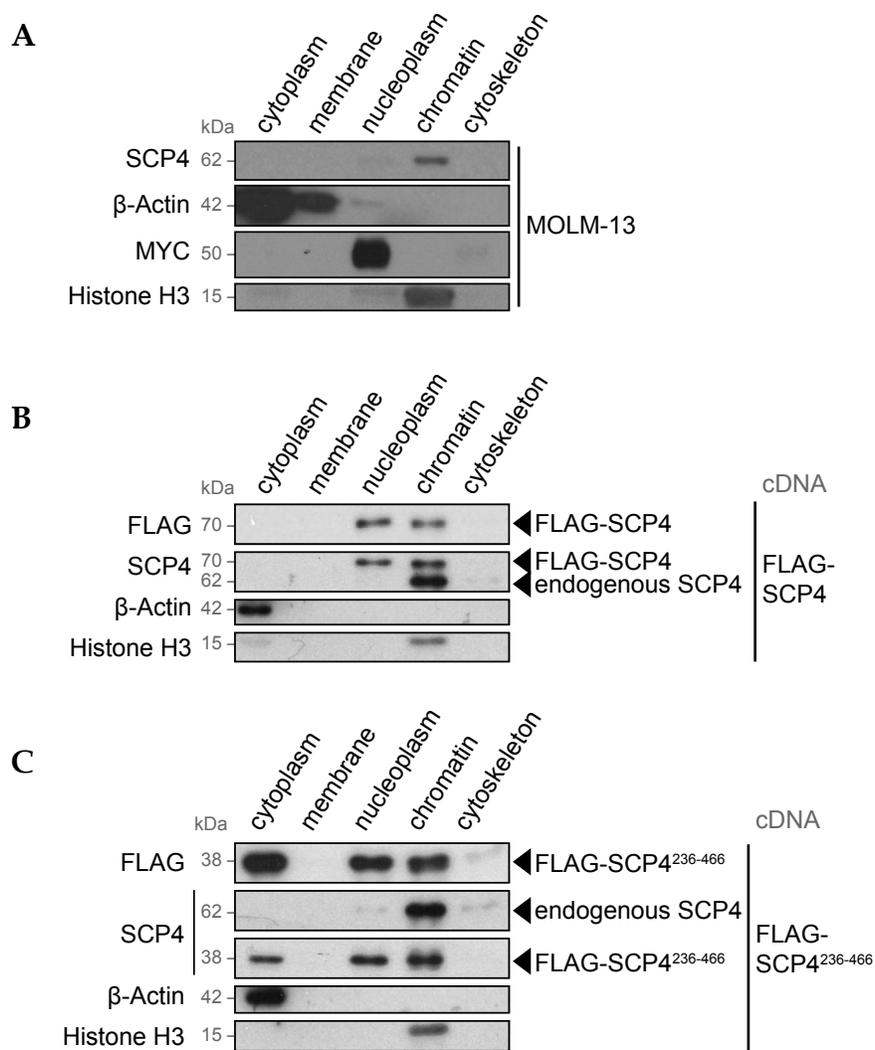


Figure 17. SCP4 is localized to the nucleus.

Western blots of the indicated fractions from (A) MOLM-13 cells and MOLM-13 cells stably expressing (B) FLAG-SCP4^{wt} and (C) FLAG-SCP4²³⁶⁻⁴⁶⁶.

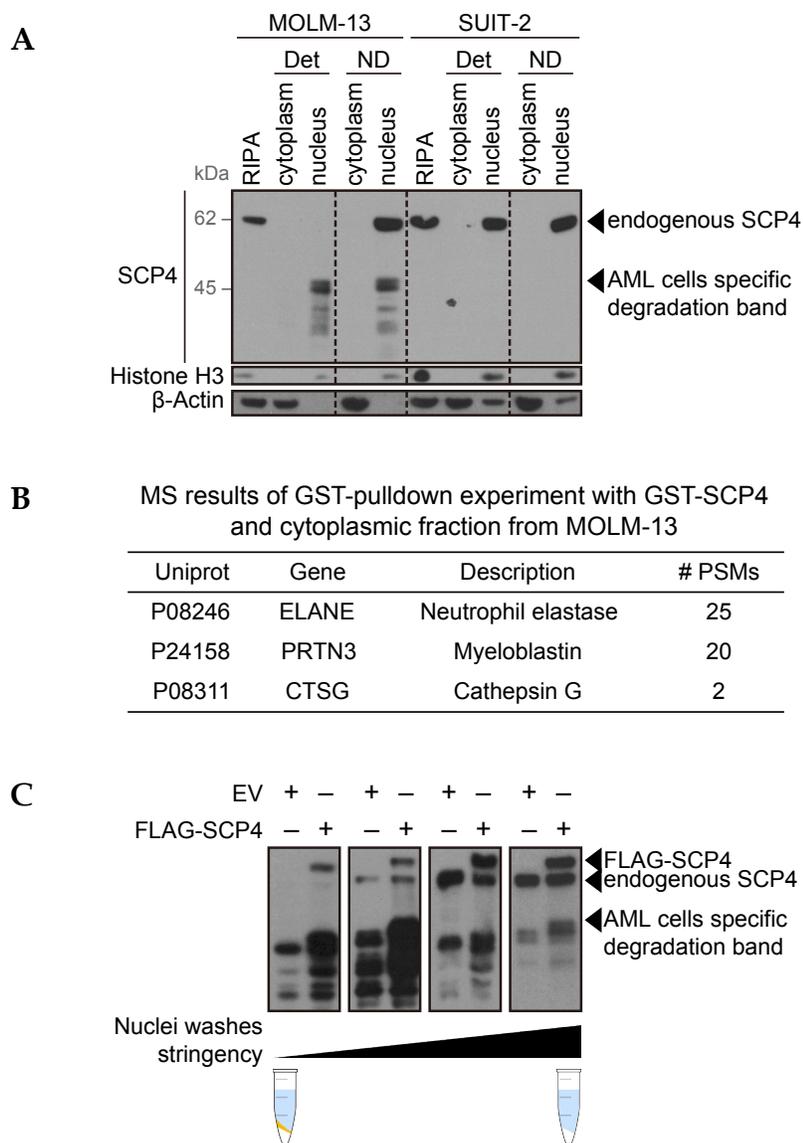


Figure 18. Optimization of SCP4 recovery from nuclear fractions.

(A) SCP4 is unstable in AML cells but not in pancreatic cancer cells. Western blots of the whole-cell lysates (RIPA) and cytoplasm/nucleus fractions from the indicated cell lines separated in parallel using the same buffers containing detergents (Det) or not (ND). Vertical black dashed lines indicate omitted lanes from the same gel, same exposure. (B) Myeloid proteases detected by MS analysis in GST-SCP4 pulldown from MOLM-13 cytoplasmic fraction. # PSMs, the total number of identified peptide spectrum matches for the protein. (C) Representative nucleus fractions from MOLM-13 cells stably expressing empty vector (EV) or FLAG-SCP4 separated using buffers not containing any detergents with increased nuclei washes (white pellet in the schematic) from cytoplasmic contamination (shown in yellow in the schematic).

Data in (B) collected by Seema Irani, Yan Jessie Zhang.

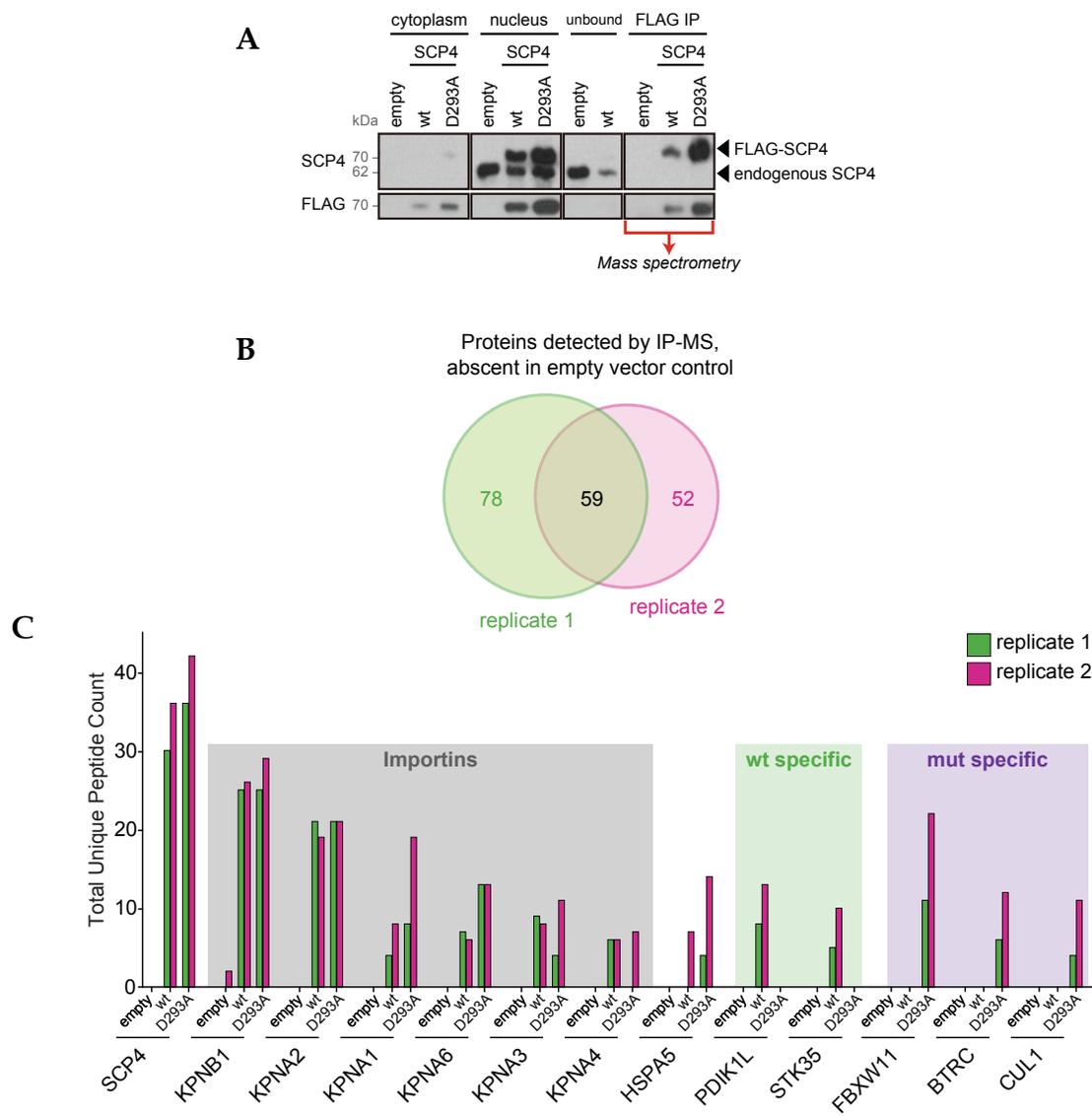


Figure 19. IP-MS identification of FLAG-SCP4 associated proteins.

(A) Representative FLAG-SCP4 immunoprecipitation Western blot analysis for the subsequent mass spectrometry (MS) analysis. Cytoplasm and nucleus fractions from MOLM-13 cells stably expressing empty vector, FLAG-SCP4 wild-type (wt), or catalytic mutant (D293A). Nuclear fraction was mixed with anti-FLAG M2 agarose overnight. The flow-through was analyzed to ensure efficient binding of the FLAG-tagged constructs (loaded as “unbound”). Following the extensive washing, the agarose amount equivalent to the cytoplasm and nucleus fractions loading was boiled in Laemmli buffer and loaded as “FLAG-IP.” The rest was sent for the MS analysis. (B) Venn diagram depicting the overlap between proteins detected by MS and absent in empty vector controls in the two independent biological replicates. (C) Total unique peptide counts for the top hits detected by MS in two independent biological replicates.

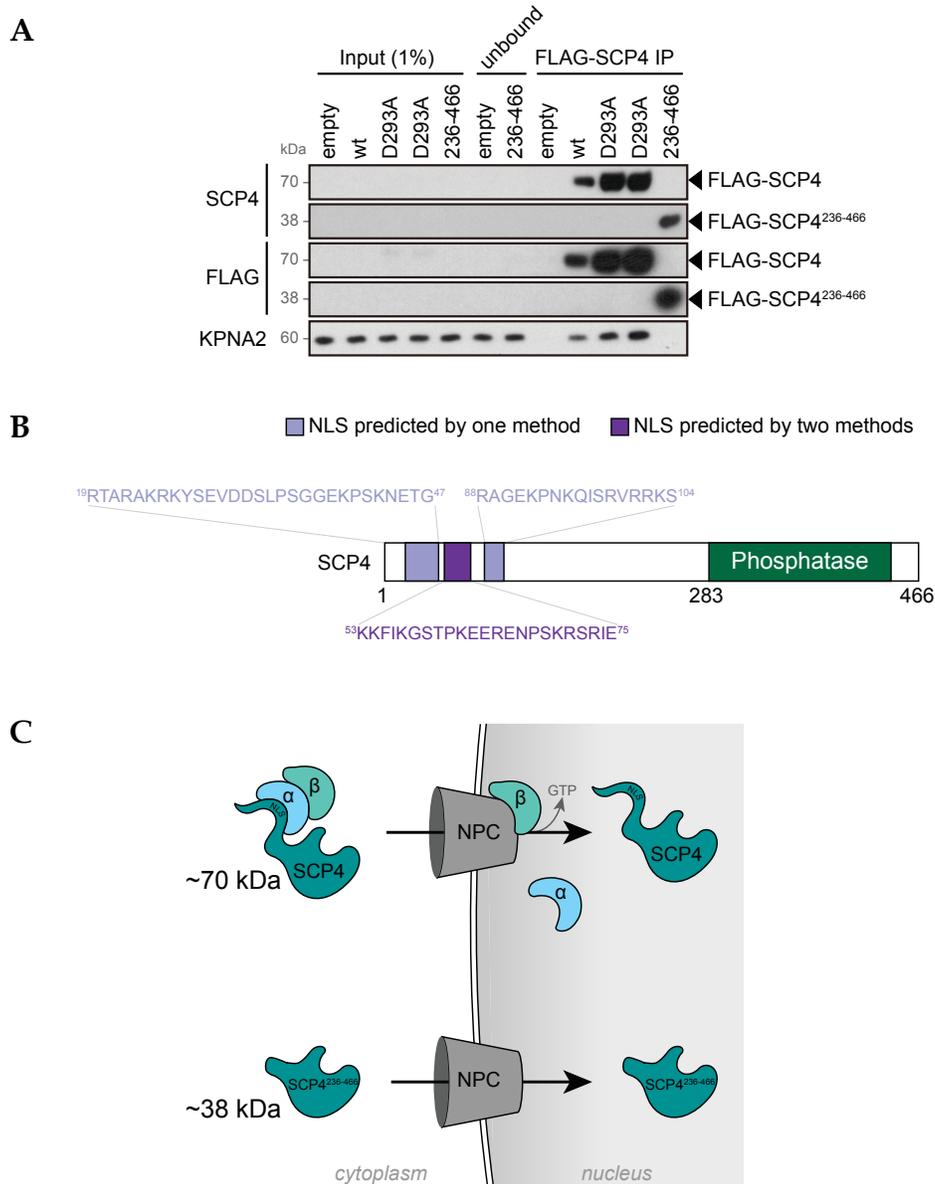


Figure 20. SCP4 N-terminus promotes nuclear localization of the full-length protein.

(A) Immunoprecipitation followed by Western blot performed with the indicated antibodies. The nuclear lysates were prepared from the human MOLM-13 cells stably expressing empty vector (empty), FLAG-SCP4^{wt} (wt), FLAG-SCP4²³⁶⁻⁴⁶⁶ (236-466), catalytic mutant FLAG-SCP4^{D293A} (D293A). The flow-through was analyzed to ensure efficient binding of the FLAG-tagged constructs (loaded as “unbound”). (B) Nuclear localization signals (NLS) predictions for SCP4 amino acid sequence performed by two different methods (Nguyen Ba et al. 2009; Kosugi et al. 2009). (C) Model of full length and truncated SCP4 transport into the nucleus.

2.6. STK35 and PDIK1L bind selectively to the catalytically active form of SCP4

Among the SCP4-associated proteins identified in our FLAG-IP-MS analysis, the paralog kinases STK35 and PDIK1L were remarkable as they were only detected in association with the wild-type SCP4 but not with the catalytically dead SCP4^{D293A} (Fig. 19C). Notably, PDIK1L and STK35 are paralogs sharing 68.9% identity and 83.3% similarity (Fig. 21A). They constitute a distinct new kinase family 4 (NKF4) and are relatively poorly studied (see Introduction §1.6.2).

As no commercially available antibodies were able to detect these two kinases (data not shown), we cloned HA-tagged proteins for their further biochemical characterization.

To confirm the association between SCP4 and PDIK1L, we first transiently expressed HA-tagged PDIK1L and FLAG-SCP4 constructs in HEK293T cells, followed by FLAG- or HA-IP. In accordance with the MS results, both SCP4^{wt} and SCP4²³⁶⁻⁴⁶⁶ proteins showed similar interaction with HA-PDIK1L (Fig. 21B,C). The interaction was mitigated by the single amino acid substitution yielding catalytically inactive protein SCP4^{D293A}.

We then established MOLM-13 cells constitutively expressing HA-PDIK1L or HA-STK35 together with empty vector or either one of the FLAG-SCP4 constructs (Fig. 21D–G). Similarly, we found that SCP4^{wt} and SCP4²³⁶⁻⁴⁶⁶ each efficiently associated with HA-STK35 and HA-PDIK1L, whereas this interaction was diminished with the SCP4^{D293A} protein.

The interaction between the endogenous SCP4 and HA-PDIK1L or HA-STK35 was detected in IP-MS experiments after stringent washes with salt concentrations ranging from 150 mM up to 500 mM NaCl (Fig. 22A,B). Moreover, an MS analysis of HA-PDIK1L complexes from MOLM-13 cells following high salt washes identified endogenous SCP4 as the most enriched protein with comparable spectral counts to the PDIK1L bait (Fig. 22C).

These biochemical experiments suggest that the kinases STK35 and PDIK1L exist in a stable complex with the catalytically active SCP4 phosphatase domain.

RESULTS

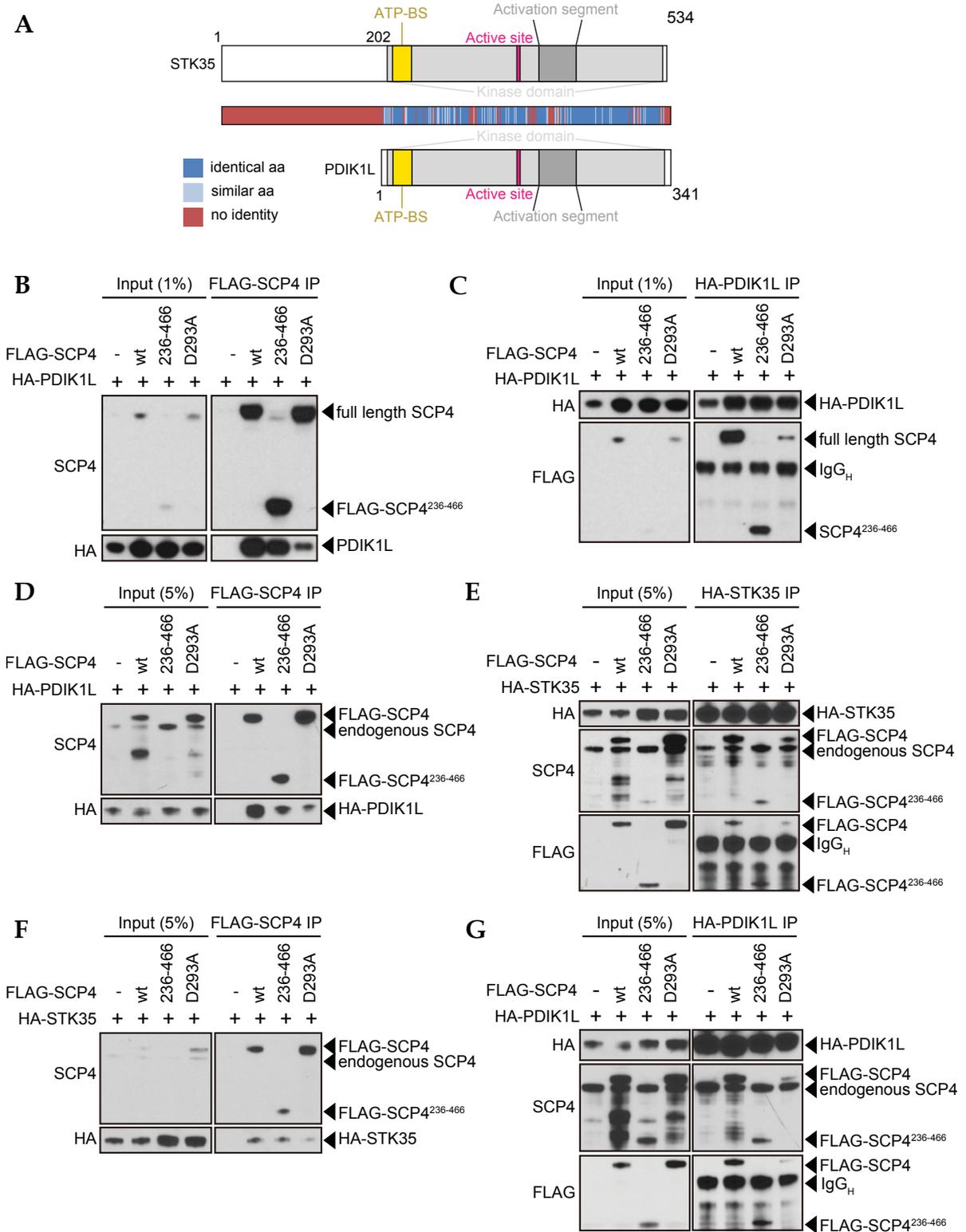


Figure 21. STK35 and PDIK1L bind selectively to the catalytically active form of SCP4.

(A) Domain architectures and homology heat-map of human STK35 and PDIK1L. ATP-BS, ATP-binding site. **(B,C)** Immunoprecipitation followed by Western blot performed with the indicated antibodies. The whole-cell lysate was prepared from HEK293T 24 hours post-transfection with the indicated constructs. **(D-G)** Immunoprecipitation followed by Western blot performed with the indicated antibodies. The nuclear lysates were prepared from the human MOLM-13 cells stably expressing the indicated constructs.

'-', empty vector; wt, wild-type FLAG-SCP4; 236-466, FLAG-SCP4²³⁶⁻⁴⁶⁶; D293A, catalytic mutant FLAG-SCP4^{D293A}; IP, immunoprecipitation. Note: degradation bands appear in the wt and D293A input at ~50 kDa and at ~40 kDa (E,G) and the wt IP at ~40 kDa (G).

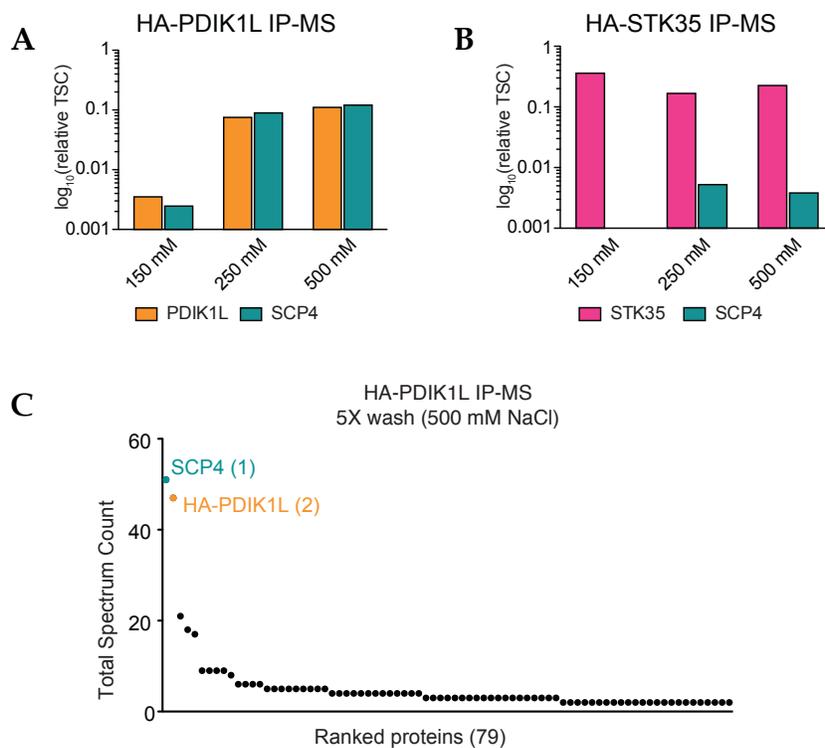


Figure 22. STK35 and PDIK1L exist in a stable complex with the endogenous SCP4.

(A,B) Total spectrum counts (TSC) for HA-PDIK1L (A), HA-STK35 (B), and endogenous SCP4 relative to the sum of total spectrum counts of all proteins detected by mass spectrometry on Pierce Anti-HA Magnetic Beads after stringent washes with washing buffers with the indicated NaCl salt concentrations. (C) Total spectrum counts (TSC) for HA-PDIK1L and endogenous SCP4 detected by mass spectrometry on Pierce Anti-HA Magnetic Beads after stringent washes with washing buffers containing 500 mM NaCl.

2.7. STK35 and PDIK1L function redundantly in the same genetic pathway as SCP4

The kinase domain-focused CRISPR screening in human cancer cell lines previously performed in our laboratory did not reveal a robust dependency on STK35 or PDIK1L in AML cell lines (Tarumoto et al. 2018). Since both kinases are expressed in AML (Fig. 23A), we hypothesized that their functional redundancy might conceal their essentiality in this context. To check this hypothesis, we first transduced AML cell lines MOLM-13 and THP1, and a CML cell line K562 with STK35 and PDIK1L sgRNAs linked to mCherry and GFP respectively in combination with each other or with an opposite color fluorophore linked to negative control and performed dual targeting competition-based proliferation assays (Fig. 23B). These experiments revealed that MOLM-13 cells experienced a synergistic proliferation defect after the dual targeting of STK35 and PDIK1L, but the cells were less sensitive to targeting either gene individually (Fig. 23C). In contrast, THP1 and K562 cells did not show any phenotype following co-transduction with STK35 and PDIK1L sgRNAs (Fig. 23C).

We then switched to using a bi-cistronic vector developed in our laboratory to express two sgRNAs simultaneously upon lentiviral infection (Klingbeil, unpublished). Using this system, we confirmed that the deleterious phenotype of STK35/PDIK1L double knockout in MOLM-13 cells could be fully rescued by expressing a single CRISPR-resistant STK35 or PDIK1L cDNA (Fig. 24A-C). Western blot experiments verified that the ectopically expressed HA-tagged wild-type STK35 and PDIK1L were effectively depleted upon their respective sgRNA and bi-cistronic vector introduction, whereas CRISPR-resistant versions were unaffected (Fig. 24D,E). Cell cycle analysis with the bi-cistronic vector showed that STK35/PDIK1L double knockout in MOLM-13 cells led to a G1/G0-arrest on day 5 post-infection, similar to SCP4 knockout (Fig. 25; Fig. 8).

We have then assayed the requirement for SCP4 and the pair of kinases in a series of

competition-based proliferation assays and confirmed that in this limited panel of cell lines, the dependency for SCP4 correlated with the growth arrest phenotype upon dual targeting of STK35 and PDIK1L (Fig. 26). For example, SCP4 and STK35/PDIK1L were not required in THP1, U937, and K562 cells compared to the four AML cell lines sensitive to both perturbations. These experiments demonstrate that STK35 and PDIK1L have redundant functions in AML and suggest that the requirement for these kinases is linked to the SCP4 dependency in this context.

We next investigated whether the catalytic kinase function of STK35 and PDIK1L was essential for AML cell proliferation. We have first performed multiple sequence alignment of STK35 and PDIK1L with other well-described human kinases and identified the conserved ATP-binding lysine residue present in other catalytically active kinases (Fig. 27A) (Cherry and Williams 2004; Lucet et al. 2006; Madeira et al. 2019). Guided by prior studies (Kamps and Sefton 1986), we replaced this residue with histidine in the CRISPR-resistant STK35 and PDIK1L cDNA constructs and performed competition-based proliferation assays evaluating whether these mutant cDNAs could rescue the lethal phenotype arising from the dual targeting of STK35/PDIK1L. Despite being expressed, albeit slightly below the level of wild-type proteins, the STK35^{K231H} and PDIK1L^{K37H} behaved as null alleles in these assays, suggesting that the presence of at least one catalytically active kinase in AML cells is necessary for leukemia cell proliferation (Fig. 27B,C).

To further investigate whether the SCP4 and STK35/PDIK1L co-dependency in AML reflected that these proteins functioned in a common genetic pathway, we performed an early time point RNA-Seq following acute genetic knockouts of SCP4 or STK35/PDIK1L. Polyadenylated RNA for the deep sequencing analysis was collected on day 5 following lentiviral infection of MOLM-13 cells with SCP4 or dual STK35/PDIK1L sgRNAs (dgSTK35-

PDIK1L). We compared the global transcriptional changes upon these perturbations across 15,095 genes to the datasets previously collected in our laboratory in the MOLM-13 cell line for the studies elucidating functional links between other proteins acting together to support leukemia cells proliferation, such as MYB and its coactivator TAF12 (Xu et al. 2018) or SIK3 and MEF2C (Tarumoto et al. 2018; see Introduction §1.2.2.6) (Fig. 28). Global correlation analysis of these datasets in all pairwise combinations recapitulated the TAF12-MYB (Pearson $R^2 = 0.56$) and SIK3-MEF2C (Pearson $R^2 = 0.35$) connectivity described previously. Notably, the transcriptional changes observed in SCP4- and STK35/PDIK1L-deficient MOLM-13 cells were closely correlated with one another compared to these other knockouts (Pearson $R^2 = 0.63$). Moreover, our data illustrate that the transcriptional changes following SCP4 knockout and STK35/PDIK1L double knockout were not non-specific effects of targeting essential genes but rather an indication that SCP4 and STK35/PDIK1L operated in the same genetic pathway.

As an orthogonal approach, we defined SCP4 signatures as the top 200 genes downregulated (sgSCP4_dn_MOLM13) and upregulated (sgSCP4_up_MOLM13) in MOLM-13 cells following SCP4 knockout (Appendix 4 & 5). The effect of STK35/PDIK1L double knockout on SCP4 target genes was more substantial than on any of the other 20,329 signatures from our laboratory and the Molecular Signatures Database (MSigDB v7.4) when interrogated via gene set enrichment analysis (GSEA) (Subramanian et al. 2005) (Fig. 29).

The correlated transcriptional changes and the co-dependency on SCP4 and STK35/PDIK1L across AML cell lines suggest that the physical interaction between SCP4 and STK35/PDIK1L reflects the functioning of these signaling molecules in a common genetic pathway.

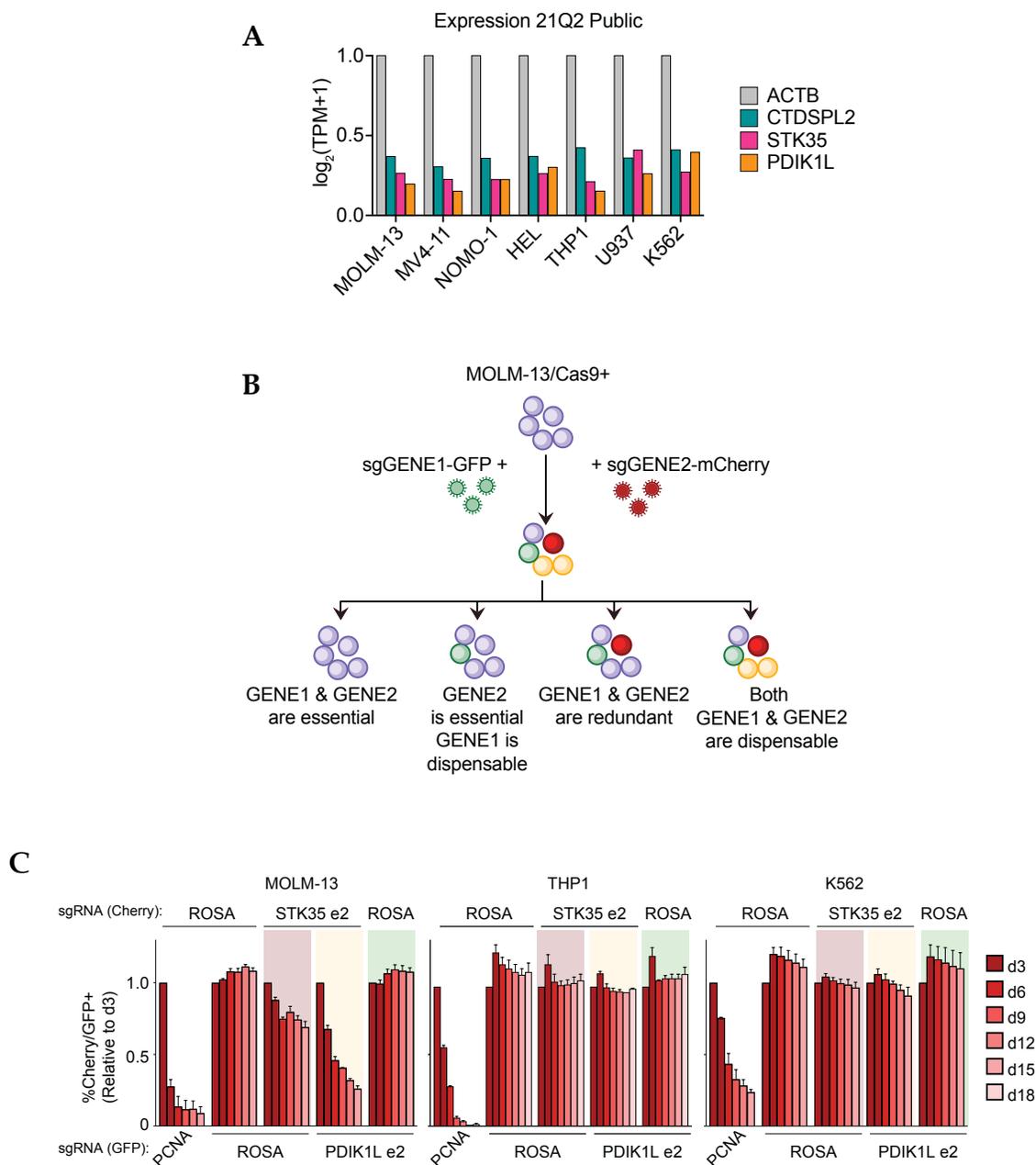


Figure 23. STK35 and PDIK1L function redundantly to support MOLM-13 proliferation.

(A) SCP4, STK35, and PDIK1L mRNA expression data relative to beta-actin expression (Meyers et al. 2017; Dempster et al. 2019). (B) Scheme of a dual targeting competition-based proliferation assay and interpretations of its possible outcomes. (C) Competition-based proliferation assays in MOLM-13, THP1, and K562 cells co-infected with GFP-linked sgRNA and mCherry-linked sgRNA. Double mCherry⁺/GFP⁺ population depletion indicates loss of cell fitness due to genetic redundancy. $n = 3$. All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. 'e' refers to the exon number targeted by each sgRNA.

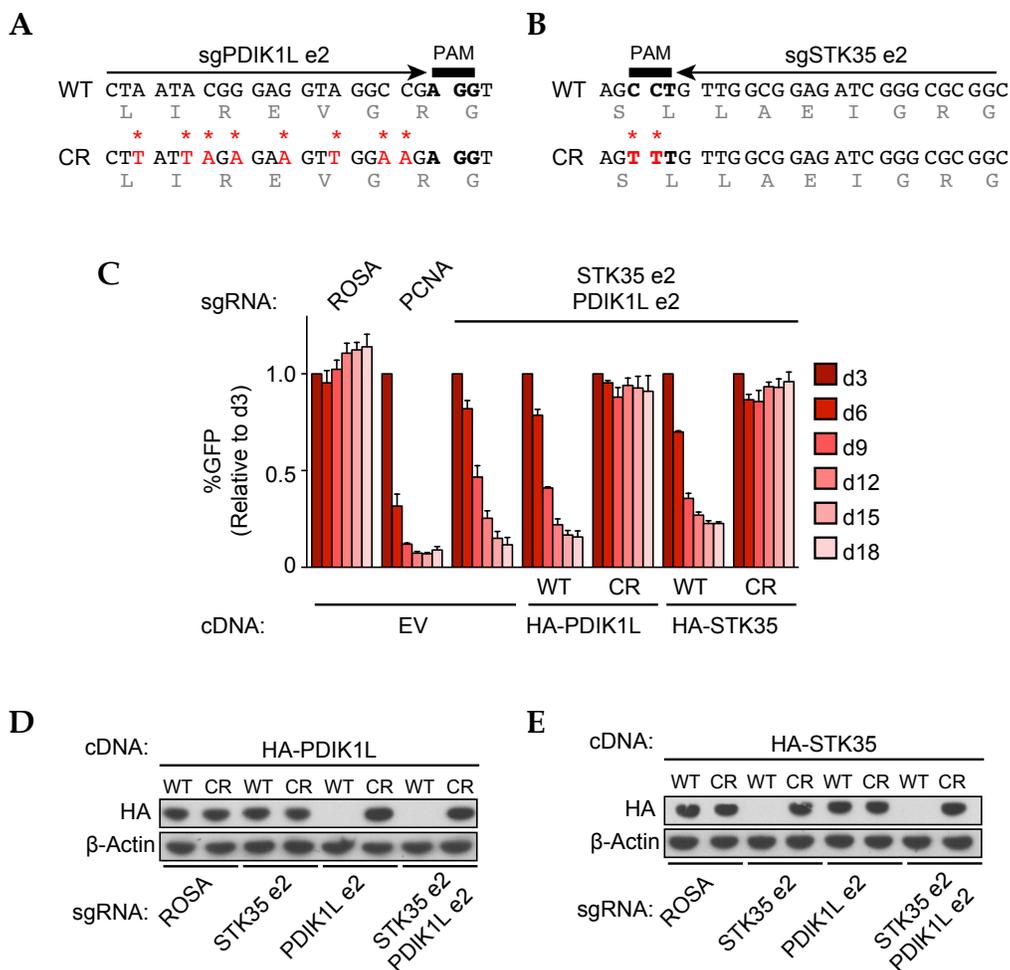


Figure 24. Either STK35 or PDIK1L is sufficient to fully rescue the endogenous STK35/PDIK1L double knockout phenotype in MOLM-13 cells.

(A,B) Design of CRISPR-resistant mutant of (A) PDIK1L (A) and (B) STK35. (C) Competition-based proliferation assay in MOLM-13 cells stably expressing empty vector (EV), wild-type (WT), or CRISPR-resistant (CR) HA-PDIK1L and HA-STK35 infected with the indicated sgRNAs. The indicated sgRNAs were linked to GFP. $n = 3$. (D,E) Western blot of whole-cell lysates from MOLM-13 cells stably expressing empty vector (EV) or CRISPR-resistant wild-type (WT) and catalytic mutants of (D) PDIK1L and (E) STK35.

All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. ‘e’ refers to the exon number targeted by each sgRNA. Starting from Figure 5C and for the rest of the study bi-cistronic vector for simultaneous targeting of STK35 and PDIK1L was used. ‘dg’ refers to the bi-cistronic vector. ROSA, negative control; PCNA, positive control.

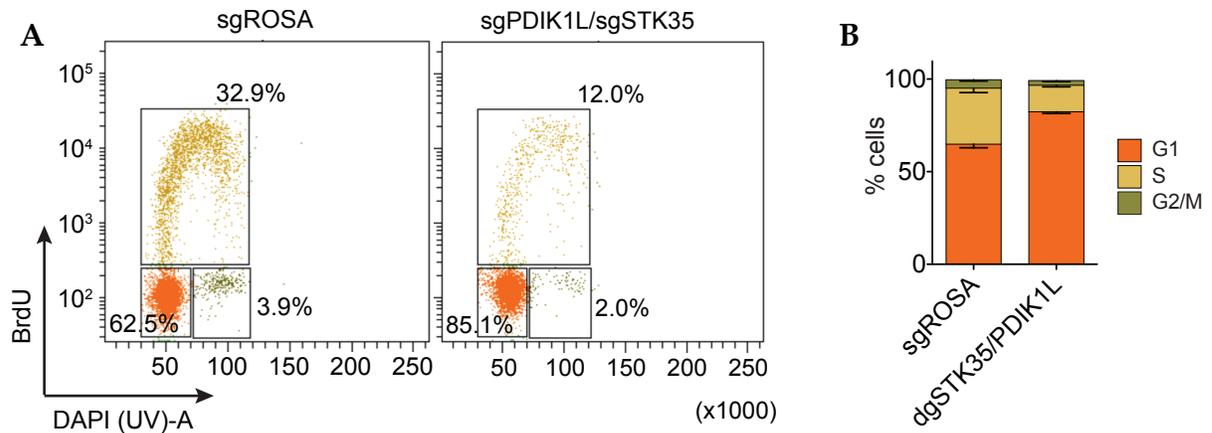


Figure 25. STK35/PDIK1L double knockout in MOLM-13 cells leads to a G1/G0-arrest.

(A) Representative flow cytometry analysis of BrdU incorporation and DNA content to infer cell cycle stage in MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. **(B)** Quantification of the different cell cycle stages in MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. $n = 3$. All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. ROSA, negative control.

Data collected by Zhaolin Yang.

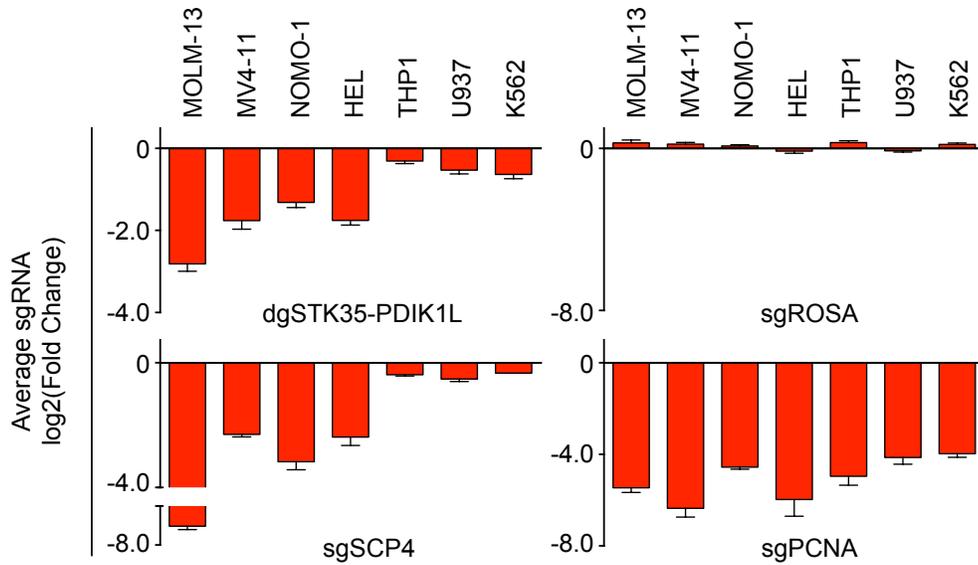


Figure 26. STK35/PDIK1L requirement is correlated with the SCP4 dependency in the context of AML.

Summary of competition-based proliferation assays in the indicated cell lines. Plotted is the fold change (\log_2) of sgRNA⁺/GFP⁺ cells after 18 days in culture (average of triplicates).

All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. The indicated sgRNAs were linked to GFP. GFP⁺ population depletion indicates loss of cell fitness caused by Cas9/sgRNA-mediated genetic mutations. 'dg' refers to the bi-cistronic vector. ROSA, negative control; PCNA, positive control.

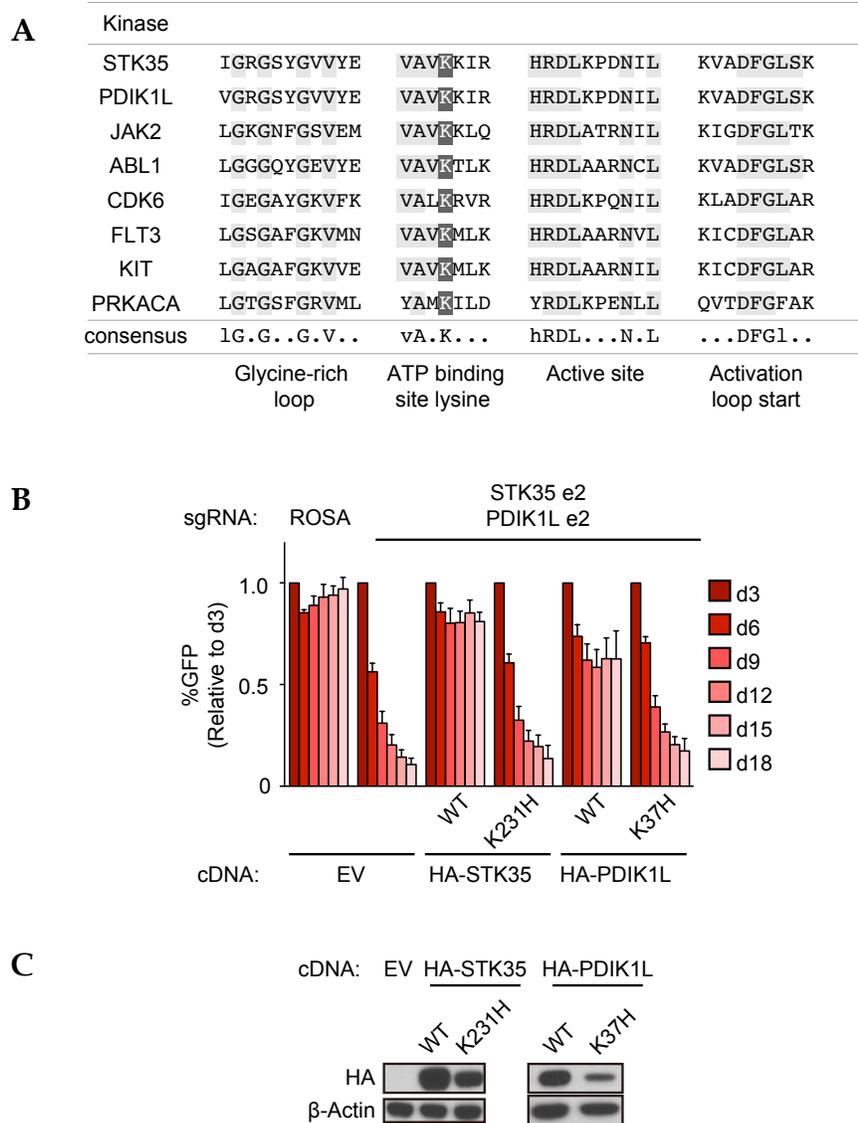


Figure 27. The kinase catalytic function of STK35/PDIK1L is essential in AML.

(A) Alignment of residues defining ATP binding site for human protein kinases. Residues that are conserved among the kinases sequences are highlighted in light gray. Dark gray indicates the conserved lysine at the ATP binding site. Consensus sequence given; uppercase = conserved, lowercase = nearly invariant. The predicted structural elements of protein kinases are indicated below the consensus sequence. (B) Competition-based proliferation assay in MOLM-13 cells stably expressing empty vector (EV) or CRISPR-resistant wild-type (WT) and catalytic mutant versions of HA-STK35 and HA-PDIK1L infected with the indicated sgRNAs. The indicated sgRNAs were linked to GFP. $n = 3$. (C) Western blot of whole-cell lysates from MOLM-13 cells stably expressing empty vector (EV) or CRISPR-resistant wild-type (WT) and catalytic mutants of STK35 and PDIK1L in MOLM-13.

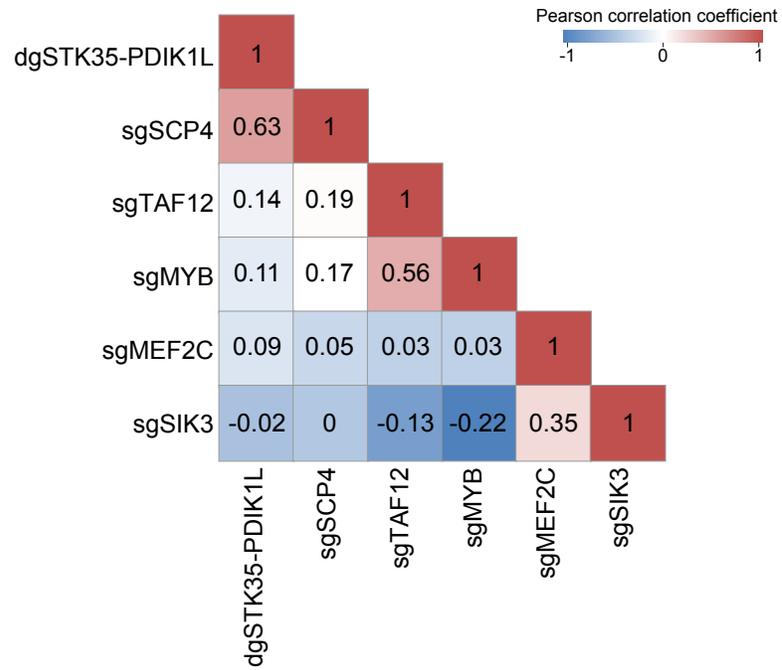


Figure 28. STK35 and PDIK1L function redundantly in the same genetic pathway as SCP4.

Correlation matrix for the global log₂ fold-changes of gene expression relative to negative controls in each independent experiment across 15,095 genes.

Data analyzed with the help from Olaf Klingbeil.

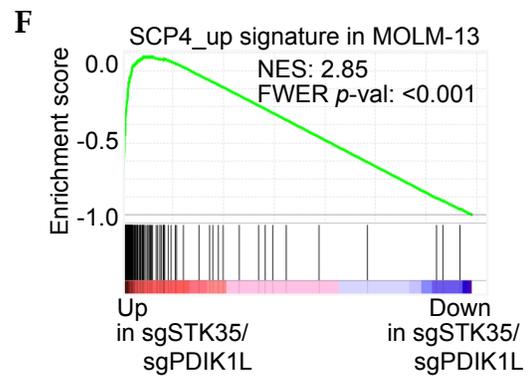
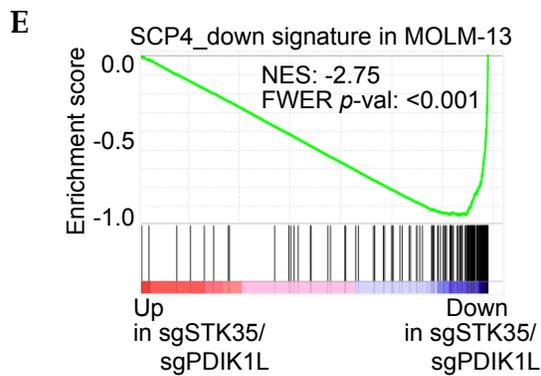
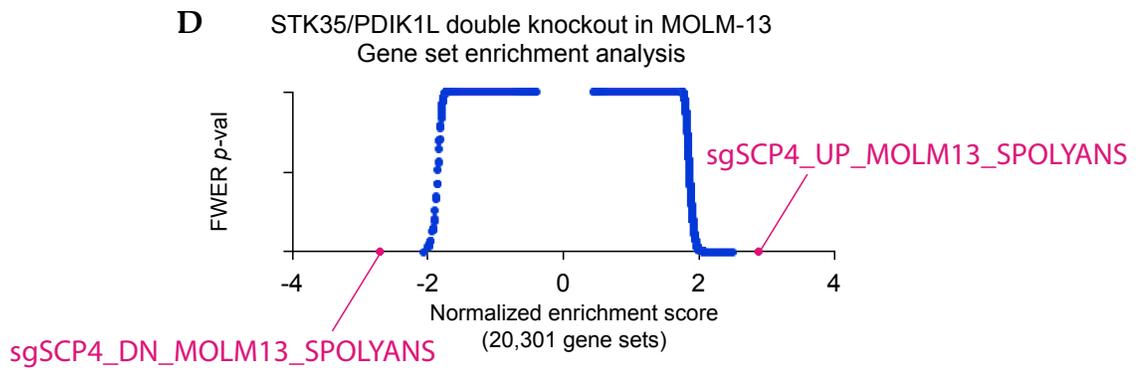
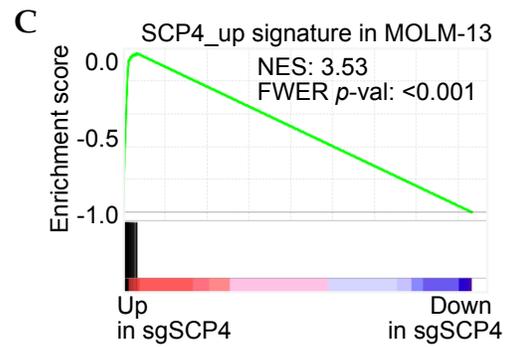
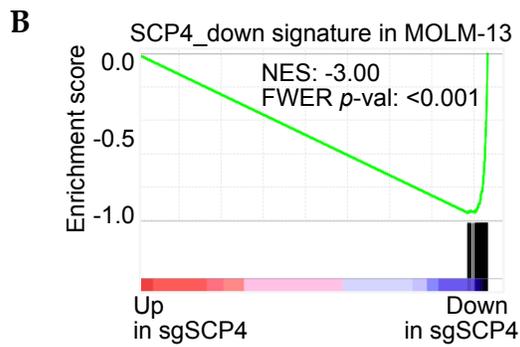
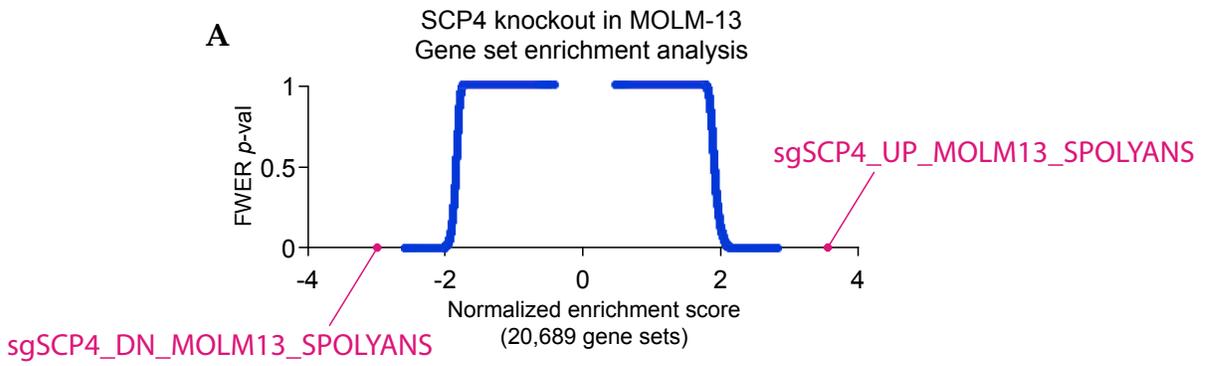


Figure 29. STK35/PDIK1L dual perturbation significantly affects SCP4 signatures.

(A) Definition of SCP4 signatures. Gene set enrichment analysis (GSEA) for SCP4 knockout (shown as a proof of principle) using all the signatures from our laboratory and all the signatures in the Molecular Signatures Database (MSigDB v7.4) together with the SCP4 signatures. Normalized enrichment score (NES) and family wise error rate (FWER) p-values are plotted for 20,689 gene sets, each represented as a single dot. **(B,C)** GSEA plots of the SCP4_down (B) and SCP4_up (C) signatures upon SCP4 knockout. NES and FWER p-values are shown. **(D)** Unbiased GSEA for dual STK35/PDIK1L knockout using all the signatures from our laboratory and all the signatures in the Molecular Signatures Database (MSigDB v7.4) together with the SCP4 signatures. NES and FWER p-values are plotted for 20,301 gene sets, each represented as a single dot. **(E,F)** GSEA plots of the SCP4_down (E) and SCP4_up (F) signatures upon SCP4 knockout. NES and FWER p-values are shown.

2.8. Biochemical evidence that SCP4 functions upstream of STK35/PDIK1L

In order to gain mechanistic insights into the mechanism that links SCP4 to STK35/PDIK1L, we have first decided to measure the protein levels of PDIK1L and STK35 upon SCP4 knockout. The caveat of this experiment was that we had to use HA-tagged ectopically expressed proteins. However, this approach proved to be informative since upon the endogenous SCP4 knockout we have observed a marked reduction in the levels of both HA-PDIK1L (Fig. 30B) and HA-STK35 (Fig. 30D) in the nuclear fractions, which effectively excluded the possibility that SCP4 affected transcriptional regulation of PDIK1L and STK35. Besides, RNA-Seq data also did not reveal significant changes in the endogenous PDIK1L and STK35 levels upon SCP4 knockout (Fig. 30A,C). Neither gene or protein levels of SCP4 were affected by the double knockout of PDIK1L and STK35 (Fig. 30E,F). The directionality of this result suggests that SCP4 functions upstream of STK35/PDIK1L to maintain the protein level of these kinases.

We next hypothesized that the physical interaction between SCP4 and STK35/PDIK1L might allow the phosphatase to remove inhibitory phosphorylation from the kinases as a mechanism of signaling cooperativity. To investigate this, we mined publicly available phosphoproteomic datasets for the reported STK35/PDIK1L phosphorylation sites (Hornbeck et al. 2015; Ochoa et al. 2020). Among the 18 phosphorylation sites on STK35/PDIK1L identified in prior studies, we noticed several located within the regulatory activation segment of both kinases (Fig. 31). Most eukaryotic protein kinases are activated by phosphorylation of specific residues in the activation loop of the activation segment (sometimes also called the T-loop) that often results in profound conformational changes (Nolen et al. 2004). However, little is known about potentially inhibitory phosphorylation sites in the activation segment. The kinase activation segment is also a protein-protein interaction region that determines the localization and regulation of a kinase and its binding

partners (Wolf et al. 2001). These considerations prompted us to investigate whether SCP4 might remove phosphorylation from this segment.

To investigate the potential role for every phosphorylated serine/threonine residue in the activation segment of STK35 and PDIK1L in their phospho-regulation, we generated MOLM-13 cells stably expressing either CRISPR-resistant phosphomutant for each of these residues (substitution to alanine, a non-phosphorylatable residue) or phosphomimetic residue (substitution to aspartate or glutamate) and performed gene complementation experiments with the endogenous STK35/PDIK1L double knockout. While all of the mutant alleles tested were expressed at comparable levels to the wild-type STK35/PDIK1L, the PDIK1L^{S194D} and STK35^{S385D} proteins were completely defective in supporting AML proliferation (Fig. 32A). In contrast, alanine substitutions of the same residues were fully functional, suggesting that mutations that mimic phosphorylation inhibit the function of these kinases. This supports the role of these residues as inhibitory for their respective kinases. Notably, the PDIK1L^{S194} and STK35^{S385} residues are located at the DFG+2 position of the activation loop for each kinase, which is reported on other kinases to form hydrogen bonds with the DFG phenylalanine to stabilize an active kinase conformation (Kornev et al. 2006; Xie et al. 2013).

The rest of the mutants showed no differential phenotypes between their phosphomutant and phosphomimetic versions. Interestingly, by aligning PDIK1L and STK35 to the other kinases with well-described phospho-regulation, we have nominated PDIK1L^{T217-P} and STK35^{S414-P} residues as the sites for activation loop activating phosphorylation (Fig. 31). However, the substitution of either residue to alanine or glutamate yielded fully functional proteins capable of supporting AML cell proliferation. Of note, simultaneous mutations of both STK35 serine 413 and serine 414 residues to alanines resulted in its complete loss of function (Fig. 32C).

To assess the possibility that SCP4 might act on either one of these phosphorylated residues, we have established collaboration with the group of Prof. Yan Jessie Zhang (College of Natural Sciences, University of Texas, Austin, Texas). Our collaborators purified the recombinant SCP4 phosphatase domain (156-466 amino acids), which was well-folded and exhibited phosphatase activity in para-nitrophenyl assays (Fig. 33A; Zhang, personal communication). We then ordered the commercial synthesis of the four phosphoserine- or phosphothreonine-containing peptides representing the activation segment phosphorylated residues (Fig. 33B). Three out of the four peptides shared an identical sequence and differed only with regard to the phosphorylation site. Zhang's group used the peptides to quantify the phosphatase activity of SCP4 with each phosphopeptide using the malachite green assay to monitor phosphate release. Notably, SCP4 showed great specificity, dephosphorylating only one out of the three peptides with the identical sequence. The peptides corresponding to phosphoserines pS194 and pS216 of PDIK1L (pS385 and pS413 of STK35) were efficiently dephosphorylated by SCP4, whereas no activity was detected for the other two phosphopeptides (Fig. 33B). Our collaborators fitted the kinetic parameters of these substrates into steady-state kinetics and calculated k_{cat}/K_M as 12.63 ± 0.79 and 45.96 ± 3.06 $\text{mM}^{-1} \text{min}^{-1}$ for phosphopeptides containing pS194 and pS216, respectively (Fig. 33C,D).

k_{cat}/K_M is a way to characterize the catalytic efficiency of a given enzyme for a particular substrate. k_{cat}/K_M values around $10^5 \text{ mM}^{-1} \text{ min}^{-1}$ indicate that the conversion of free substrate to product is diffusion controlled, in other words, the reaction happens at collision of enzyme and substrate. The lower the ratio, the slower the catalysis is. However, it is worth noting that the intrinsic phosphatase activity can vary greatly for different enzymes, even between close homologs. Since a phosphatase can also be targeted to its substrate via regulatory subunits (as in the case of PP1) or an allosteric site binding, specificity for a given substrate over another one is a very important parameter.

SCP4 has low intrinsic catalytic efficiency with both peptides *in vitro* compared with some other HAD phosphatases and PTPs (Table 1). Still, it is comparable to an HAD phosphatase Dullard and only slightly weaker than FCP1 (Zhang, personal communication). Nevertheless, SCP4 showed excellent specificity by preferentially dephosphorylating one of the three peptides with the same sequence (Fig. 33B) and overall only exhibiting activity with 2/7 peptides tested (data not shown). Given the strong physical interaction between SCP4 and STK35/PDIK1L (Fig. 22), it is possible that using full-length proteins could give us much higher k_{cat}/K_M values.

Overall, this data supports a model in which SCP4 as part of the phosphatase/kinase complex removes inhibitory phosphorylation from the activation loop of STK35/PDIK1L, which may allow for kinases activation.

Table 1. Kinetic constants of different phosphatases towards their substrates

Gene	Enzyme	Substrate	k_{cat}/K_M	Reference
CTDSPL2	SCP4	KVADFGLP <u>S</u> SKVCSAS	0.2105	This study
CTDSPL2	SCP4	KCFLP <u>S</u> TACGTDFYM	0.766	This study
PPP1CA	PP1wt	KRp <u>T</u> IRR	0.623	Wang et al. 2011
CTDSP1	SCP1	pNPP	0.724	Zhang et al. 2006
CTDSP1	SCP1	YSPTSPSYp <u>S</u> PTSPS	0.166	Zhang et al. 2006
CTDSP1	SCP1	YSPTSPSYSPTp <u>S</u> PS	11.4	Zhang et al. 2006
CTDSP1	SCP1	EDLP <u>S</u> PPSPPLPK	140	Burkholder et al. 2018
PTPN1	PTP1B	YDEDFp <u>Y</u> DYEF	55,000	Selner et al. 2014
PTPN7	HePTP	YDEDFp <u>Y</u> DYEF	25	Selner et al. 2014
PTPN11	SHP-2	WDEDFp <u>Y</u> DWEF	6,800	Ren et al. 2011
PTPN21	PTPD1	YDEDFp <u>Y</u> DYEF	1.2	Selner et al. 2014

Values for k_{cat}/K_M are in $\text{mM}^{-1} \text{sec}^{-1}$

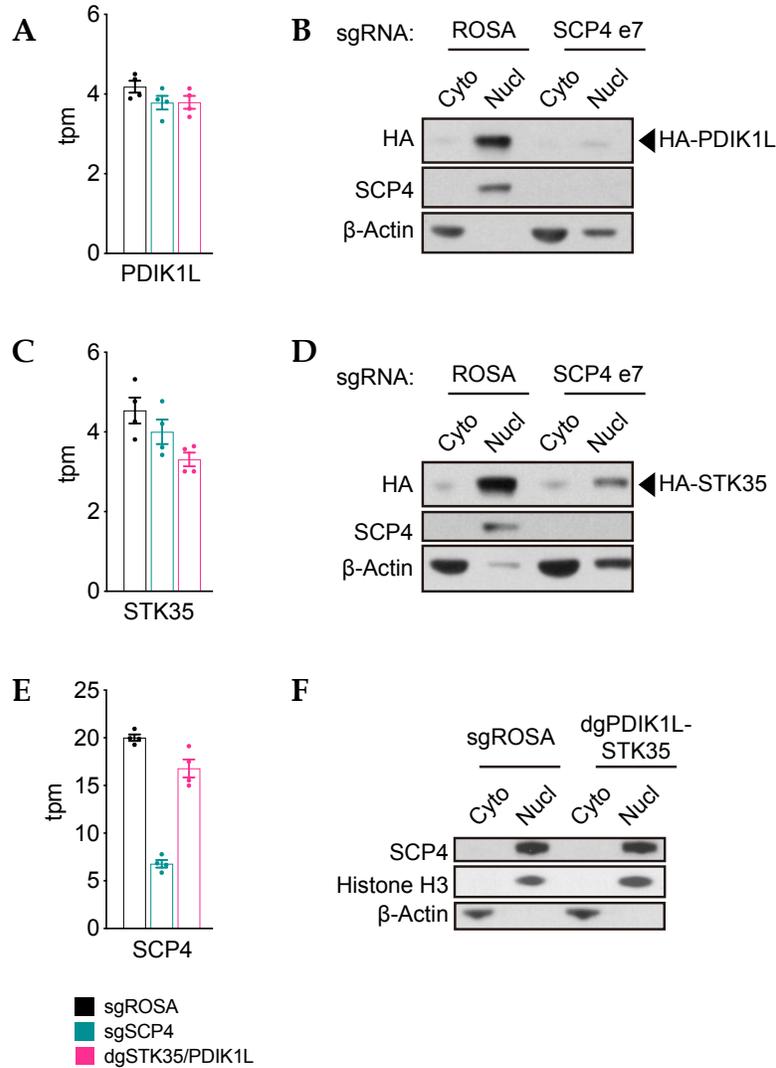


Figure 30. Biochemical evidence that SCP4 functions upstream of STK35/PDIK1L.

(A,C,E) RNA-Seq transcripts per million (tpm) data for PDIK1L, STK35, and SCP4 from MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. Plotted are data from each individual RNA-Seq replicate, the mean \pm SEM. (B,D,F) Western blots of cytoplasm (Cyto) and nucleus (Nucl) fractions of MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. $n = 3$, representative images shown.

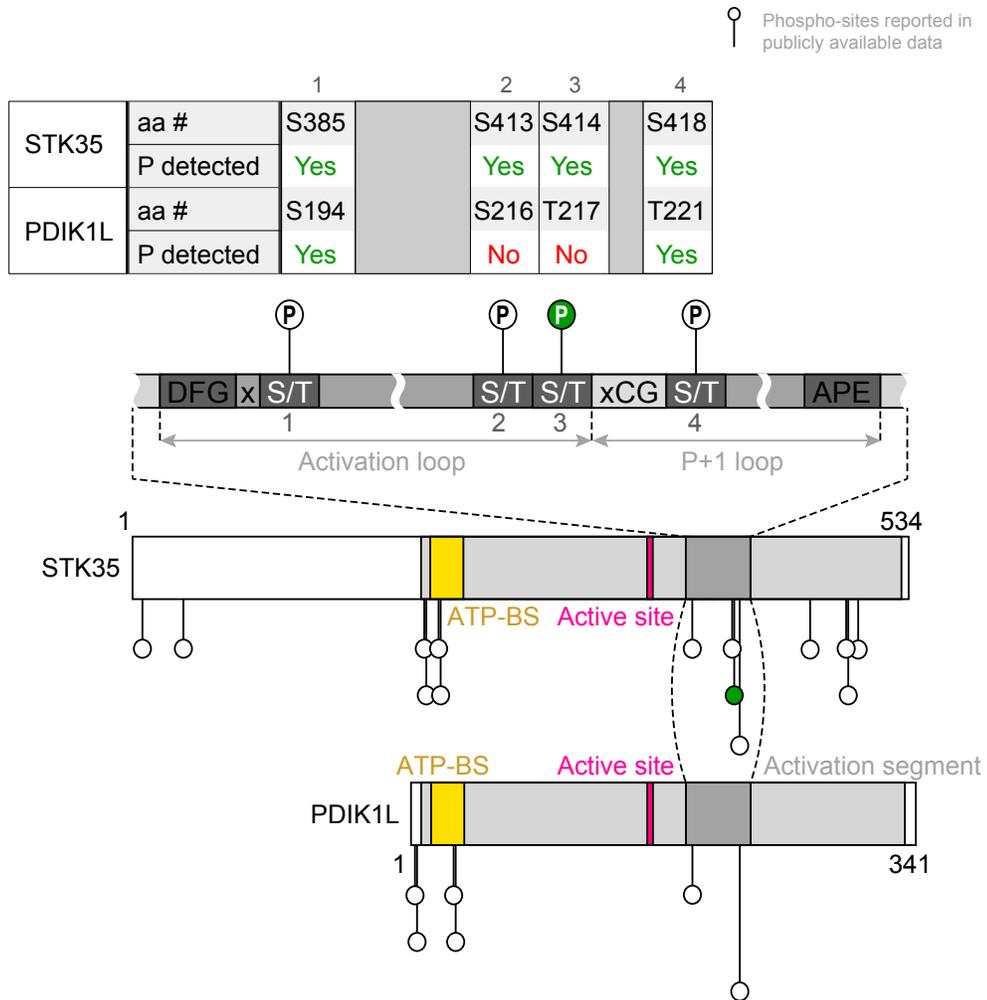


Figure 31. Schematics of STK35/PDIK1L phosphorylation sites.

Schematics of STK35/PDIK1L phosphorylation sites reported in the publicly available phospho-proteomics datasets on STK35 and PDIK1L relative to their domain architectures.

Based on data from Hornbeck et al. 2015 and Ochoa et al. 2020.

aa #, amino acid number; P, phosphorylation; ATP-BS, ATP-binding site; predicted activating phosphorylation shown in green.

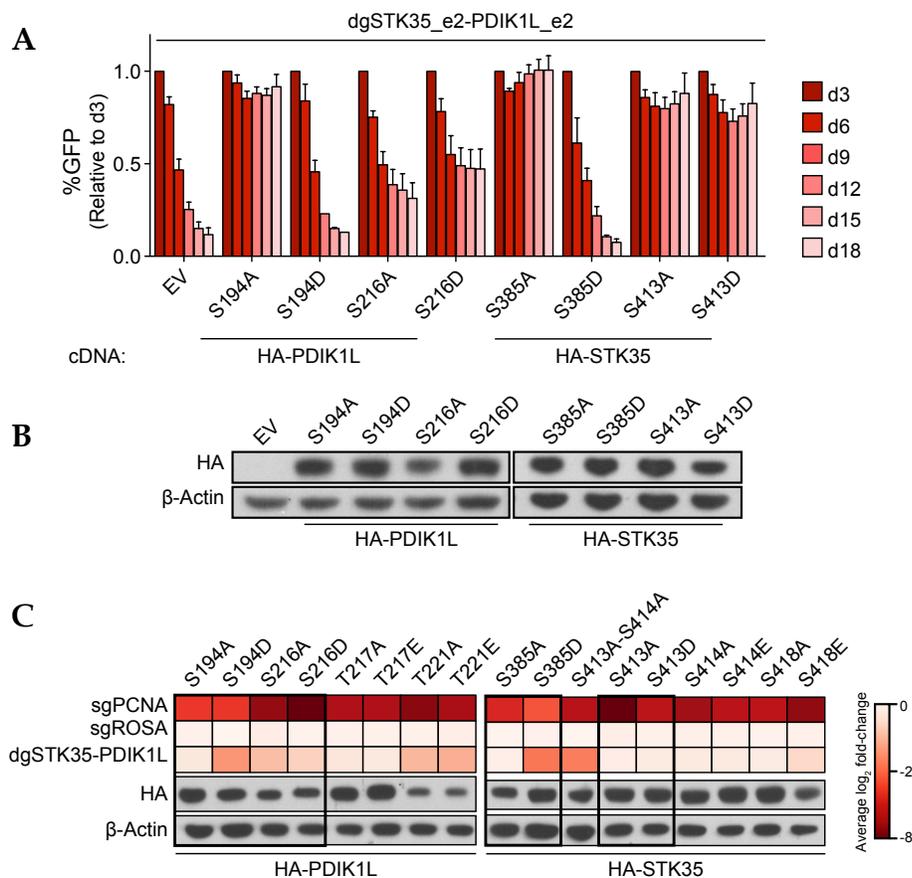


Figure 32. S194 on PDIK1L and S385 on STK35 could serve as inhibitory phosphorylation residues. (A) Competition-based proliferation assay in MOLM-13 cells stably expressing empty vector (EV) or the CRISPR-resistant HA-PDIK1L or HA-STK35 constructs harboring the indicated amino acid substitutions infected with the indicated sgRNAs. $n = 3$. (B) Western blot of whole-cell lysates from MOLM-13 cells stably expressing empty vector (EV) or CRISPR-resistant HA-PDIK1L or HA-STK35 constructs harboring the indicated amino acid substitutions. (C) Summary of competition-based proliferation assays in MOLM-13 cells stably expressing CRISPR-resistant HA-PDIK1L or HA-STK35 constructs harboring the indicated amino acid substitutions infected with the indicated sgRNAs. Plotted is the fold change (\log_2) of sgRNA⁺/GFP⁺ cells after 18 days in culture (average of triplicates). Below are shown Western blots of whole-cell lysates of the cells stably expressing the indicated constructs. Black frame indicates data corresponding to panel A.

All bar graphs represent the mean \pm SEM. All sgRNA experiments were performed in Cas9-expressing cell lines. ‘e’ refers to the exon number targeted by each sgRNA. ‘dg’ refers to the bi-cistronic vector for simultaneous targeting of STK35 and PDIK1L. The indicated sgRNAs were linked to GFP. GFP⁺ population depletion indicates loss of cell fitness caused by Cas9/sgRNA-mediated genetic mutations.

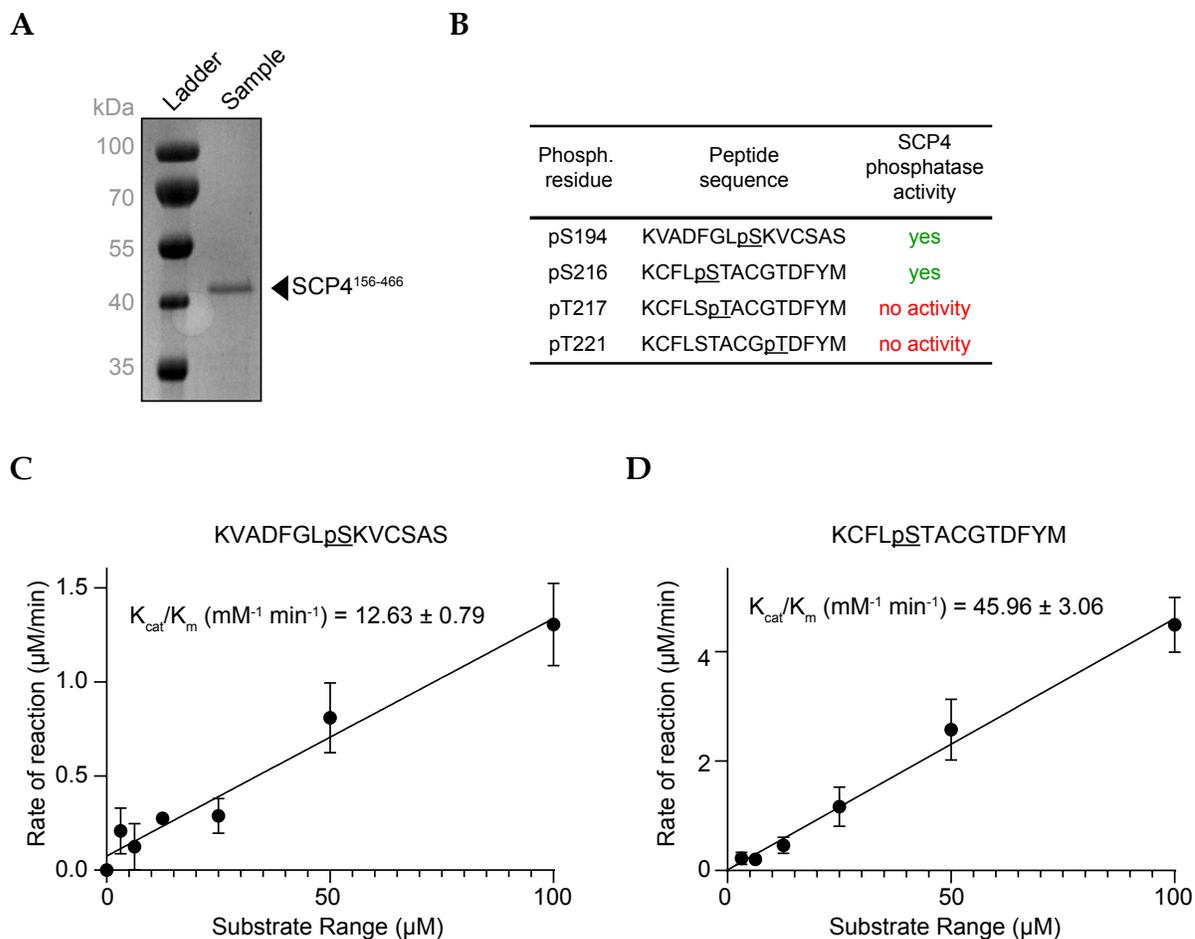


Figure 33. SCP4 exhibits exquisite substrate specificity *in vitro*.

(A) Recombinant SCP4 protein purity as assessed by SDS-PAGE and Coomassie blue staining. His-SUMO-(TEV)-SCP4 was expressed in BL21 (DE3) cells and purified by affinity, anion exchange, and gel filtration chromatography. Molecular weight markers are shown for reference. (B) Phosphorylated peptides assayed for SCP4 phosphatase activity *in vitro*. (C,D) Phosphatase activity of SCP4 with the indicated peptides plotted for kinetic fitting. Each measurement was conducted in triplicate with standard deviations shown as error bars.

Data in (A), (C), and (D) was collected and analyzed by Rosamaria Y. Moreno, Yan Jessie Zhang.

2.9. SCP4-STK35/PDIK1L sustains the aberrant metabolic state of AML

We next sought to understand the downstream output of this phosphatase/kinase complex that could account for its essential function in AML. From the aforementioned RNA-Seq experiments, we identified 678 genes that were significantly downregulated following both SCP4 knockout and STK35/PDIK1L double knockout in MOLM-13 cells and 892 – significantly upregulated (Fig. 34A). The overlap between downregulated genes upon both perturbations was significant, $\chi^2(1, N = 3536) = 490.5493$, $p < .00001$. Similarly, the overlap between genes upregulated upon SCP4 knockout and STK35/PDIK1L double knockout was significant, $\chi^2(1, N = 3841) = 385.4878$, $p < .00001$. Overall, significant gene expression changes upon SCP4 depletion were more likely to overlap with significant gene expression changes upon dual STK35/PDIK1L targeting than not. This serves as yet another piece of evidence for the remarkable correlation among the global gene expression changes we observed (Fig. 28).

Ontology analysis of the downregulated genes revealed enrichment for biological pathways concerning amino acid metabolism and transport functions (Fig. 34B). For example, all the major transaminases that catalyze reversible interconversion of glutamate into other amino acids were downregulated as well as many of the solute carriers that import amino acids into the cell (Fig. 34C).

Given the profound ways amino acid biosynthesis and transport are known to be deregulated to support the proliferation of AML cells (see Introduction §1.4.1.2), we decided to measure total amino acid levels in AML cells upon SCP4 knockout and STK35/PDIK1L double knockout via untargeted liquid chromatography-mass spectrometry (LC-MS) analysis based on an in-house library of 236 hydrophilic metabolites.

Both SCP4 knockout and PDIK1L-STK35 double knockout resulted in global effects on leukemia cell metabolism that were consistent between the two conditions in the three

independent biological replicates (Fig. 35A). The levels of several metabolites remained unchanged in these knockout cells, but many amino acids were present at reduced levels in cells deficient for either SCP4 or STK35/PDIK1L (Fig. 35B).

For example, one of the most deficient amino acids consistent between replicates was proline (Fig. 35C), which could be related to the decrease in pyrroline-5-carboxylate reductase 1 (*PYCR1*) expression, encoding the rate-limiting enzyme for proline biosynthesis (Xiao et al. 2020) (Fig. 35D). The decreased expression of *PYCR1* was shown to be implicated in the rewiring of proline metabolism from the TCA cycle to the urea cycle (Sharif et al. 2019). In line with this hypothesis, we have observed consistent downregulation of several TCA cycle metabolites, whereas many urea cycle metabolites were among the least affected in all three biological replicates (Fig. 36A). The urea cycle is closely linked to the TCA cycle through the reversible reaction providing fumarate for the TCA cycle and *de novo* arginine biosynthesis (Fig. 36B). The rate-limiting step for this connection, argininosuccinate biosynthesis, also was affected by SCP4 knockout and STK35/PDIK1L double knockout both at the enzyme argininosuccinate synthetase 1 (*ASS1*) expression level (Fig. 36C) and argininosuccinate levels.

Since other amino acids are also known to fuel the TCA cycle (Fig. 36B), as well as serving biosynthetic blocks for the rapidly growing AML cells, we propose that SCP4 knockout and STK35/PDIK1L double knockout may result in multifaceted deregulation of metabolic homeostasis of AML cells.

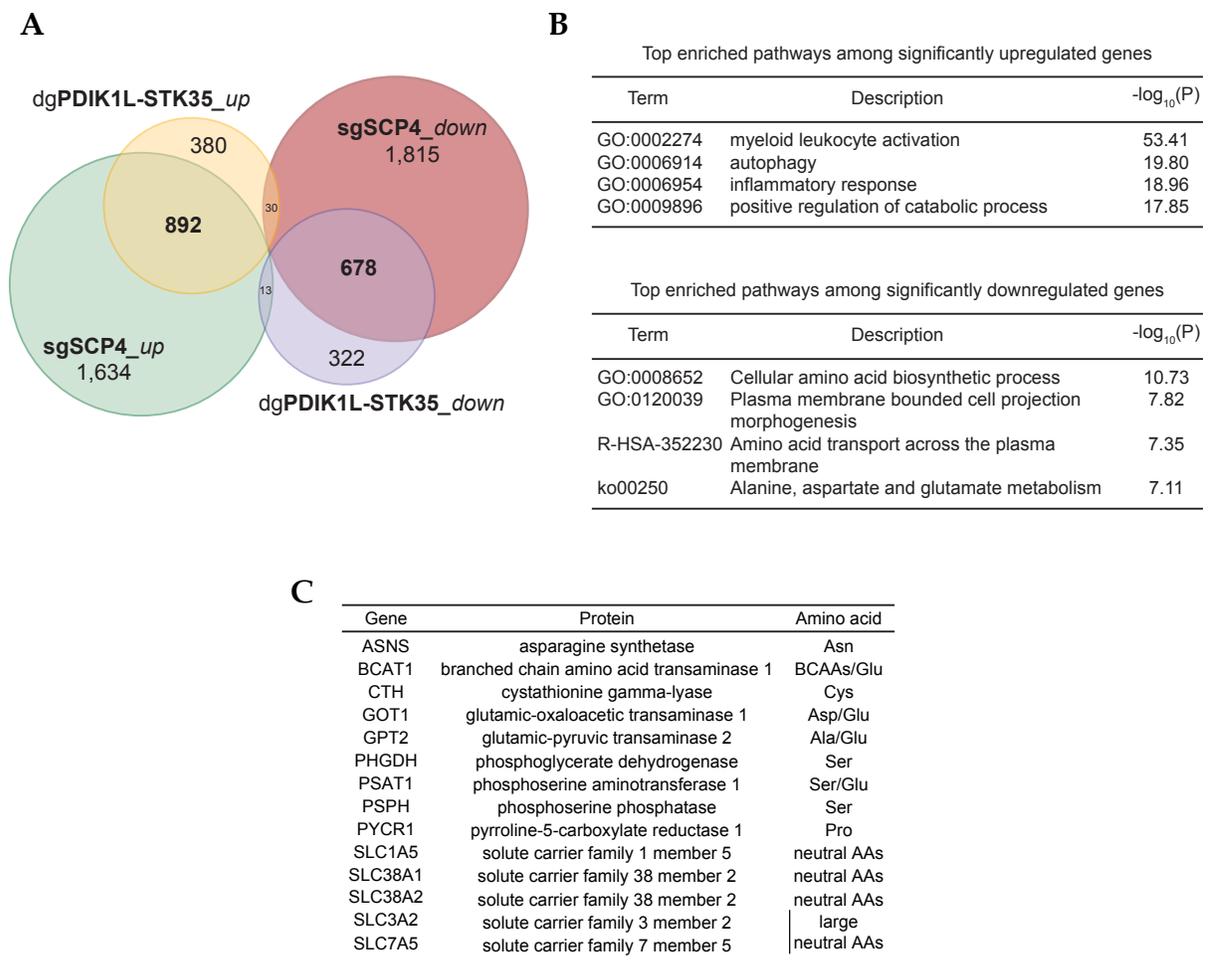


Figure 34. Nominating the downstream output of the SCP4-STK35/PDIK1L complex.

(A) Venn diagram depicting the overlap between statistically significant up- and downregulated genes in MOLM-13 cells upon SCP4 knockout and STK35/PDIK1L double knockout. DeSeq2 ($n = 4$). (B) Ontology analysis of overlapping statistically significant up- and downregulated genes in MOLM-13 cells upon both SCP4 knockout and STK35/PDIK1L double knockout. (C) Selected statistically significant downregulated genes in MOLM-13 cells upon both SCP4 knockout and STK35/PDIK1L double knockout and the amino acids they are involved in biosynthesis or transport.

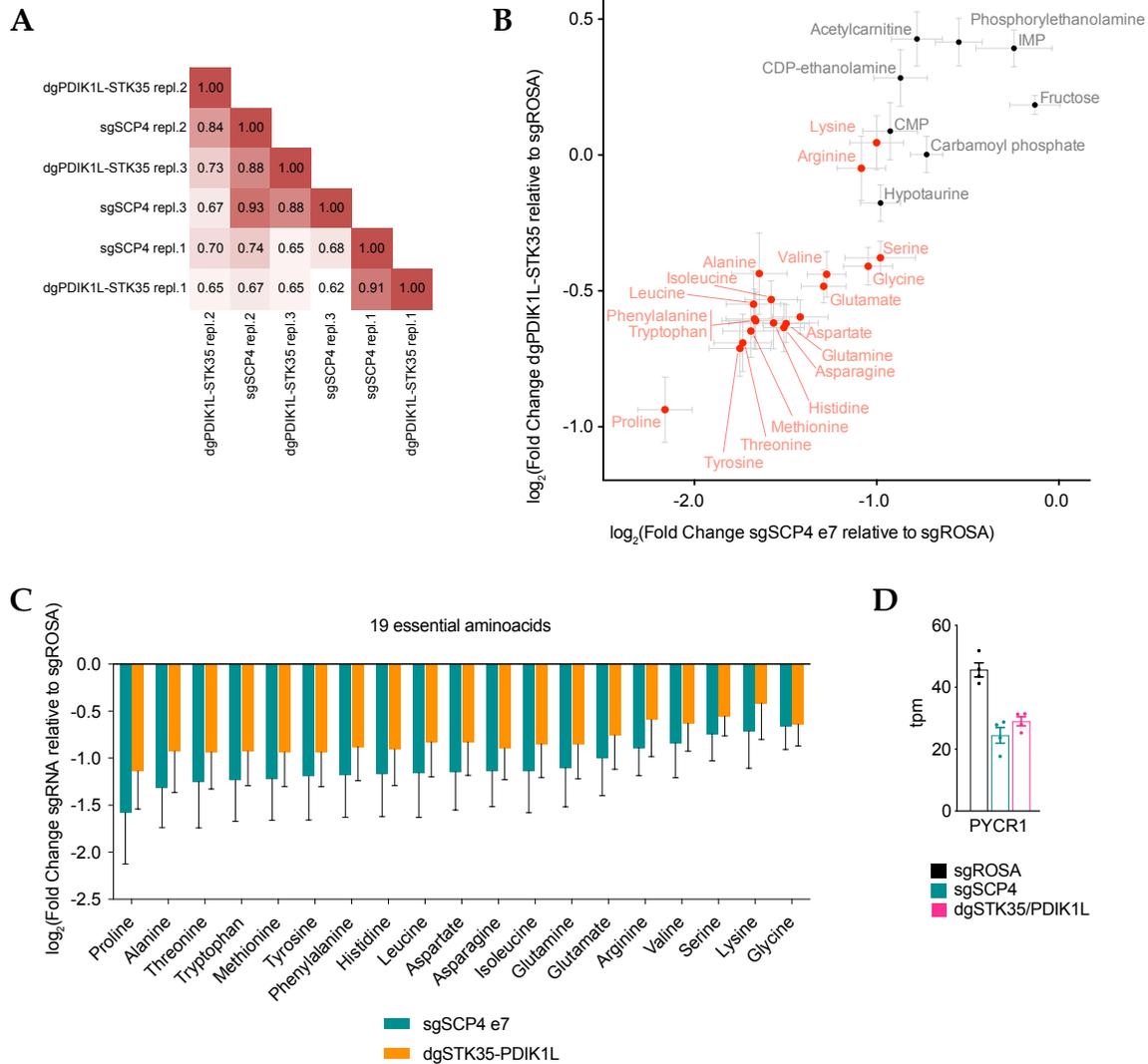


Figure 35. SCP4 knockout and PDIK1L-STK35 double knockout result in consistent changes in the leukemia cell metabolism.

(A) Correlation matrix for the global log₂ fold-changes of metabolites levels upon SCP4 knockout and STK35/PDIK1L double knockout relative to negative control as measured in each independent biological replicate across 80 metabolites. (B) The correlation between log₂ fold changes in the levels of selected metabolites upon SCP4 knockout and STK35/PDIK1L double knockout relative to negative control as measured by the MS analysis. Every dot represents the mean ± SEM (n_{tech} = 6). Amino acids are in red, few unchanged metabolites in black for reference. Shown is a representative experiment of three independent biological replicates. (C) log₂ fold-changes of amino acids ranked from lowest to highest. n_{bio} = 3. (D) RNA-Seq transcripts per million (tpm) data for *PYCR1* from MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. Plotted are data from each individual RNA-Seq replicate, the mean ± SEM.

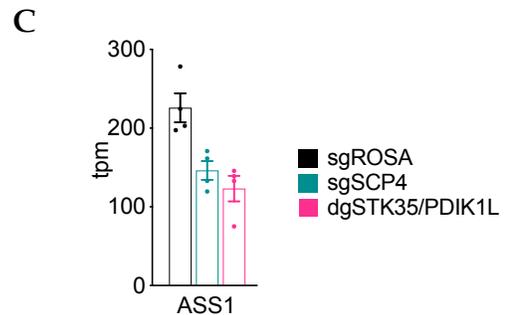
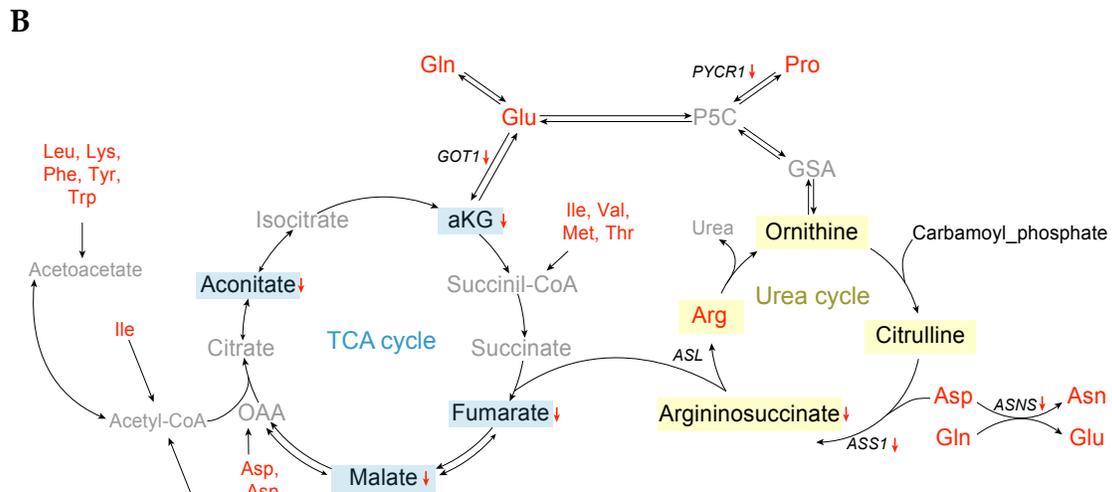
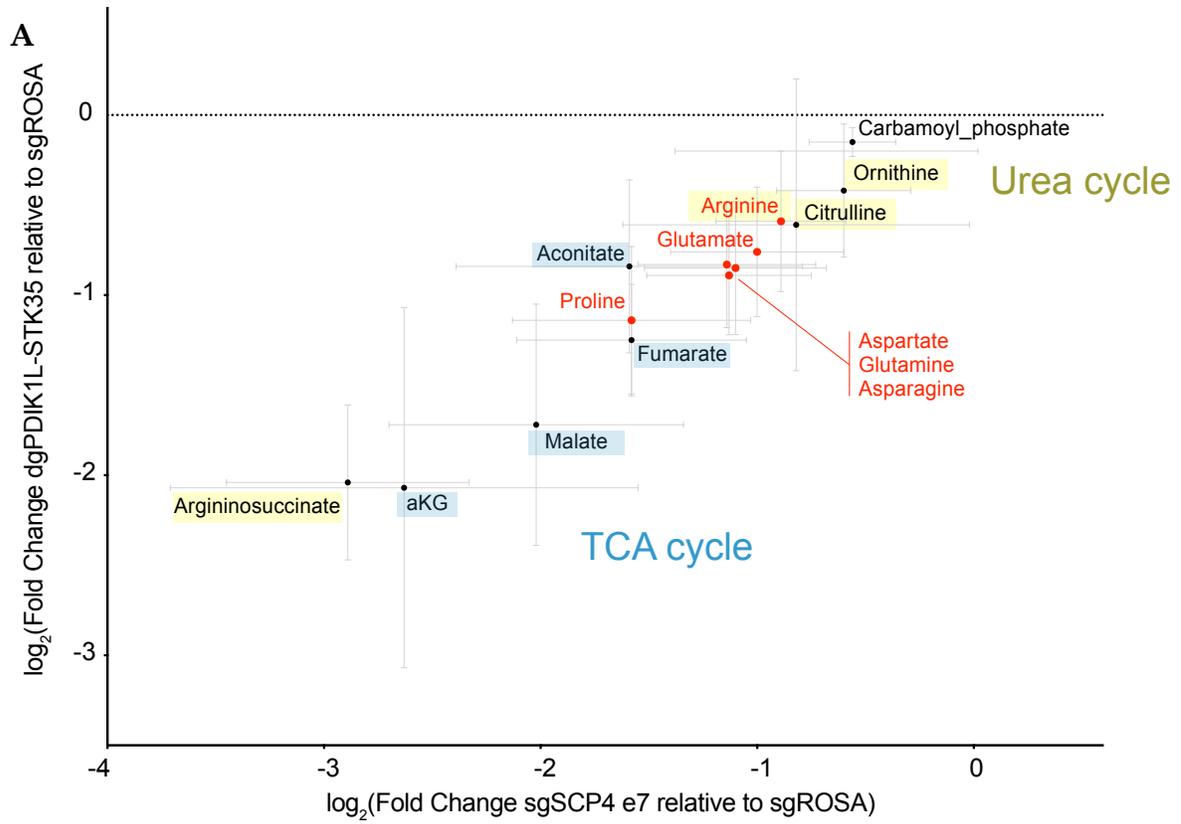


Figure 36. SCP4-STK35/PDIK1L signaling supports the aberrant metabolic state in AML.

(A) The correlation between \log_2 fold changes in the levels of selected metabolites upon SCP4 knockout and STK35/PDIK1L double knockout relative to negative control as measured by the MS analysis. Every dot represents the mean \pm SEM ($n_{\text{bio}} = 3$). **(B)** Selected amino acids and their functional links to TCA, glucose metabolism, and urea cycles. Selected downregulated enzymes are shown. Adapted from Albaugh et al. 2017. **(C)** RNA-Seq transcripts per million (tpm) data for *ASS1* from MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. Plotted are data from each individual RNA-Seq replicate, the mean \pm SEM.

Amino acids are in red; metabolites of urea cycle are in yellow; TCA cycle – in blue; metabolites not detected by LC-MS – in gray.

2.10. Results summary

In this study, we have employed phosphatase domain-focused CRISPR screens to comprehensively assay the requirements for human phosphatases in different cancer cell lines with a specific focus on AML. We have discovered that a poorly studied nuclear phosphatase SCP4 is an acquired dependency largely specific to AML cell lines. We showed that SCP4 role in AML is conserved between species. SCP4 knockout slowed down the AML progression in a mouse xenograft model. We have provided preliminary evidence suggesting that SCP4 is not required for proliferation and differentiation of human HSPCs, nominating it for further exploration as a potential therapeutic target in AML.

Using biochemical and genetic approaches, we showed that the functional catalytic domain of SCP4 in the context of its C-terminal fragment SCP4²³⁶⁻⁴⁶⁶ is necessary and sufficient for supporting leukemia cell proliferation. Unbiased SCP4 interactome analysis revealed that SCP4 tightly associates with the kinase paralogs STK35 and PDIK1L. We showed that STK35 and PDIK1L play redundant roles in supporting AML cell proliferation, and their kinase catalytic activity is required.

Our global transcriptome studies implicated SCP4 and STK35/PDIK1L as functioning in the same AML maintenance pathway. We provided evidence that SCP4/kinases complex might serve signaling cooperativity between these two enzymes by (1) regulating protein levels and (2) activating the kinases. Moreover, we have discovered a novel site for potentially inhibitory phosphorylation at the eukaryotic kinase activation loop that has not been previously explored to our knowledge and might serve as a conserved kinases regulation mechanism.

We reveal that one downstream output of this signaling complex is to maintain the expression of amino acid biosynthetic enzymes and transporters. These findings lead us to speculate that the SCP4-STK35/PDIK1L complex is required in AML to support the

pathological metabolic requirements in this disease. Given the novelty of our research in the context of the current data available for these relatively poorly studied enzymes, we have outlined the important mechanistic details for SCP4 and STK35/PDIK1L biochemistry that draw upon and expand the general principles of inferring biological pathways and mechanisms of protein regulation.

Overall, we propose a model in which SCP4 interaction with STK35/PDIK1L in the nucleus results in stabilization and activation of the kinases. Together, SCP4-STK35/PDIK1L signaling supports the aberrant metabolic state of AML (Fig. 37).

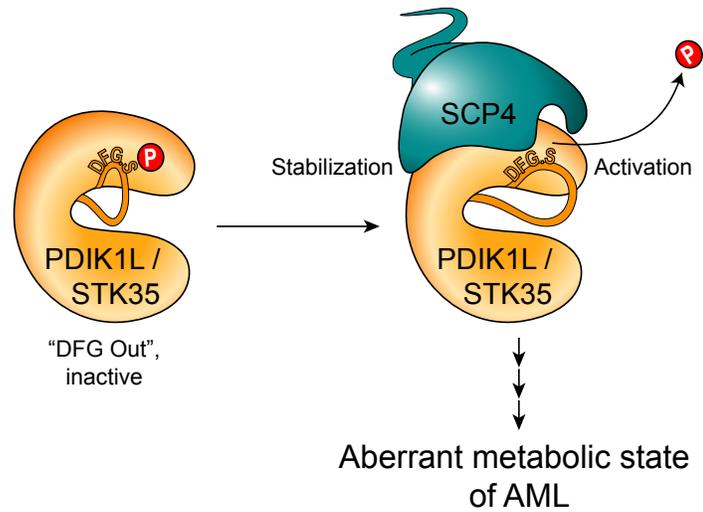


Figure 37. Model for SCP4-STK35/PDIK1L cooperativity in supporting the aberrant metabolic state in AML.

STK35/PDIK1L in the cell are inactive when phosphorylated on the DFG+2 serine residue at the start of the activation loop of the activation segment. Docking of SCP4 on STK35/PDIK1L leads to STK35/PDIK1L stabilization and activation through dephosphorylation of the inhibitory serine residue. The complex downstream genetic program is at least in part concerned with supporting the aberrant metabolic state in AML.

3. DISCUSSION AND PERSPECTIVES

3.1. The novelty

The Human Genome Project was completed in April 2003 with 99% of the human genome sequenced and mapped to an accuracy rate of less than one error every 10,000 base pairs. The decade of genome sequencing stimulated an explosion in cloning and characterization of full-length cDNAs, aiding identification of transcriptional units and initial description of unknown proteins (Zhang et al. 2000; Kemmer et al. 2006; Wood et al. 2019). However, the precise biochemical and cellular functions of 20–30% of human genes, by different estimates, remain an enigma to this day (Ellens et al. 2017; Wood et al. 2019). Among the difficulties in characterizing the unknowns of the human proteome are: (1) lack of a clear link to a phenotype; (2) absence of orthologs in tractable model systems; (3) lack of sequence homology to other characterized genes; (4) lack of predicted domain organization. There are also epistemological considerations, such as research bias in providing more details for the previously described proteins. For example, Edwards et al. in a 2011 essay “Too many roads not taken” observed that around 70% of human protein research was still focused on the 10% of proteins known even before the human genome was sequenced (Edwards et al. 2011).

Cancer cells streamline all the cellular processes for the ultimate goal of unlimited survival and proliferation and, as a result, sometimes expose roles for proteins otherwise hidden by redundancy and cooperativity between biological pathways in the healthy cells. In this regard, systematic knockout screens in a diverse set of cancer cell lines can be instrumental in revealing novel functions of both poorly and well-studied proteins. Standardized data portals, such as DepMap, provide an invaluable tool for researchers for simultaneous assessment of dependencies on a given gene across hundreds of human cancer cell lines and data integration with analogous characteristics for other genes, including co-

dependencies, expression, and mutations profiles. However, these comprehensive databases have their blind spots that can additionally impede research progress in the case of poorly studied genes.

For example, the topic of this dissertation, SCP4, could not be adequately assessed as a dependency in DepMap since all the sgRNAs used in Project Achilles/DepMap genome-wide screens exclusively targeted its N-terminus. We have conclusively shown that the 1-235 amino acid segment of SCP4 was dispensable for the protein function and “invisible” for the dropout screening efforts, at least in the context of human leukemia (Fig. 14; Fig. 15).

An additional technological limitation well-recognized in the field is the functional redundancy of paralogous genes, which also occluded the utility of DepMap data for STK35/PDIK1L. In this dissertation, we showed that catalytically active STK35 and PDIK1L could substitute for one another in the context of AML, and the requirement for their function correlated with that of SCP4 (Fig. 23-27). Curiously, even with the confounding factors mentioned above, the latest release of DepMap co-dependencies data (DepMap 21Q2 Public + Score, CERES) has identified SCP4/*CTDSPL2* as the top positively correlated co-dependency of STK35 and *vice versa*.

While this dissertation illustrates the capability of thoughtfully designed CRISPR-Cas9 genetic screens to reveal novel signaling pathways in cancer, linking the dependencies on SCP4 and STK35/PDIK1L to any other features of the examined cell lines proved challenging. Based on our screens and the publicly available data of the driver mutations in the assayed cancer cell lines, we can generate hypotheses for the synthetic lethal hits in the context of the most common and well-described oncogenes and tumor suppressors such as *RAS* or *TP53* (Fig. 4). In contrast, the precise molecular underpinnings of the dependency on the SCP4-STK35/PDIK1L signaling complex require further studies. Nevertheless, this work provides for the first time the phenotype link and the relevant mechanistic details for SCP4

and STK35/PDIK1L that will hopefully reveal the roles that these enzymes play in health and disease, specifically in the context of AML.

3.2. Dissecting the roles for the N-terminus in SCP4 regulation

Our mechanistic studies of SCP4 started by contrasting the CRISPR exon scanning approach with amino acid evolutionary conservation and sequence-based prediction of intrinsically unstructured regions (Chang et al. 2018; Meszaros et al. 2018). We observed a remarkable positive correlation of the likelihood for CRISPR-induced indel mutagenesis to disrupt SCP4 functionality with the degree of amino acid conservation and a negative correlation with the prediction for region disorder (Fig. 14). All three approaches gave very similar estimates for the SCP4 segment starting from amino acid 236 being highly conserved, structured, and essential for AML cell proliferation. These observations serve as another testament to the utility of the domain-focused CRISPR-screening approach developed in our laboratory (Shi et al. 2015).

However, the N-terminal domain of SCP4 is present across many species, and it is reasonable to hypothesize that it serves specific functions in SCP4 regulation and maintenance. Our work provides evidence that the N-terminus is recognized by importins α and β to aid its nuclear localization (Fig. 19C; Fig. 20). There are a few definitive experiments we would propose for future studies in this regard. For example, we could mutate the predicted nuclear localization signals to sequences lacking characteristic arginines and lysines and enriched for hydrophobic residues and look for a deficiency in nuclear localization upon fractionation and the absence of complementation for the knockout of the endogenous gene by the misplaced SCP4. The caveat of this approach is the peculiar degradation we observed for the full-length SCP4 upon cytoplasmic exposure in AML cells (Fig. 18). An alternative approach would be to fuse the functional SCP4²³⁶⁻⁴⁶⁶ segment to an

N-terminal GFP with a nuclear export signal that would result in its size-restricted inability to freely pass through the nuclear pore and additional active sequestration in the cytoplasm.

Another potential role for the N-terminus emerged from our data suggesting that the full-length protein could be a target of myeloid proteases (Fig. 18). In fact, Zhong et al., who reported myeloid cells specific degradation of HOXA9, had to perform triple knockout of *Elane/Prtn3/Ctsg* myeloid proteases in mice to be able to precipitate enough HoxA9 protein for ChIP-Seq experiments since no effective inhibitors exist for these proteases (Zhong et al. 2018). Whether our efforts at performing SCP4 ChIP-Seq remained unsuccessful due to the same issue remains to be seen. However, SCP4²³⁶⁻⁴⁶⁶ protein was present in the cytoplasm (Fig. 17C), which leads us to speculate that the N-terminus of SCP4 might also serve as the landing site for the myeloid proteases. Myeloid proteases have been mostly investigated with regard to their extracellular roles in the context of infection or inflammation (Benarafa and Simon 2017). However, the competition between the importins binding the N-terminus to aid SCP4 nuclear localization and the myeloid proteases-mediated cytoplasmic degradation of SCP4 could serve its regulation in the cell. One piece of evidence in favor of this hypothesis would be if we detected SCP4 in the cytoplasm upon triple proteases knockout. Another experiment would involve blocking nuclear import through targeting the importins identified in SCP4 IP-MS via gene editing or pharmacologically and measuring total SCP4 levels. Of note, *CTSG* is one of the significantly down-regulated genes upon both SCP4 knockout and *STK35/PDIK1L* double knockout that could indicate a potential negative feedback loop if our hypothesis about the role of myeloid proteases in SCP4 regulation was true (Supplemental Fig. S1).

3.3. Phosphatase/kinase complexes

The interactions between SCP4 and PDIK1L or STK35 were detected in IP-MS experiments after stringent washes with salt concentrations ranging from 150 mM up to 500 mM NaCl (Fig. 22A,B). Moreover, our data suggested a potentially 1:1 complex between ectopically expressed HA-PDIK1L and endogenous SCP4 (Fig. 22C), possibly due to PDIK1L being less stable in the absence of SCP4 (Fig. 30B). Nevertheless, this interaction was weakened by a single point mutation of the SCP4 active site (Fig. 21). Since the interactions of enzymes with their substrates are often transient (Fahs et al. 2016), it is unlikely that the strong binding we detect between SCP4 and STK35/PDIK1L is entirely due to an active site interaction with phosphorylated residues. While our experiments indicate that the activation loop of STK35/PDIK1L could be a substrate of SCP4, we speculate that the binding mechanism between SCP4 and STK35/PDIK1L represents docking involving non-catalytic surfaces of these two proteins (Remenyi et al. 2006).

Phosphatase/kinase complexes, while not extremely common, have been reported in cell biology, for example, for the kinases involved in the cell cycle progression (Helps et al. 2000), MAPK/ERK pathway (see below). Other examples include phosphoinositide (PI) regulatory network (Lecompte et al. 2008) and cAMP-dependent protein kinase (PKA) complexes (Colledge and Scott 1999) that employ scaffold proteins for forming stable signaling complexes. Among the potential advantages of a stable phosphatase/kinase complex formation are simultaneous localization of both enzymes to a specific substrate and the precise temporal regulation of the phospho-signaling output of the complex due to the fixed phosphatase/kinase balance (Botelho 2009; Szomolay and Shahrezaei 2012).

The A-Kinase Interacting Proteins, or AKAPs, were among the first families of scaffold proteins recognized for their role in cell signaling specificity (Colledge and Scott 1999). AKAPs bind to the regulatory subunit of PKA and other enzymes in the cAMP

pathway and target the resulting signaling complexes to their subcellular compartments. One such AKAP, AKAP79/150, binds PKA, protein kinase C, and calcineurin (*PPP3CB*), a calcium-dependent protein phosphatase, and directs them to calcium entry sites, such as neuronal L-type calcium channels (Oliveria et al. 2007). Another AKAP, AKAP9 (AKAP350/450, or Yotiao), provides a scaffold for the signaling complex between PKA, protein phosphatase 1, and PDE4D3 and targets it to NMDA receptors (Westphal et al. 1999). Both complexes allow for the acute bidirectional regulation of their cellular targets critical for the precise coordination of certain synaptic and cytoskeletal signaling events.

Protein phosphatase 5 (PP5) was found to interact with Raf1 catalytic domain and ERK1/2 directly, and these interactions were additionally facilitated by the heat-shock protein Hsp90 (von Kriegsheim et al. 2006; Mazalouskas et al. 2014). Of note, in our own MS data, we observed SCP4 association with Hsp70, encoded by *HSPA5* (Fig. 19C). Heat-shock proteins are molecular chaperones that catalyze appropriate protein folding and are often present in functional heterocomplexes. Identifying and disrupting Hsp70 binding surface could be informative on its functional role in SCP4 complexes formation, if any. PP5 inhibited Raf1 by dephosphorylating pS338 outside of its activation loop, and its association with Raf1 kinases ERKs was reported to precisely regulate Raf1 phosphorylation output.

One prominent example of a direct phosphatase/kinase complex was discovered at Cold Spring Harbor Laboratory in 1994 for the cyclin-dependent kinase 2 (CDK2) and kinase-associated phosphatase (KAP), encoded by *CDKN3* gene (cyclin-dependent kinase inhibitor 3) (Hannon et al. 1994; Poon and Hunter 1995). The complex between KAP and CDK2 was, in fact, so robust that it allowed solving the crystal structure for these two proteins by the group of David Barford, who also was a Cold Spring Harbor Laboratory Fellow between 1991 and 1994 (Song et al. 2001). While the binding interface between KAP and CDK2 involved extensive interactions remote from the active sites of both enzymes, the

activation segment of CDK2 was reported to interact with the catalytic site of KAP via the phosphate group of the phosphorylated threonine residue (pT160) in the CDK2 activation loop. KAP inhibits CDK2 through dephosphorylation of its activating residue, thus promoting timely mitotic exit.

Ca²⁺/CaM-dependent protein kinase IV (CaMKIV) forms a signaling complex with B α - and B δ -containing PP2A holoenzymes (Reece et al. 2009). This complex allows the characteristic transient nature of CaMKIV activation in response to its stimuli since PP2A rapidly dephosphorylates the pT200 residue within the CaMKIV activation loop and returns its activity to the baseline level within minutes upon activation (Westphal et al. 1998; Anderson et al. 2004). This case illustrates how the formation of multiprotein complexes regulates the timing of signaling events in addition to specificity (Pawson and Nash 2000; Hunter 2000).

In yet another example of a stable phosphatase/kinase complex, mitogen-activated protein kinase ERK2 bears a docking domain composed of the negatively charged amino acids located opposite the kinase active site, which was shown to interact with the positively charged residues in the kinase interaction motives of the MAP kinase phosphatase MKP3. These interactions ensure exquisite specificity of ERK2 regulation by MKP3. MKP3 also deactivates ERK2 via dephosphorylation of pT183 and pY185 in its activation loop. Importantly, additional interactions between the ERK2 activation segment and MKP3 facilitate MKP3 activation by realigning its catalytic residues. Single amino acid substitutions of the residues of MKP3 directly involved in ERK2 binding to alanines could disrupt the MKP3-ERK2 complex (Camps et al. 1998; Liu et al. 2006; Zhou et al. 2006).

In contrast to the diminished interaction of SCP4^{D293A} with the kinases (Fig. 21), catalytically inactive phosphatase MKP3^{C293S} has a similar binding affinity with ERK2 to that of the wild type MKP3 (Zhou et al. 2001). In addition, the MKP3^{C293S} mutant enhances ERK2

phosphorylation and limits its localization to the cytoplasm (Brunet et al. 1999). An important experiment would be to assess phosphorylation levels and any differences in expression or localization of STK35/PDIK1L upon the expression of the SCP4 inactive mutant.

The SCP4^{D293A} protein in our study may adopt an inactive conformation, which simultaneously impairs both catalysis and docking with STK35/PDIK1L. One piece of evidence in favor of this hypothesis could be that SCP4^{D293A} catalytic mutant has its specific binding partners: FBXW11, BTRC, and CUL1 (Fig. 19C). FBXW11 and BTRC are paralogs, F-box proteins, that together with CUL1 constitute two of the four subunits of the ubiquitin ligase complex called SCF (SKP1-CUL1-F-box protein). We have confirmed these MS results by IP-Western blot (Supplemental Fig. S2). Future structural studies of the SCP4-STK35/PDIK1L complex coupled with mutagenesis could shed further light on the specific binding interactions between these catalytic domains.

Lastly, the stable complex formation between an HAD-phosphatase SCP4 and either one of its macromolecular kinase partners could be synergistic, providing the resulting complex access to additional substrates not accessible for each enzyme separately. Given the disordered nature of SCP4 N-terminus (Fig. 14B), we can assume it belongs to the ‘capless’ C0-type HAD phosphatases (see Introduction §1.5.2). The capless HAD phosphatases are known to oligomerize to process low molecular weight substrates (Parsons et al. 2002; Wu and Woodard 2003; Lu et al. 2009; Biswas et al. 2009). Structural studies of SCP4 in the complex with STK35/PDIK1L are required to determine whether the unstructured N-terminus adopts a new conformation or whether the kinase subunit of the complex could potentially serve as a “pseudocap” to sterically restrict the solvent access to the otherwise exposed HAD active site and hence allow SCP4 to act on low molecular weight substrates.

3.4. Phospho-regulation of eukaryotic kinases

In each of the examples of phosphatase/kinase complexes discussed above, the docking interactions between phosphatases and kinases led to allosteric modulation of catalytic kinase activity (Remenyi et al. 2006). Our study provides evidence that a similar docking interaction and potential for allosteric regulation exist between SCP4 and STK35/PDIK1L, leading to stabilization of the kinase substrates and positioning them for dephosphorylation. Our biochemical assays and mutational analyses implicate analogous serines at the activation loops of STK35 (pS385) and PDIK1L (pS194) as substrates of SCP4.

In terms of kinases regulation, activating phosphorylation at conserved serine, threonine, or tyrosine residues of the activation segment is a well-described mechanism of activation for multiple kinases (Nolen et al. 2004; Ten Eyck et al. 2008). Interestingly, for STK35/PDIK1L, phosphorylations on the predicted activating S/T residues were likely dispensable since the kinases were functional when either residue was mutated one by one to alanines (Fig. 32C). However, for STK35, we showed that mutating both serines at the activating position to alanines at the same time renders a null mutant (STK35^{S413A-S414A} in Fig. 32C). The latter suggests that either one of the adjacent S/T residues at the activating position could be required for the kinase activation, although additional mutagenesis experiments are required to prove that. Similar observations were made for the Aurora A kinase (Rowan et al. 2013). On the contrary, for the AMPK α , a mechanism is delineated where one of the S/T residues adjacent to the activating S/T is inhibitory (Djouder et al. 2010).

Of note, PDIK1L was more sensitive to mutations in the residues of its activation segment though only S194 exhibited a clear pattern of being a potential site for the inhibitory phosphorylation (Fig. 32C). Phyre2 predictions of three-dimensional structures for STK35 and PDIK1L sequences based on homology modeling (Kelley et al. 2015) revealed

that STK35, unlike PDIK1L, does not have a prospective β -turn in its activation loop (Supplemental Fig. S3). While there are conceivable limits to the accuracy of homology modeling for protein structures, if this prediction holds true in structural studies, it could explain why the STK35 structure would allow for more variation in the composition of its activation segment. In contrast, any perturbations around the activation loop of PDIK1L could peel its β -turn apart, distorting the kinase structure and disrupting its activity (Nolen et al. 2004).

A lot less is known about inhibitory phosphorylation at the activating segment. For example, S/T residues at +4 position to the activating S/T residue were described as likely inhibitory in several studies. However, this position was also extremely sensitive to any substitutions, making it difficult to study via phosphomimetic/phosphomutant types of experiments (Timm et al. 2008; McCartney et al. 2016). For STK35/PDIK1L kinases, we did not confirm that analogous T221 in PDIK1L and S418 in STK35 residues were likely to act as sites for inhibitory phosphorylation (Fig. 32C).

The conserved DFG motif of the activation loop in an active kinase conformation binds to a magnesium ion that interacts directly with ATP (usually reflected in the “DFG-in” structures of the kinases) (Modi and Dunbrack 2019; Xie et al. 2020). In an inactive conformation, the activation loop is often collapsed onto the surface of the kinase, blocking substrate binding, with the “DFG-out” position for the DFG motif also incompatible with binding magnesium ion and ATP. The structural differences in “DFG-in” and “DFG-out” states serve as the basis for developing specific kinase inhibitors, with type I inhibitors often targeting the active, “DFG-in” conformations, and type II – “DFG-out” state. While it was known that the DFG+2 serine residue was critical for the correct orientation of the “DFG-in” state in the active kinase via a 3-turn and hydrogen bonds (Kornev et al. 2006; Xie et al. 2013), our work for the first time to our knowledge provides genetic evidence that this

residue serves as a site for inhibitory phosphorylation.

We hypothesize that incorporation of phosphorylation at the DFG+2 serine residue of STK35/PDIK1L locks the activation loop in an inactive conformation that is incompatible with the precise ATP-magnesium coordination (Fig. 37). However, this can only be determined by further structural studies of STK35/PDIK1L alone or in the complex with SCP4. Additionally, *in vitro* measurements of kinase activity with or without inhibitory phosphorylation would be essential evidence. One of the main caveats of most *in vitro* approaches is that the kinase that phosphorylates these residues on STK35/PDIK1L is unknown. Nevertheless, based on the publicly available phospho-proteomics data, we could speculate that the large portion of STK35/PDIK1L kinases in the cell are inactive and require SCP4 for their activation.

3.5. AML metabolism

One of the critical questions that our work proposes is what are the immediate effectors of the SCP4/kinases phospho-signaling cascade. Identifying substrates of SCP4/kinases complex could potentially allow us to explain why AML cell lines show varying degrees of fitness defects upon SCP4 knockout or STK35/PDIK1L double knockout. While we ruled out the most apparent genetic explanations (such as differential gene expression, correlation with other known dependencies, cytogenetic and mutational background), there are many more potential causes for non-oncogene dependencies. For example, DNA hypermethylation and transcriptional silencing in AML can result in distinct synthetic lethal vulnerabilities (see Introduction §1.4.1.1).

The neonatal viability of SCP4^{-/-} mice can be extended from 20 to 42 hours by injecting glucose, which suggests a fundamental role of SCP4 in regulating metabolic pathways *in vivo* (Cao et al. 2018). In this prior study, SCP4 was found to positively regulate

FoxO1/3a transcriptional activity in hepatocytes, which promoted gluconeogenesis (for example, through the upregulation of phosphoenolpyruvate carboxykinase *PCK1* and glucose-6-phosphatase *G6PC*). While AML cells do not express *PCK1* or *G6PC* at high levels, we did observe some downregulation in isoforms *PCK2* and *G6PC3* upon SCP4 knockout in MOLM-13 (Supplemental Fig. S4).

Global transcriptional changes upon acute SCP4 knockout and STK35/PDIK1L double knockout showed consistent downregulation of multiple genes involved in amino acid biosynthesis and transport (Fig. 34B,C). We confirmed that most amino acid levels were indeed depleted upon SCP4 knockout and STK35/PDIK1L double knockout (Fig. 35). This nutrient deficiency was supported by the concurrent upregulation of genes involved in autophagy and catabolic pathways (Fig. 34B).

AML cells drive complex metabolic reprogramming to rapidly generate energy and biosynthesis intermediates and hijack the nutrient-sensing signaling pathways operating in these pathological conditions (Castro et al. 2019). Amino acids are essential for this rewiring since they can be used as biosynthetic building blocks and as carbon sources for energy production, control mitochondrial OXPHOS, protect cells against ROS, and regulate major nutrient sensors such as mTOR (Kreitz et al. 2019; see Introduction §1.4.1.2).

A 2018 study from Pereira et al. showed that different AML cell lines have different metabolic profiles (Pereira et al. 2018). Applying similar methods to the cell lines used in our study could be potentially informative if the differential dependency on SCP4 and STK35/PDIK1L reflected the baseline relationship between the ratio of glycolytic and oxidative metabolism in these cell lines. The hypothesis would be that the cells heavily relying on OXPHOS for energy production would be more sensitive to SCP4/kinases complex perturbation since they are more dependent on high amino acid supplies to support their TCA cycle. In our metabolomics data, we contrasted the downregulation of TCA cycle

metabolites with a relatively unaffected urea cycle (Fig. 36).

Another potential avenue for further research would be to investigate the mTOR kinase's biochemical response to the loss of SCP4/kinases since mTOR is a central node in nutrient sensing and transcriptional outputs of metabolic distress (Aramburu et al. 2014). mTORC1 also actively inhibits autophagy; hence upon amino acid starvation, autophagy is activated (Rothe et al. 2019). Watson et al. showed that autophagy induction limits glycolytic metabolism in the AML context (Watson et al. 2015). Additionally, many of the downregulated genes overlapping between SCP4 knockout and STK35/PDIK1L double knockout that are involved in amino acid biosynthesis and transport are some of the most sensitive targets of the stress-induced transcription factor ATF4 (Seo et al. 2009; Torrence et al. 2021). Characterizing the state of these proteins upon SCP4/kinases complex perturbation is essential to gain more insight into the nature of the following metabolic defects experienced by the cells.

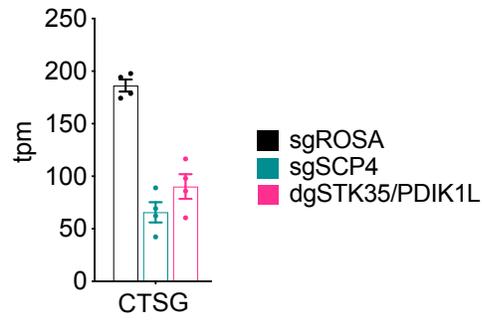


Figure S1. Transcriptional downregulation of the myeloid protease cathepsin G in response to SCP4 knockout and STK35/PDIK1L double knockout.

RNA-Seq transcripts per million (tpm) data for *CTSG* from MOLM-13 cells on day 5 post-infection with the indicated sgRNAs.

Plotted are data from each individual RNA-Seq replicate, the mean \pm SEM.

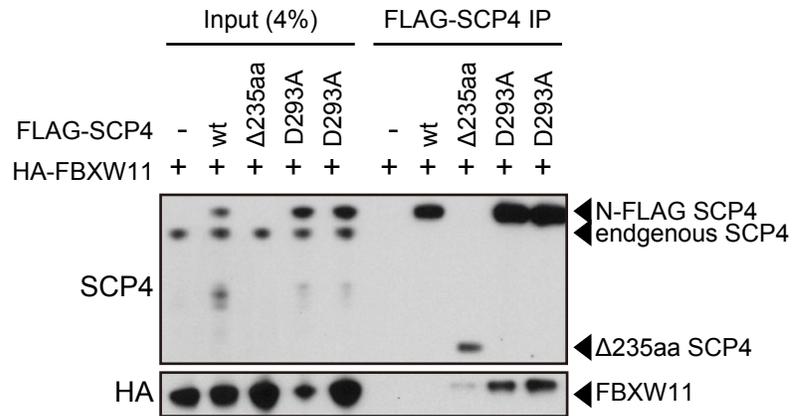


Figure S2. Catalytic mutants of SCP4, but not the wild-type SCP4, interact with FBXW11.

Immunoprecipitation followed by Western blotting performed with the indicated antibodies. The nuclear lysates were prepared from the human MOLM-13 cells stably expressing empty vector (empty), FLAG-SCP4^{wt} (wt), FLAG-SCP4²³⁶⁻⁴⁶⁶ (236-466), catalytic mutant FLAG-SCP4^{D293A} (D293A).

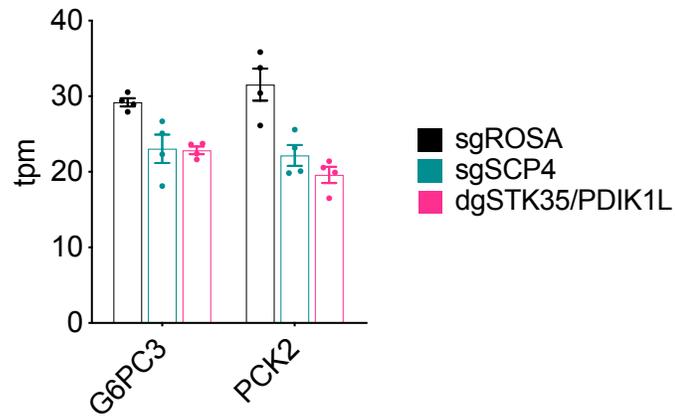


Figure S4. Gluconeogenesis genes upon SCP4 knockout and STK35/PDIK1L double knockout.

RNA-Seq transcripts per million (tpm) data for *G6PC3* and *PCK2* from MOLM-13 cells on day 5 post-infection with the indicated sgRNAs.

Plotted are data from each individual RNA-Seq replicate, the mean \pm SEM.

4. FUTURE DIRECTIONS

4.1. Immediate degradation

One of the major developments for understanding the functions of SCP4 and STK35/PDIK1L in the cell would be adopting and optimizing their specific and acute degradation strategy. An important caveat of depletion of SCP4 and STK35/PDIK1L at the DNA level that we have performed throughout this study via CRISPR-mediated knockout is the 5 days lag period between the infection of the cells with the sgRNAs and the sample collection, for example, for RNA-Seq or metabolomics experiments. Immediate SCP4 and STK35/PDIK1L protein degradation would allow us to discriminate between the direct consequences of their depletion and any confounding secondary effects.

There are a few strategies that we could try and implement for our model systems, such as fusing constructs of interest to (1) a destabilizing FKBP12 mutant (Banaszynski et al. 2006); (2) an auxin-inducible degron (AID) tag (Natsume et al. 2016); (3) an FKBP12^{F36V} mutant that could be degraded via cell-permeable degraders called dTAGs (Nabet et al. 2018). Destabilizing FKBP12 mutant fusion protein will be effectively degraded at baseline without the addition of stability-inducing ligand (conditional activation system). The other two approaches allow for conditional degradation of the protein of interest. Some of the caveats for all systems include cell line-specific degradation kinetics, incomplete target degradation, unconditional degradation due to ligand analogs in the media, potential ligand off-target effects, the requirement for co-expression of additional elements of the degradation system (for the AID approach).

Ideally, we would knock-in the degradation system into the endogenous loci of the proteins of interest, for example, the N-terminus of SCP4, but initial validation of either approach could be performed with the exogenous expression of a construct and stable CRISPR-mediated knockout of the endogenous gene.

4.2. The role of SCP4 on chromatin

Several studies, including ours, pointed out the remarkable affinity of SCP4 towards chromatin fraction. One possibility is that the biochemical properties of SCP4 mediate this association since chromatin is a highly charged biopolymer and upon cell lysis might bind proteins through nonspecific electrostatic interactions (Kustatscher et al. 2014). Interestingly, unlike endogenous protein, FLAG-tagged SCP4 in our study was equally distributed between the nuclear soluble and chromatin fractions (Fig. 17A,B). This can indicate that FLAG-tag somehow disrupts SCP4 loading on chromatin or that there is a competition between the exogenous and endogenous SCP4. A clarification experiment in this regard that could also point towards the specificity of SCP4 function on chromatin would be observing FLAG-SCP4 preferential chromatin localization upon the endogenous SCP4 knockout.

Another necessary experiment is RNase A treatment of the chromatin fraction prior to chromatin solubilization that would help us to determine whether SCP4 is anchored or bridged to chromatin via RNA (Hiragami-Hamada et al. 2018).

Additional evidence could come from immunostaining experiments with specific antibodies against endogenous SCP4 and histone marks for transcriptional repression or activation. In this way, we could clarify whether SCP4 association with heterochromatin, described in the literature, is true for our model system (Wani et al. 2016).

Our preliminary results suggested that at least HA-PDIK1L might be enriched in the chromatin fraction (data not shown) and less present in the other fractions with an obvious caveat of being exogenously expressed. These data were encouraging considering the tight association between PDIK1L and SCP4 and could suggest a model in which SCP4 attracts PDIK1L to chromatin for performing its downstream functions. Of note, HA-PDIK1L degradation in all fractions upon endogenous SCP4 knockout did not allow us to observe any relocalization of HA-PDIK1L. Going forward, the development of specific immune

reagents against endogenous proteins is crucial for providing an additional characterization of the SCP4-STK35/PDIK1L complex and its association to chromatin.

Given SCP4 affinity with chromatin is functional, we could attempt to optimize ChIP-Seq protocols to preserve endogenous SCP4 via trouble-shooting fractionation, chromatin shearing, and triple knockout of myeloid proteases, if necessary. Depending on the mode of chromatin association, SCP4 ChIP-Seq could benefit from two-step crosslinking methods (Tian et al. 2012). Alternatives to ChIP-Seq could include Cut&Run or Cut&Tag, however, more evidence is needed to assess the utility of these methods for the proteins not directly associated with DNA or for the heterochromatin regions.

The development of protein depletion techniques from §4.1 would also allow us to perform ChIP-Seq with antibodies against histone epigenetic modifications of interest or Pol II upon acute SCP4 degradation. This could be informative in nominating immediate consequences of SCP4 loss on transcription factors and epigenetic modifiers or global transcription. Another experiment that could potentially help us to identify genes specifically affected by the acute SCP4 depletion would be ATAC-seq.

4.3. Understanding the phenotype

Competition-based proliferation assays are a powerful tool to detect cell fitness defects upon sgRNA mediated knockout of a gene of interest. However, additional experiments are required to understand the nature of the deleterious phenotype. Our cell cycle analysis upon SCP4 knockout and STK35/PDIK1L double knockout and apoptosis measurements showed both evidence for an increase in G1/G0 arrest and apoptosis (Fig. 8; Fig. 25). These effects were measured on day 5 post-infection with sgRNAs, and immediate protein degradation could help us distinguish primary from secondary effects of the essential genes knockout.

Various techniques have been developed to delineate the precise molecular mode of cell death (Galluzzi et al. 2018). As such, blocking the key cell death mediators via genetic means or with specific pharmacological inhibitors could be informative if the full or partial rescue of SCP4/kinases knockout phenotype was to be observed. Optimizing immunostaining methods for SCP4/kinases would also benefit our understanding of the knockout phenotype through various imaging and cytochemical techniques available for delineating the precise cell death phenotype (Taatjes et al. 2008).

Another interesting question would be the relationship of SCP4/kinases complex with the cell cycle progression. The publicly available resource Cyclebase does not report any evidence for transcriptional regulation for STK35/PDIK1L dependency on the cell cycle, with their mRNAs expression measured in the high-throughput studies being non-periodic, unlike for the cell-cycle related proteins PCNA or AURKA (Santos et al. 2015). However, there is no expression data for SCP4 and no proteomics data available for either protein in relation to cell cycle measurements.

While we have not observed any effects of SCP4 knockout on differentiation in human HSPCs, the top upregulated signature in our RNA-Seq studies was myeloid leukocyte activation (Fig. 11-13; Fig. 34). We could combine immediate SCP4/kinases depletion with immune activation markers flow-cytometry panels to gain further insights into this phenotype and any changes in cell morphology. Additionally, conditioned media experiments could be informative if SCP4/kinases loss phenotype induces some paracrine signaling in addition to its cell-autonomous effects.

We also have to clarify the effect that SCP4 has on the stability of PDIK1L/STK35. Appreciating the full extent of this phenotype is difficult in the absence of specific antibodies against the endogenous proteins. Additional experiments using inhibitors of the ubiquitin-proteasome or lysosomal proteolysis pathways are required to differentiate

between these modes of protein degradation for cytoplasmic SCP4 and exogenous or endogenous STK35/PDIK1L upon the endogenous SCP4 knockout. The association with Hsp70 that we noticed in our MS data could potentially target proteins towards the autophagy pathway (Fernandez-Fernandez et al. 2017).

4.4. Identifying and validating SCP4 substrates

We have nominated STK35/PDIK1L as both interacting partners and potential substrates of SCP4 due to their stable interaction and the body of research on analogous phosphatase/kinase complexes (see Discussion §3.3–3.4). However, the question remains of whether these kinases are the only relevant substrates of SCP4 in AML and what are the immediate downstream effectors of the complex.

Like ours, many studies started research into potential substrates of a given phosphatase through detecting protein-protein interactions in hypothesis-driven functional studies or through unbiased interactome analysis. For example, SCP1 over the last decade has been found to dephosphorylate and regulate the repressor element 1 (RE1) silencing transcription factor (REST) (Yeo et al. 2005; Nesti et al. 2014; Burkholder et al. 2018). In other cases, enzyme-substrate complexes tend to be transient and make substrate identification difficult (Fahs et al. 2016). One way to overcome this for protein tyrosine phosphatases has been the discovery of “substrate-trapping” mutants (Flint et al. 1997; Blanchetot et al. 2005). Nonetheless, “substrate-trapping” techniques for other families of phosphatases, including HAD-phosphatases, are yet to be described in the literature.

Another standard method of discovering and validating physiological substrates of phosphatases is expression screening. It is not unbiased in the sense that specific antibodies against both phosphorylated residues and total protein levels of a prospective substrate have to be established. The following exogenous expression of a phosphatase versus its

catalytic mutant and the associated decrease in the specific phosphorylation versus total protein levels could indicate the potential enzyme-substrate relationship between them. Along these lines, antibodies against pS194 in PDIK1L and pS385 in STK35 are required to provide more evidence of these kinases being physiological SCP4 substrates in the cell. A more comprehensive way to look for additional substrates would be differential phosphoproteomics coupled with a labeling technique such as stable isotope labeling by amino acids in cell culture (SILAC) and recent advances in the differential tandem mass spectrometry proteomics, such as SAMPEI (Cifani et al. 2021), upon SCP4 acute depletion or the addition of the recombinant catalytically active versus mutant SCP4 into the cell lysates.

4.5. Identifying kinases substrates

Given our focus on the SCP4/kinases complex signaling outcome, a valid research path would be focusing on the substrates of the kinases. Similarly to identifying phosphatase substrates, the approaches here are limited to mapping kinase interactions and the indirect phosphoproteomics-based strategies (Jurcik et al. 2020). The caveats are also the same: transient enzyme-substrate interactions, high noise, secondary effects associated with a gain or a loss of a phospho-signaling enzyme. For all of these methods, it is essential to compare the kinases with their catalytically dead mutants and versions with phosphomutant or phosphomimetic residues at the sites of inhibitory phosphorylation we discovered.

An alternative strategy recently described for kinase-substrates identification is employing their analog-sensitive (AS) mutants. AS mutants harbor several mutations in their ATP-binding pocket that allow them to accept in addition to ATP bulky ATP analogs, such as ATP with a γ -thiophosphate (Michowski et al. 2020). This approach could help us accomplish two goals: (1) identify direct kinase substrates that would be uniquely labeled

with a thiophosphate (labeling); and (2) selectively block the mutant kinase with the available specific bulky inhibitors (inhibition).

Additionally, it would be interesting to compare the outputs for either substrate nomination approach upon disrupting the interaction between SCP4 and STK35/PDIK1L (see below).

4.6. Structural and biochemical studies of purified proteins

One of the main downstream developments for our research would be optimizing expression and purification strategies for STK35/PDIK1L alone or in the complex with SCP4. First, it is important to show that SCP4 dephosphorylates full-length STK35/PDIK1L *in vitro*. Among the challenges of this experiment would be making sure that the kinases were phosphorylated on the residues of interest. We can try purification strategies from insect or mammalian cell cultures to preserve naturally occurring phosphorylation. Again, phosphorylation-specific antibodies would be required. In our study, SCP4 additionally efficiently dephosphorylated pS216 residue that did not show a differential phenotype characteristic for potentially inhibitory phosphorylation upon phosphomimetic/phosphomutant substitution experiments (Fig. 32, 33). It remains to be seen whether SCP4 can also dephosphorylate this residue in the context of fully folded kinase substrates.

SCP4/kinases complex purification is necessary to determine its stoichiometric composition. The analytical gel filtration step during purification would provide additional evidence for the integrity of this phosphatase/kinase complex. The purified complex would also allow us to map the interaction surfaces between them, for example using hydrogen/deuterium exchange mass spectrometry. These data would allow us to design additional experiments involving interaction surface perturbations through genetic mutations and rationally designed peptides.

The co-IP experiments for SCP4 with different kinase mutants (catalytically dead, phosphomimetic/phosphomutant substitutions) would be informative for the model refinement and proposed structural studies. For example, any of the phosphomimetic/phosphomutant kinase versions we used could have weakened or abolished the phosphatase/kinase association. This would require us to reevaluate the original conclusions from these experiments. Finally, structural studies of the purified SCP4 in the complex with STK35/PDIK1L wild type and with phosphomimetic/phosphomutant substitutions (as long as they do not disrupt the complex formation in co-IP studies) would be instrumental in supporting the model we propose. This could be the first tangible way to obtain the structures for these proteins, given that the previous attempts in SCP4 crystallization have been unsuccessful (Zhang, personal communication). We would also be able to see the position of the DFG-loop relative to the SCP4 catalytic site for the active versus inactive kinase conformations, which would provide additional insight into the physiological regulation of eukaryotic kinases.

4.7. Therapeutic implications

All the experiments described above are important for deeper mechanistic insights into the role of SCP4/kinases complex in AML cells. It is likely that understanding the immediate effects of the complex perturbation, nominating downstream substrates, and characterizing the precise cell fitness defects will provide enough evidence to support or deny our current hypothesis that SCP4/kinases complex is implicated in supporting AML cell metabolism. The metabolomics MS analysis specifically requires acute protein degradation and careful labeling techniques to assess precise metabolic defects and their extent accurately. Additional validation for this hypothesis could come from a partial rescue of the SCP4/kinases loss phenotype upon adding the critical metabolites into the media.

Our findings that SCP4 is not required for normal HSPCs differentiation and thus could serve as a therapeutic opportunity in human AML were further supported by the published SCP4 knockout mouse model phenotype (Cao et al. 2018). However, Cao et al. did not methodically characterize the abnormalities in the embryonic development of homozygous SCP4^{-/-} mice besides neonatal lethality and the glucose sensitivity. The mouse studies are a crucial next step for understanding the role of SCP4. Accurately describing the systematic SCP4 deficiency, specifically in the context of hematopoietic development, is essential for establishing its potential functional role in AML. Conditional knockout animal models are required, possibly with the simultaneous introduction of AML drivers, such as MLL-AF9, to characterize the AML onset and phenotypes. For the potential therapeutic implications, conditional knockdown in mice is required to access the target therapeutic potential. In this regard, our data concerning the mouse AML xenograft model are encouraging (Fig. 9).

Several HAD family phosphatases have been successfully targeted using small-molecule inhibitors (Freschauf et al. 2009; Zhang et al. 2011; Krueger et al. 2013; Krueger et al. 2014; Wang et al. 2016b). If the structural studies we propose are successful, virtual screening approaches could be exploited similarly to those described for the evolutionary close SCP1 (Park et al. 2017). Another option to explore would be developing STK35/PDIK1L dual inhibitor using small molecules that compete with ATP binding to these targets. Mapping the interaction surface between SCP4 and the kinases would allow us to design allosteric peptide inhibition disrupting their interaction.

Overall, our study justifies the development of such compounds targeting this phospho-catalytic signaling complex, which would be expected to suppress AML cells whereas having minimal effects on normal hematopoiesis, yield additional mechanistic insights into the SCP4/kinases cellular functions, and enrich our understanding of AML.

5. MATERIALS AND METHODS

5.1. Cell lines

Human cell lines MOLM-13, NOMO-1, MV4-11, ML-2, HEL, SET-2, THP1, U937 (acute myeloid leukemia, AML); K562 (chronic myelogenous leukemia, CML); PANC-1, MIAPaCa-2, SUIT-2, AsPC-1 (pancreatic ductal adenocarcinoma, PDAC); RH4, RH30 (rhabdomyosarcoma, RMS); NCI-H1048 (small cell lung cancer, SCLC); and murine cell line RN2 (MLL-AF9/NRas^{G12D} AML) (Zuber et al. 2011a) were cultured in RPMI-1640 supplemented with 10% fetal bovine serum (FBS). Human cell line SEM (AML) was cultured in IMDM with 10% FBS. Human cell line OCI-AML3 (AML) was cultured in Alpha MEM with 20% FBS. Transformed human HSPC MLL-AF9 (MA9) cells (Mulloy et al. 2008; Wei et al. 2008) derivatives MA9-NRAS^{G12D} and MA9-FLT3^{ITD} (AML) were cultured in IMDM with 20% FBS. Human cell line Kasumi-1 (AML) was cultured in RPMI with 20% FBS. Human cell lines RD (RMS), A549 (lung adenocarcinoma), and HEK293T were cultured in DMEM with 10% FBS. Murine cell line NIH3T3 was cultured in DMEM with 10% bovine calf serum (BCS). 1% penicillin/streptomycin was added to all media. All cell lines were cultured at 37°C with 5% CO₂ and were regularly tested mycoplasma negative.

5.2. Plasmid construction and sgRNA cloning

The sgRNA lentiviral expression vector with optimized sgRNA scaffold backbone (LRG2.1, Addgene plasmid # 108098) was derived from a lentiviral U6-sgRNA-EFS-GFP expression vector (LRG, Addgene plasmid # 65656) by replacing the original wild-type sgRNA scaffold with our optimized version (Tarumoto et al. 2018). LRCherry2.1 was produced from LRG2.1 by replacing GFP with mCherry coding sequence. LRG2.1-Blast was generated by the introduction of the blasticidin resistance gene after GFP coding sequence

with a P2A linker. In this study, LRG2.1, LRG2.1-Blast, or LRCherry2.1 were used for introducing sgRNA into human cell lines, whereas the original LRG vector was used for sgRNA expression in murine cell lines. All the sgRNAs were cloned into the LRx2.1-derivatives or LRG vectors using a BsmBI restriction site. Sequences of all sgRNAs used in this study are provided in the Appendix 6.

LentiV_Neo vector was derived from LentiV_Cas9_puro vector (Addgene plasmid # 108100) by removing Cas9 and replacing a puromycin resistance gene with a neomycin resistance gene. *CTDSPL2* cDNA (Horizon Discovery, Clone ID: 5744745), partial *STK35* cDNA (Horizon Discovery, Clone ID: 9021751), and *PDIK1L* cDNA (Horizon Discovery, Clone ID: 4828997) were cloned into LentiV_Neo vector using In-Fusion cloning system (Takara Bio; Cat. No. 121416). The N-terminus of *STK35* cDNA (4-296 bp from translation start site) was added to obtain *STK35* cDNA corresponding to NM_080836. The N-terminal FLAG and HA-tags were added to the cDNAs in-frame. For the generation of all the mutants, the base substitutions were introduced into the cDNA via PCR mutagenesis. The list of all the primers used for the generation of mutants used in this study can be found in the Appendix 7.

5.3. Construction of sgRNA libraries

All the genes contained within the gene group “Phosphatases” reported by the HUGO Gene Nomenclature Committee (HGNC) were selected for designing the phosphatase domain-focused sgRNA library. The phosphatase domain information was retrieved from the NCBI Conserved Domains Database. 6 to 30 sgRNAs were designed against exons encoding phosphatase domains. For the CRISPR exon-scan library 85 sgRNAs were selected against every exon of *SCP4*. All sgRNAs were designed using <https://mojica.cshl.edu> and filtered for the minimal predicted off-target effects (Hsu et al. 2013).

The domain targeting and positive/negative control sgRNAs oligonucleotides were synthesized in a pooled format on an array platform (Twist Bioscience) and then amplified by PCR, using Phusion® Hot Start Flex DNA Polymerase (NEB; Cat. No. M0535). The PCR products were cloned into the BsmBI-digested lentiviral vector LRG2.1 (Addgene plasmid # 108098) using a Gibson Assembly® Cloning Kit (New England BioLabs; Cat. No. E2611). The library was produced in MegaX DH10B T1R Electrocomp Cells (Invitrogen; Cat. No. C640003).

5.4. Lentivirus production and infection

The lentivirus was packed in HEK293T cells by transfecting lentiviral expression vector plasmids together with the lentiviral packaging plasmid (psPAX2, Addgene plasmid # 12260) and the envelope protein expressing plasmid (VSV-G) using polyethylenimine (PEI 25K™; Polysciences; Cat. No. 23966) transfection reagent. HEK293T cells were transfected when ~80–90% confluent.

For the pooled sgRNA phosphatase library lentivirus production, five 10 cm plates of HEK293T were used. For one 10 cm plate of HEK293T cells, 5 µg of plasmid DNA, 5 µg of VSV-G, 7.5 µg psPAX2, and 64 µL of 1 mg/mL PEI were mixed with 1 mL of Opti-MEM®, incubated at room temperature for 20 minutes, and then added to the cells. The media was changed to 5 mL of fresh media 6–8 hours post-transfection. The lentivirus-containing media was collected at 24, 48, and 72 hours post-transfection, pooled together and filtered through a 0.45 µM non-pyrogenic filter.

For lentivirus infection, target cells were mixed with the virus (volume empirically determined to result in the desired infection rate) and 4 µg/ml polybrene and then centrifuged at 600 x g for 30–45 minutes at room temperature. For adherent cell lines, media was changed at 24 hours post-infection. If selection was required for stable cell line

establishment, corresponding antibiotics (1 µg/mL puromycin, 1 µg/mL blasticidin, 1 mg/mL G418) were added 24 hours post-infection.

5.5. Pooled negative-selection CRISPR screening and data analysis

Cas9-expressing stable cell lines were produced by infection with the lentivirus encoding human codon-optimized *Streptococcus pyogenes* Cas9 protein (LentiV_Cas9_puro, Addgene plasmid # 108100) and subsequent selection with 1 µg/mL puromycin. Lentivirus of pooled sgRNA phosphatase library titer was estimated by mixing the cells with the serially diluted virus and measuring GFP% on day three post-infection using a Guava® easyCyte™ Flow Cytometer (Merck Millipore). The cells were then infected with the volume of virus estimated to result in a 30–40% infection rate to increase the probability of a single sgRNA introduction event per cell. The number of cells was kept 1000 times more than the sgRNA number in the library to maintain the representation of sgRNAs during the screen. Cells were harvested at initial (day three post-infection) and final (10 or more population doublings after the initial) time points. Genomic DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN; Cat. No. 51304).

Sequencing libraries were prepared as described previously (Shi et al. 2015). Briefly, genomic DNA fragments (~200 bp) containing sgRNAs were amplified by PCR, using Phusion® Hot Start Flex DNA Polymerase (NEB; Cat. No. M0535). The PCR products were end-repaired with T4 DNA polymerase (NEB; Cat. No. B02025), DNA Polymerase I, Large (Klenow) fragment (NEB; Cat. No. M0210L), and T4 polynucleotide kinase (NEB; Cat. No. M0201L). The 3' adenine overhangs were added to the blunt-end DNA fragments by Klenow Fragment (3'-5' exo; NEB; Cat. No. M0212L). The DNA fragments were then ligated with diversity-increased custom barcodes (Shi et al. 2015), using Quick Ligation Kit (NEB; Cat. No. M2200L). The ligated DNA was PCR amplified with primers containing Illumina

paired-end sequencing adaptors, using Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher; Cat. No. F5485). The final libraries were quantified using Bioanalyzer Agilent DNA 1000 (Agilent 5067-1504) and were pooled together in an equimolar ratio for paired-end sequencing using MiSeq platform (Illumina) with MiSeq Reagent Kit V3 150-cycle (Illumina).

The sgRNA sequences were mapped to the reference sgRNA library to discard any mismatched sgRNA sequences. The read counts were calculated for each individual sgRNA. The following analysis was performed with a custom Python script: sgRNAs with read counts less than 50 in the initial time point were discarded; the total read counts were normalized between samples; the average \log_2 fold-change in the reads corresponding to sgRNAs targeting a given gene (CRISPR score) was calculated, as described (Wang et al. 2015). AML-specific dependency was determined by subtracting the average of CRISPR scores in non-AML cell lines from the average of CRISPR scores in AML cell lines, and that score was ranked in ascending order. The phosphatase CRISPR screening data is provided in the Appendix 2. The SCP4 CRISPR exon scanning screening data is provided in the Appendix 3.

5.6. Competition-based cell proliferation assays

Cas9-expressing cell lines were infected with lentivirus encoding sgRNAs linked with either GFP or mCherry reporters at the infection rate of ~20–60% as described above. Percentage of GFP- or mCherry-positive cells was measured every three days from day 3 to day 18 post-infection using Guava® easyCyte™ Flow Cytometer (Merck Millipore). The percentage of fluorescent cells in the population on each day was divided by their percentage on day three to calculate fold-change corresponding to the effect of a given sgRNA on cell proliferation.

5.7. Western Blot

For knockout experiments, cells were harvested on day 5 post-infection with sgRNA.

For whole-cell lysates, a precise cell number was calculated before sample preparation for equal loading. Cells were washed 1X with ice-cold PBS. Cell pellets were then resuspended in PBS and 2X Laemmli Sample Buffer (Bio-Rad, Cat. No. 1610737) containing β -mercaptoethanol (BME), boiled for 30 minutes, and cleared by centrifugation at room temperature at 13,000 rpm.

For fractionation experiments, cell fractions were prepared as described below. The exact protein concentrations in the fractions were measured with the Pierce™ BCA Protein Assay Kit (Thermo Fischer, Cat. No. 23225). The cell fractionation samples were boiled for 5–10 minutes with a 2X Laemmli Sample Buffer containing BME.

The samples were separated by SDS-PAGE (NuPAGE 4-12% Bis-Tris Protein Gels, Thermofisher), followed by transfer to nitrocellulose membrane via wet transfer at 90 V for 2–2.5 hours. Membranes were then blocked for 1 hour with 5% non-fat milk in TBST and incubated with primary antibodies overnight. On the next day, the membranes were washed 3X with TBST followed by incubation with HRP-conjugated secondary antibodies (Rabbit Cytiva/Amersham NA934; Mouse Agilent/Dako P026002-2) for one hour at room temperature. After 3X washes with TBST, an enhanced chemiluminescence solution containing HRP substrate was added to the membranes, and the signal was visualized using the darkroom development techniques for chemiluminescence. Antibodies used in this study included SCP4 (*CTDSPL2*) (CST, # 6932, 1:500), FLAG (Sigma Aldrich, F1804, 1:10,000), HA-HRP (Sigma Aldrich, clone 3F10, 1:10,000), H3K4me3 (Sigma Aldrich, 07-473, 1:1,000), β -Actin-HRP (Sigma A3854-200UL; 1:50,000).

5.8. Cell cycle arrest and apoptosis analysis

Cell cycle analysis was performed using the BrdU Flow Kit according to the manufacturer protocol (BD, FITC BrdU Flow Kit; Cat. No. 559619), with cells pulsed with BrdU for 1 hour at 37°C. Annexin V apoptosis staining was performed using the Apoptosis Detection Kit according to the manufacturer protocol (BD, FITC Annexin V Apoptosis Detection Kit; Cat. No. 556547). Cells were co-stained with 4',6-diamidino-2-phenylindole (DAPI) to measure DNA content and analyzed with a BD LSRFortessa flow cytometer (BD Biosciences) and FlowJo software (TreeStar).

5.9. Editing and differentiation of human peripheral blood CD34⁺ cells

Circulating G-CSF-mobilized human CD34⁺ cells were obtained from two de-identified healthy donors (KeyBiologics). Enrichment of CD34⁺ cells was performed by immunomagnetic bead selection using an CliniMACS instrument (Miltenyi Biotec). After enrichment, the CD34⁺ cell fraction was 95% of the total, with <0.2% CD3⁺ cells and <2% CD19⁺ cells.

Purified recombinant Cas9 protein was obtained from Berkeley Macrolabs. Two sgRNAs against SCP4 and one sgRNA against MYC were synthesized by TriLink BioTechnologies. The sgRNAs included 2'-O-methyl 3'-phosphorothioate (MS) modifications at 3 terminal nucleotides at both the 5' and 3' ends (Hendel et al. 2015).

Editing in human CD34⁺ cells was performed as described before (Métais et al. 2019). Briefly, cryopreserved CD34⁺ cells were thawed and pre-stimulated for 48 hours in the maintenance medium (Appendix 8). The cells were then washed 1X in PBS, resuspended in T buffer included in kit (Thermo Fisher Scientific, MPK10025), mixed with RNP (described below), and electroporated with 1600 V, 3 pulses of 10 ms, using a Neon Transfection System (Thermo Fisher Scientific, Cat. No. MPK5000). As a control, cells were not electroporated or

electroporated with Cas9 and Non-targeting gRNA.

Ribonucleoprotein complexes (RNPs) were formed by incubating 160 pmol of Cas9 with 320 pmol sgRNA in 50 μ L of 10 mM HEPES (Thermo Fisher Scientific, catalog # 15630080), 150 mM NaCl (Thermo Fisher Scientific, catalog # 9759) for 35 minutes. 1 million CD34⁺ cells resuspended in 50 μ L of T buffer were mixed with RNP complex in a final volume of 100 μ L.

After electroporation, cells were immediately added to the media with cytokines required for differentiation into erythroid, myeloid, or megakaryocyte lineages or to methylcellulose for the colony forming cell (CFC) assay (see below). Erythroid, myeloid, and megakaryocyte maturation were monitored by immune-flow cytometry for the appropriate cell surface markers (Appendix 8) on days 2, 5, 8, and 16 post-electroporation (final day for differentiation) using BD LSRFortessa Dual™ Flow Cytometer. The genomic DNA from each of the lineages was collected in parallel to the immune-flow cytometry, and Western blot samples were prepared on days 5 and 8.

The on-target indel frequencies were quantified using QuantStudio™ 12K Flex Real-time PCR System (Thermo Fisher Scientific) with Power SYBR Green Master Mix (Thermo Fisher). Two primer pairs were developed for each targeted region using Primer3 (Untergasser et al. 2012). One primer in each pair was chosen to span the Cas9 cleavage site 3-nt upstream of the PAM site, and the second primer was selected at a 100–150 bp distance. Primer sequences are provided in Appendix 8. Quantitative PCR was performed on genomic DNA from gene-edited cells or control cells. Based on the consistency of amplification, either chr10 or untargeted region of CTDSPL2 was used as an endogenous control for calculating Δ CT values. The percentage of editing in CD34⁺ cells electroporated with sgRNAs vs. Cas9 alone was approximated as $1 - 2^{(-\Delta\Delta CT)}$. For the myeloid lineage, Surveyor nuclease assay was performed with primers in Appendix 8, according to the

manufacturer protocol (Integrated DNA Technologies, Cat. No. 706020) (Appendix 8. sgRNAs; media and cytokines; and antibodies used in flow cytometry panels in CD34⁺ experiments).

For the CFC assay, ~800 CD34⁺ gene-edited and control cells were mixed with 1.0 mL of MethoCult™ SF H4636 methylcellulose (Stemcell Technologies, Cat. No. 04636) containing recombinant cytokines for human embryonic stem (ES) cell-derived hematopoietic progenitor cells. The cultures were incubated in 35-mm tissue culture dishes. After 14 days, single colonies were classified and quantified.

All the reagents used in these experiments can be found in the Appendix 8.

5.10. *In vivo* transplantation of MOLM-13 cells into NSG mice

All animal procedures and studies were approved by the Cold Spring Harbor Laboratory Animal Care and Use Committee in accordance with the Institutional Animal Care and Use Committee. First, MOLM-13/Cas9⁺ cell line stably expressing luciferase was established via lentiviral infection with a Lenti-luciferase-P2A-Neo (Addgene # 105621) vector followed by G418 (1 mg/mL) selection. These cells were infected with lentivirus encoding GFP-linked sgRNAs either sgROSA (negative control) or sgSCP4. On day 3 post-infection, the infection rate was checked by the percentage of GFP positive cells, and all samples had over 90% infection rate. 0.5 million cells were transplanted into sublethally (2.5Gy) irradiated NSG mice (Jax 005557) through tail vein injection. Mice were imaged with IVIS Spectrum system (Caliper Life Sciences) on days 12 and 15 post-injection for visualizing the disease progression.

5.11. Sub-cellular fractionation and Immunoprecipitation in MOLM-13 cells

Subcellular protein fractionation was performed either by using Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Scientific; Cat. No. 78840) according to the manufacturer protocol or following the protocol below.

For subsequent mass spectrometry (MS) analysis, ~200 million cells were collected and the protocol below was used. For immunoprecipitation (IP) coupled with Western blot, 40-60 million cells were collected and the protocol below was used.

The cells were washed 2X with ice-cold PBS, then resuspended in 5X Packed Cell Volume (PCV) of Hypotonic cell lysis buffer (10 mM HEPES pH 7.9; 1.5 mM MgCl₂; 10 mM KCl; 1 mM DTT; supplemented with proteinase inhibitors). Note: SCP4 might be a target of myeloid proteases in MOLM-13 cells (data not shown) (Zhong et al. 2018). Therefore, we found that the best protein preservation was achieved when avoiding the use of detergents in the protocol and thoroughly washing the nuclei pellet from any traces of the cytoplasm.

After 15 minutes incubation at 4°C with rotation, the cells were pelleted and resuspended in 2X PCV of Hypotonic cell lysis buffer. The cell walls were disrupted using a glass tissue homogenizer on ice with 10 up-and-down strokes using a type B pestle. The disrupted cells in suspension were pelleted for 20 minutes at 11,000 x g at 4°C. The supernatant was saved as the *cytoplasmic fraction*. The pellet was thoroughly washed 4X with Hypotonic cell lysis buffer. The crude nuclei pellet was resuspended in 2/3X PCV of Extraction Buffer (20 mM HEPES pH 7.9; 1.5 mM MgCl₂; 420 mM NaCl; 25% Glycerol; 1 mM DTT; 1:1000 benzonase (Sigma E1014-25KU); supplemented with proteinase inhibitors). The nuclear lysates were incubated at room temperature for 1 hour on the rotor and centrifuged at 4°C for 10 minutes at 21,000 x g. The supernatant was diluted to 150 mM NaCl using No-salt dilution buffer (20 mM HEPES pH 7.9; 1.5 mM MgCl₂; 0.31 mM EDTA; 1 mM DTT; supplemented with proteinase inhibitors) and saved as the *nuclear fraction*.

The exact protein concentrations in the fractions were measured with the Pierce™ BCA Protein Assay Kit (Thermo Fischer, Cat. No. 23225).

For immunoprecipitation, NP-40 was added up to 0.25%, and the nuclear fractions were incubated with anti-FLAG M2 agarose (Sigma A2220-10ML) or anti-HA agarose (Sigma E6779-1ML) or Pierce Anti-HA Magnetic Beads (Thermofisher 88837) at 4°C overnight. Next day, the beads were washed 5X with Wash buffer (20 mM HEPES pH 7.9; 150 mM NaCl; 0.2 mM EDTA; 0.25% NP-40; 1 mM DTT; supplemented with proteinase inhibitors). For Western blot the beads were boiled for 10 minutes in 1X Laemmli Sample Buffer (Bio-Rad, Cat. No. 1610737) containing BME. For MS, the beads were send on ice to the Proteomics and Metabolomics Facility Center for Biotechnology at the University of Nebraska-Lincoln.

5.12. Protein identification by mass spectrometry (MS)

Bead samples were made up in 60µL 1x non-reducing LDS sample buffer and incubated at 95°C for 10 min prior to loading 50µL of sample onto a Bolt™ 12% Bis-Tris-Plus gel and running briefly into the top of the gel. The gel was fixed and stained with colloidal coomassie blue G250 stain. Gel containing the proteins was reduced and alkylated, then washed to remove SDS and stain before digestion with trypsin (500ng) overnight at 37°C. Peptides were extracted from the gel pieces, dried down, and samples were re-dissolved in 30µL, 2.5% acetonitrile, 0.1% formic acid. 5µL of each digest was run by nanoLC-MS/MS using a 2h gradient on a 0.075 mm x 250 mm CSH C18 column feeding into a Q-Exactive HF mass spectrometer.

All the samples were analyzed using Mascot (Matrix Science, London, UK; version 2.6.2). Mascot was set up to search the cRAP_20150130.fasta (123 entries); uniprot-human_20191024 database (selected for Homo sapiens, unknown version, 74,034 entries)

assuming the digestion enzyme trypsin. Mascot was searched with a fragment ion mass tolerance of 0.060 Da and a parent ion tolerance of 10.0 parts-per-million. Deamidation of asparagine and glutamine, oxidation of methionine, carbamidomethylation of cysteine, phosphorylation of serine, threonine and tyrosine were specified in Mascot as variable modifications.

Scaffold (version Scaffold_4.8.9, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 80.0% probability by the Peptide Prophet algorithm (Keller et al. 2002) with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al. 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters.

5.13. Co-IP in HEK293T cells

One 80% confluent 10-cm plate of 293T cells was transfected with 10 μ g HA-PDIK1L and 10 μ g of either Empty vector, or FLAG-SCP4, or FLAG-SCP4²³⁶⁻⁴⁶⁶, or FLAG-SCP4^{D293A}. Total 20 μ g DNA was mixed with 60 μ L of 1 mg/mL PEI and 1 mL of Opti-MEM®, incubated at room temperature for 15 minutes, and then added to the cells. 24 hours post-transfection the whole-cell lysates were prepared as follows.

Collected cells by incubation for 5 minutes with 2 mL Trypsin, then neutralized with 6 mL DMEM media. Washed cells 2X with ice-cold PBS. Resuspended cells in 500 μ L of ice-cold Cell Lysis Buffer (CST, #9803). Incubated on ice for 30 minutes. Centrifuged at 4°C for

10 minutes at 14,000 x g.

The exact protein concentrations in the fractions were measured with the Pierce™ BCA Protein Assay Kit (Thermo Fischer, Cat. No. 23225).

Whole-cell lysates were incubated with anti-FLAG M2 agarose (Sigma A2220-10ML) or anti-HA agarose (Sigma E6779-1ML). Next day, the beads were washed 5X with the Cell Lysis Buffer. Afterwards, the beads were boiled for 10 minutes in 1X Laemmli Sample Buffer (Bio-Rad, Cat. No. 1610737) containing BME, followed by Western blot analysis.

5.14. RNA-Seq

For knockout experiments, cells were harvested on day 5 post-infection with sgRNA. RNA was purified with the guanidinium thiocyanate-phenol-chloroform extraction method. Briefly, cells were washed 1X with PBS and lysed with 1 mL of TRIzol (Thermo Scientific; Cat. No. 15596018). The samples were vigorously mixed with 200 μ L chloroform and incubated for 3 minutes at room temperature, followed by centrifugation at 10,000 x g for 15 minutes at 4 °C. The aqueous phase containing RNA was added to the equal volume of isopropanol, and RNA was precipitated after incubation of 10 minutes at room temperature. Precipitated RNA was then pelleted by centrifugation at 10,000 x g for 10 minutes at 4 °C and washed 1X with 1 mL of 75% ethanol. After air-drying for 10 minutes, RNA was resuspended in 20 μ L of DEPC-treated (RNase-free) water.

RNA-seq libraries were prepared using TruSeq sample prep kit v2 (Illumina) according to the manufacturer protocol. Briefly, polyA RNA was enriched with oligo-dT beads and fragmented enzymatically. The cleaved RNA fragments were reverse transcribed into first strand cDNA using SuperScript II reverse transcriptase and random primers. The RNA template was then removed, and a replacement strand to generate double-stranded (ds) cDNA was synthesized. The ds cDNA was purified with AMPure XP beads, end-

repaired, 3'-adenylated, and ligated with indexing adapters, preparing them for hybridization onto a flow cell. The DNA fragments that have adapter molecules on both ends were selectively enriched by PCR-amplification and purified with AMPure XP beads. RNA-seq libraries were pooled together in equimolar concentrations and analyzed by single-end sequencing using NextSeq (Illumina). Four independent biological replicates for control samples, samples from cells with SCP4 knockout, and cells with STK35-PDIK1L double knockout were sequenced in the same flow cell.

5.15. RNA-Seq data analysis

Sequencing reads were mapped into reference human genome hg19 using STAR v. 2.5.2b-1 (Dobin et al. 2013). The mapped reads were assigned to genes using HTSeq-Count v. 0.6.1p1 (Anders et al. 2015). The differential expression gene analysis was performed using DESeq2, with four replicates for each sample (Love et al. 2014).

Genes were ranked by their log₂ fold change multiplied by the adjusted p-value (corrected for multiple testing) as input for GSEA Preranked analysis with all available signatures in the Molecular Signature Database v.7.2, complemented with various signatures generated in our laboratory.

Pseudo-alignment to the reference human transcriptome (gencode v.35), assignment of reads to genes, and TPM calculation were all performed using Salmon (v.1.0) (Patro et al. 2017).

For comparison with other RNA-Seq datasets, raw RNA-Seq files from GSE109491 and GSE104308 datasets corresponding to the available replicates of negative controls, sgTAF12, sgMYB, sgSIK3, and sgMEF2C were downloaded. All the raw datasets, including ours, were re-aligned (pseudo-alignment) using Kallisto to hg38, bootstrap 100 (Bray et al. 2016). The differential expression gene analysis was performed using DESeq2, with four

replicates for each sample (Love et al. 2014).

Pearson's product-moment correlations between global log₂ fold changes observed upon different knockouts in MOLM-13 cells were calculated using *cor()* function in R with default parameters. Transcripts Per Million (TPM) values were calculated using Kallisto (Bray et al. 2016).

5.16. Protein Expression and Purification

The SCP4 phosphatase domain (encoding residues 156–466) was sub-cloned into a pET28a (Novagene) derivative vector encoding a 6xHis-tag followed by a SUMO-tag and a TEV protease site. BL21 (DE3) cells expressing SCP4 were grown in one-liter cultures at 37°C in Luria-Bertani (LB) broth (Thermo) containing 50 µg/ml kanamycin. Once the cultures reached an OD₆₀₀ value of 0.6–0.8, the protein expression was induced with 0.2 M isopropyl-β-D-thiogalactopyranoside (IPTG), and the cultures were grown an additional 16 hours at 18°C.

The cells were pelleted and resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 15 mM imidazole, 10% glycerol, 0.1% Triton X-100, and 10 mM BME) and sonicated at 90 A for 2.5 minutes of 1 second on / 5 seconds off cycles on ice. The lysate was cleared by centrifugation at 15000 rpm for 45 min at 4°C. The supernatant was loaded over 3 ml of Ni-NTA beads (Qiagen) equilibrated in lysis buffer, then washed through with wash buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 30 mM imidazole, and 10 mM BME. The recombinant SCP4 was finally eluted with elution buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 300 mM imidazole, and 10 mM BME. Protein fractions were pooled and dialyzed overnight at 4°C in a 10.0 kDa dialysis membrane (Thermo) against dialysis buffer (50 mM Tris HCl pH 7.5, 100 mM NaCl, and 10 mM BME) supplemented with 0.5 mg of TEV protease to remove the 6xHis and SUMO-tag.

The protein was then loaded onto a DEAE anion exchange column (GE) equilibrated with Buffer A (50 mM Tris-HCl pH 7.5, 10 mM NaCl, 10 mM BME) and eluted with a 0-100% gradient of Buffer A to Buffer B (50 mM Tris-HCl pH 7.5, 500 mM NaCl, and 10 mM BME). Peak fractions were analyzed by SDS-PAGE before the fractions containing SCP4 were pooled and dialyzed overnight in gel filtration buffer containing 25 mM Tris HCl pH 8.0, 200 mM NaCl, and 10 mM BME. The protein was finally polished using gel filtration chromatography with a Superdex 75 size exclusion column. Fractions containing SCP4 were pooled, concentrated, and flash-frozen at -80°C .

5.17. Phosphatase activity assay

Initial rates of phosphatase activity were determined by incubating SCP4 ($1\mu\text{M}$) with various concentrations of phosphopeptide (0–100 μM) in SCP4 assay buffer (50 mM Tris-acetate pH 5.0, 10 mM MgCl_2) in a $40\mu\text{l}$ reaction volume for 10 minutes at 37°C . $20\mu\text{l}$ of the reaction was then quenched with $40\mu\text{l}$ of BIOMOL Green reagent (Enzo Life Sciences) within a clear, flat-bottom 96-well plate. The mixtures were allowed to sit at room temperature for 30 minutes for color development before reading the absorbance signal at A_{620} in a Tecan Plate reader 200. The A_{620} values were corrected by subtracting the A_{620} values of substrate only reactions at each substrate concentration. The readings obtained were converted to the amount of phosphate released through a Phosphate standard curve determined using the BIOMOL Green assay following the manufacturer instructions. The reaction rate was plotted for kinetic fitting to derive k_{cat}/K_M . Biological triplicates were prepared and analyzed independently to guarantee replication of experiments.

5.18. Metabolomics analysis by liquid chromatography coupled to mass spectrometry (LC-MS)

Cells (0.6 mln cells/well) were seeded in 6 well plates after lentiviral infection with sgRNAs. On Day 5 post-infection, cells were quickly washed in PBS before adding 1 mL of ice-cold extraction solution (50% methanol, 30% acetonitrile, 20% H₂O) per million cells. The cell suspension was snap frozen in liquid nitrogen. Samples were agitated using a tube rotator at 4°C for 15 minutes followed by incubation at -80°C overnight. Samples were then centrifuged at 15,000 rpm, 4°C for 10 minutes. The supernatants were collected and stored in autosampler vials at -80°C until analysis.

Intracellular extracts from five independent cell cultures were analyzed for each condition. Samples were randomized in order to avoid bias due to machine drift and processed blindly. LC-MS analysis was performed using a Vanquish Horizon UHPLC system coupled with a Q Exactive HF mass spectrometer (both Thermo Fisher Scientific). Sample extracts (8 µL) were injected onto a Sequant ZIC-pHILC column (150 mm × 2.1 mm, 5 µm) and guard column (20 mm × 2.1 mm, 5 µm) from Merck Millipore kept at 45°C. The mobile phase was composed of 20 mM ammonium carbonate with 0.1% ammonium hydroxide in water (solvent A), and acetonitrile (solvent B). The flow rate was set at 200 µL/min with the previously described gradient (Mackay et al. 2015). The mass spectrometer was operated in full MS and polarity switching mode. The acquired spectra were analyzed using XCalibur Qual Browser and XCalibur Quan Browser software (Thermo Fisher Scientific) by referencing to an internal library of compounds.

5.19. Quantification and statistical analysis

All statistical analysis and graphical representation of data was performed using GraphPad PRISM5 software. Error bars represent the mean plus or minus standard error of the mean, unless otherwise stated.

5.20. Data and software availability

The RNA-seq data from this study have been uploaded to GEO database with accession number GSE174169.

REFERENCES

- Ahmed TA, Adamopoulos C, Karoulia Z, Wu X, Sachidanandam R, Aaronson SA, Poulidakos PI. 2019. SHP2 Drives Adaptive Resistance to ERK Signaling Inhibition in Molecularly Defined Subsets of ERK-Dependent Tumors. *Cell Rep* **26**: 65-78 e65.
- Albaugh VL, Mukherjee K, Barbul A. 2017. Proline Precursors and Collagen Synthesis: Biochemical Challenges of Nutrient Supplementation and Wound Healing. *J Nutr* **147**: 2011-2017.
- Alfayez M, Issa GC, Patel KP, Wang F, Wang X, Short NJ, Cortes JE, Kadia T, Ravandi F, Pierce S et al. 2021. The Clinical impact of PTPN11 mutations in adults with acute myeloid leukemia. *Leukemia* **35**: 691-700.
- Alfayez M, Kantarjian H, Kadia T, Ravandi-Kashani F, Daver N. 2020. CPX-351 (vyxeos) in AML. *Leuk Lymphoma* **61**: 288-297.
- Alter BP. 2014. Fanconi anemia and the development of leukemia. *Best Pract Res Clin Haematol* **27**: 214-221.
- Alvira D, Naughton R, Bhatt L, Tedesco S, Landry WD, Cotter TG. 2011. Inhibition of protein-tyrosine phosphatase 1B (PTP1B) mediates ubiquitination and degradation of Bcr-Abl protein. *J Biol Chem* **286**: 32313-32323.
- Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166-169.
- Anderson KA, Noeldner PK, Reece K, Wadzinski BE, Means AR. 2004. Regulation and function of the calcium/calmodulin-dependent protein kinase IV/protein serine/threonine phosphatase 2A signaling complex. *J Biol Chem* **279**: 31708-31716.
- Apel A, Moshe Y, Ofra Y, Gural A, Wolach O, Ganzel C, Canaani J, Zektser M, Duek A, Stemer G et al. 2021. Venetoclax combinations induce high response rates in newly diagnosed acute myeloid leukemia patients ineligible for intensive chemotherapy in routine practice. *Am J Hematol*.
- Aramburu J, Ortells MC, Tejedor S, Buxade M, Lopez-Rodriguez C. 2014. Transcriptional regulation of the stress response by mTOR. *Sci Signal* **7**: re2.
- Armstrong SA, Staunton JE, Silverman LB, Pieters R, den Boer ML, Minden MD, Sallan SE, Lander ES, Golub TR, Korsmeyer SJ. 2002. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* **30**: 41-47.
- Arora D, Stopp S, Bohmer SA, Schons J, Godfrey R, Masson K, Razumovskaya E, Ronnstrand L, Tanzer S, Bauer R et al. 2011. Protein-tyrosine phosphatase DEP-1 controls receptor tyrosine kinase FLT3 signaling. *J Biol Chem* **286**: 10918-10929.
- Avramis VI. 2012. Asparaginases: biochemical pharmacology and modes of drug resistance. *Anticancer Res* **32**: 2423-2437.
- Ayatollahi H, Shajiei A, Sadeghian MH, Sheikhi M, Yazdandoust E, Ghazanfarpour M, Shams SF, Shakeri S. 2017. Prognostic Importance of C-KIT Mutations in Core Binding Factor Acute Myeloid Leukemia: A Systematic Review. *Hematol Oncol Stem Cell Ther* **10**: 1-7.

REFERENCES

- Bachelor MA, Lu Y, Owens DM. 2011. L-3-Phosphoserine phosphatase (PSPH) regulates cutaneous squamous cell carcinoma proliferation independent of L-serine biosynthesis. *J Dermatol Sci* **63**: 164-172.
- Badr P, Elsayed GM, Eldin DN, Riad BY, Hamdy N. 2018. Detection of KIT mutations in core binding factor acute myeloid leukemia. *Leuk Res Rep* **10**: 20-25.
- Banaszynski LA, Chen LC, Maynard-Smith LA, Ooi AG, Wandless TJ. 2006. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell* **126**: 995-1004.
- Barford D, Neel BG. 1998. Revealing mechanisms for SH2 domain mediated regulation of the protein tyrosine phosphatase SHP-2. *Structure* **6**: 249-254.
- Benarafa C, Simon HU. 2017. Role of granule proteases in the life and death of neutrophils. *Biochem Biophys Res Commun* **482**: 473-481.
- Bill M, Mrozek K, Kohlschmidt J, Eisfeld AK, Walker CJ, Nicolet D, Papaioannou D, Blachly JS, Orwick S, Carroll AJ et al. 2020. Mutational landscape and clinical outcome of patients with de novo acute myeloid leukemia and rearrangements involving 11q23/KMT2A. *Proc Natl Acad Sci U S A* **117**: 26340-26346.
- Biswas T, Yi L, Aggarwal P, Wu J, Rubin JR, Stuckey JA, Woodard RW, Tsodikov OV. 2009. The tail of KdsC: conformational changes control the activity of a haloacid dehalogenase superfamily phosphatase. *J Biol Chem* **284**: 30594-30603.
- Bjelosevic S, Gruber E, Newbold A, Shembrey C, Devlin JR, Hogg SJ, Kats L, Todorovski I, Fan Z, Abrehart TC et al. 2021. Serine biosynthesis is a metabolic vulnerability in FLT3-ITD-driven acute myeloid leukaemia. *Cancer Discov*.
- Blanchetot C, Chagnon M, Dube N, Halle M, Tremblay ML. 2005. Substrate-trapping techniques in the identification of cellular PTP targets. *Methods* **35**: 44-53.
- Bochtler T, Frohling S, Kramer A. 2015. Role of chromosomal aberrations in clonal diversity and progression of acute myeloid leukemia. *Leukemia* **29**: 1243-1252.
- Bohmer SA, Weibrecht I, Soderberg O, Bohmer FD. 2013. Association of the protein-tyrosine phosphatase DEP-1 with its substrate FLT3 visualized by in situ proximity ligation assay. *PLoS One* **8**: e62871.
- Boivin B, Zhang S, Arbiser JL, Zhang ZY, Tonks NK. 2008. A modified cysteinyl-labeling assay reveals reversible oxidation of protein tyrosine phosphatases in angiomyolipoma cells. *Proc Natl Acad Sci U S A* **105**: 9959-9964.
- Boj SF, Hwang CI, Baker LA, Chio, II, Engle DD, Corbo V, Jager M, Ponz-Sarvisse M, Tiriack H, Spector MS et al. 2015. Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**: 324-338.
- Botelho RJ. 2009. Changing phosphoinositides "on the fly": how trafficking vesicles avoid an identity crisis. *Bioessays* **31**: 1127-1136.
- Bower H, Bjorkholm M, Dickman PW, Hoglund M, Lambert PC, Andersson TM. 2016. Life Expectancy of Patients With Chronic Myeloid Leukemia Approaches the Life Expectancy of the General Population. *J Clin Oncol* **34**: 2851-2857.

REFERENCES

- Brautigian DL. 2013. Protein Ser/Thr phosphatases--the ugly ducklings of cell signalling. *FEBS J* **280**: 324-345.
- Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**: 525-527.
- Brunet A, Roux D, Lenormand P, Dowd S, Keyse S, Pouyssegur J. 1999. Nuclear translocation of p42/p44 mitogen-activated protein kinase is required for growth factor-induced gene expression and cell cycle entry. *EMBO J* **18**: 664-674.
- Bullinger L, Dohner K, Dohner H. 2017. Genomics of Acute Myeloid Leukemia Diagnosis and Pathways. *J Clin Oncol* **35**: 934-946.
- Bunda S, Burrell K, Heir P, Zeng L, Alamsahebpoor A, Kano Y, Raught B, Zhang ZY, Zadeh G, Ohh M. 2015. Inhibition of SHP2-mediated dephosphorylation of Ras suppresses oncogenesis. *Nat Commun* **6**: 8859.
- Burkholder NT, Mayfield JE, Yu X, Irani S, Arce DK, Jiang F, Matthews WL, Xue Y, Zhang YJ. 2018. Phosphatase activity of small C-terminal domain phosphatase 1 (SCP1) controls the stability of the key neuronal regulator RE1-silencing transcription factor (REST). *J Biol Chem* **293**: 16851-16861.
- Burroughs AM, Allen KN, Dunaway-Mariano D, Aravind L. 2006. Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. *J Mol Biol* **361**: 1003-1034.
- Cagatay T, Chook YM. 2018. Karyopherins in cancer. *Curr Opin Cell Biol* **52**: 30-42.
- Caldwell JT, Ge Y, Taub JW. 2014. Prognosis and management of acute myeloid leukemia in patients with Down syndrome. *Expert Rev Hematol* **7**: 831-840.
- Camps M, Nichols A, Gillieron C, Antonsson B, Muda M, Chabert C, Boschert U, Arkininstall S. 1998. Catalytic activation of the phosphatase MKP-3 by ERK2 mitogen-activated protein kinase. *Science* **280**: 1262-1265.
- Cao J, Yu Y, Zhang Z, Chen X, Hu Z, Tong Q, Chang J, Feng XH, Lin X. 2018. SCP4 Promotes Gluconeogenesis Through FoxO1/3a Dephosphorylation. *Diabetes* **67**: 46-57.
- Carter JL, Hege K, Yang J, Kalpage HA, Su Y, Edwards H, Huttemann M, Taub JW, Ge Y. 2020. Targeting multiple signaling pathways: the new approach to acute myeloid leukemia therapy. *Signal Transduct Target Ther* **5**: 288.
- Castro I, Sampaio-Marques B, Ludovico P. 2019. Targeting Metabolic Reprogramming in Acute Myeloid Leukemia. *Cells* **8**.
- Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, Morozova O, Newton Y, Radenbaugh A, Pagnotta SM et al. 2016. Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. *Cell* **164**: 550-563.
- Chang KT, Guo J, di Ronza A, Sardiello M. 2018. Aminode: Identification of Evolutionary Constraints in the Human Proteome. *Sci Rep* **8**: 1357.
- Chen CC, Li B, Millman SE, Chen C, Li X, Morris JPt, Mayle A, Ho YJ, Loizou E, Liu H et al. 2020. Vitamin B6 Addiction in Acute Myeloid Leukemia. *Cancer Cell* **37**: 71-84 e77.

REFERENCES

- Chen MJ, Dixon JE, Manning G. 2017. Genomics and evolution of protein phosphatases. *Sci Signal* **10**.
- Chen R, Amoui M, Zhang Z, Mardon G. 1997. Dachshund and eyes absent proteins form a complex and function synergistically to induce ectopic eye development in *Drosophila*. *Cell* **91**: 893-903.
- Chen WL, Wang JH, Zhao AH, Xu X, Wang YH, Chen TL, Li JM, Mi JQ, Zhu YM, Liu YF et al. 2014. A distinct glucose metabolism signature of acute myeloid leukemia with prognostic value. *Blood* **124**: 1645-1654.
- Chen YN, LaMarche MJ, Chan HM, Fekkes P, Garcia-Fortanet J, Acker MG, Antonakos B, Chen CH, Chen Z, Cooke VG et al. 2016. Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. *Nature* **535**: 148-152.
- Cheng A, Uetani N, Simoncic PD, Chaubey VP, Lee-Loy A, McGlade CJ, Kennedy BP, Tremblay ML. 2002. Attenuation of leptin action and regulation of obesity by protein tyrosine phosphatase 1B. *Dev Cell* **2**: 497-503.
- Cherry M, Williams DH. 2004. Recent kinase and kinase inhibitor X-ray structures: mechanisms of inhibition and selectivity insights. *Curr Med Chem* **11**: 663-673.
- Chian R, Young S, Danilkovitch-Miagkova A, Ronnstrand L, Leonard E, Ferrao P, Ashman L, Linnekin D. 2001. Phosphatidylinositol 3 kinase contributes to the transformation of hematopoietic cells by the D816V c-Kit mutant. *Blood* **98**: 1365-1373.
- Choudhary C, Brandts C, Schwable J, Tickenbrock L, Sargin B, Ueker A, Bohmer FD, Berdel WE, Muller-Tidow C, Serve H. 2007. Activation mechanisms of STAT5 by oncogenic Flt3-ITD. *Blood* **110**: 370-374.
- Chow RD, Chen S. 2018. Cancer CRISPR Screens In Vivo. *Trends Cancer* **4**: 349-358.
- Colledge M, Scott JD. 1999. AKAPs: from structure to function. *Trends Cell Biol* **9**: 216-221.
- Cook PJ, Ju BG, Telese F, Wang X, Glass CK, Rosenfeld MG. 2009. Tyrosine dephosphorylation of H2AX modulates apoptosis and survival decisions. *Nature* **458**: 591-596.
- Coombs CC, Zehir A, Devlin SM, Kishtagari A, Syed A, Jonsson P, Hyman DM, Solit DB, Robson ME, Baselga J et al. 2017. Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell* **21**: 374-382 e374.
- Cortes JE, Heidel FH, Fiedler W, Smith BD, Robak T, Montesinos P, Candoni A, Leber B, Sekeres MA, Pollyea DA et al. 2020. Survival outcomes and clinical benefit in patients with acute myeloid leukemia treated with glasdegib and low-dose cytarabine according to response to therapy. *J Hematol Oncol* **13**: 92.
- Cossa G, Parua PK, Eilers M, Fisher RP. 2021. Protein phosphatases in the RNAPII transcription cycle: erasers, sculptors, gatekeepers, and potential drug targets. *Genes Dev* **35**: 658-676.
- Cunningham I, Kohno B. 2016. 18 FDG-PET/CT: 21st century approach to leukemic tumors in 124 cases. *Am J Hematol* **91**: 379-384.
- Daver N, Schlenk RF, Russell NH, Levis MJ. 2019. Targeting FLT3 mutations in AML: review of current knowledge and evidence. *Leukemia* **33**: 299-312.

REFERENCES

- Dawson MA, Bannister AJ, Gottgens B, Foster SD, Bartke T, Green AR, Kouzarides T. 2009. JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. *Nature* **461**: 819-822.
- de Klein A, van Kessel AG, Grosveld G, Bartram CR, Hagemeijer A, Bootsma D, Spurr NK, Heisterkamp N, Groffen J, Stephenson JR. 1982. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. *Nature* **300**: 765-767.
- Deeks ED. 2016. Venetoclax: First Global Approval. *Drugs* **76**: 979-987.
- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**: 602-607.
- Dempster JM, Rossen J, Kazachkova M, Pan J, Kugener G, Root DE, Tsherniak A. 2019. Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines. *bioRxiv*: 720243.
- Dhillon S. 2019. Gilteritinib: First Global Approval. *Drugs* **79**: 331-339.
- DiNardo CD, Jonas BA, Pullarkat V, Thirman MJ, Garcia JS, Wei AH, Konopleva M, Dohner H, Letai A, Fenaux P et al. 2020a. Azacitidine and Venetoclax in Previously Untreated Acute Myeloid Leukemia. *N Engl J Med* **383**: 617-629.
- DiNardo CD, Ravandi F, Agresta S, Konopleva M, Takahashi K, Kadia T, Routbort M, Patel KP, Mark B, Pierce S et al. 2015. Characteristics, clinical outcome, and prognostic significance of IDH mutations in AML. *Am J Hematol* **90**: 732-736.
- DiNardo CD, Tiong IS, Quaglieri A, MacRaidl S, Loghavi S, Brown FC, Thijssen R, Pomilio G, Ivey A, Salmon JM et al. 2020b. Molecular patterns of response and treatment failure after frontline venetoclax combinations in older patients with AML. *Blood* **135**: 791-803.
- Djouder N, Tuerk RD, Suter M, Salvioni P, Thali RF, Scholz R, Vaahtomeri K, Auchli Y, Rechsteiner H, Brunisholz RA et al. 2010. PKA phosphorylates and inactivates AMPKalpha to promote efficient lipolysis. *EMBO J* **29**: 469-481.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.
- Dohner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Buchner T, Dombret H, Ebert BL, Fenaux P, Larson RA et al. 2017. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* **129**: 424-447.
- Dohner H, Estey EH, Amadori S, Appelbaum FR, Buchner T, Burnett AK, Dombret H, Fenaux P, Grimwade D, Larson RA et al. 2010. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood* **115**: 453-474.
- Edwards AM, Isserlin R, Bader GD, Frye SV, Willson TM, Yu FH. 2011. Too many roads not taken. *Nature* **470**: 163-165.
- Eguchi M, Minami Y, Kuzume A, Chi S. 2020. Mechanisms Underlying Resistance to FLT3 Inhibitors in Acute Myeloid Leukemia. *Biomedicines* **8**.

REFERENCES

- Eisner A, Pazyra-Murphy MF, Durreesi E, Zhou P, Zhao X, Chadwick EC, Xu PX, Hillman RT, Scott MP, Greenberg ME et al. 2015. The Eya1 phosphatase promotes Shh signaling during hindbrain development and oncogenesis. *Dev Cell* **33**: 22-35.
- Elchebly M, Payette P, Michaliszyn E, Cromlish W, Collins S, Loy AL, Normandin D, Cheng A, Himms-Hagen J, Chan CC et al. 1999. Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science* **283**: 1544-1548.
- Elgehama A, Chen W, Pang J, Mi S, Li J, Guo W, Wang X, Gao J, Yu B, Shen Y et al. 2016. Blockade of the interaction between Bcr-Abl and PTB1B by small molecule SBF-1 to overcome imatinib-resistance of chronic myeloid leukemia cells. *Cancer Lett* **372**: 82-88.
- Ellens KW, Christian N, Singh C, Satagopam VP, May P, Linster CL. 2017. Confronting the catalytic dark matter encoded by sequenced genomes. *Nucleic Acids Res* **45**: 11495-11514.
- Elson A. 2018. Stepping out of the shadows: Oncogenic and tumor-promoting protein tyrosine phosphatases. *Int J Biochem Cell Biol* **96**: 135-147.
- Fahs S, Lujan P, Kohn M. 2016. Approaches to Study Phosphatases. *ACS Chem Biol* **11**: 2944-2961.
- Fanidi A, Muller DC, Yuan JM, Stevens VL, Weinstein SJ, Albanes D, Prentice R, Thomsen CA, Pettinger M, Cai Q et al. 2018. Circulating Folate, Vitamin B6, and Methionine in Relation to Lung Cancer Risk in the Lung Cancer Cohort Consortium (LC3). *J Natl Cancer Inst* **110**.
- Farrell AS, Allen-Petersen B, Daniel CJ, Wang X, Wang Z, Rodriguez S, Impey S, Oddo J, Vitek MP, Lopez C et al. 2014. Targeting inhibitors of the tumor suppressor PP2A for the treatment of pancreatic cancer. *Mol Cancer Res* **12**: 924-939.
- Farrington CC, Yuan E, Mazhar S, Izadmehr S, Hurst L, Allen-Petersen BL, Janghorban M, Chung E, Wolczanski G, Galsky M et al. 2020. Protein phosphatase 2A activation as a therapeutic strategy for managing MYC-driven cancers. *J Biol Chem* **295**: 757-770.
- Fernandez-Fernandez MR, Gragera M, Ochoa-Ibarrola L, Quintana-Gallardo L, Valpuesta JM. 2017. Hsp70 - a master regulator in protein degradation. *FEBS Lett* **591**: 2648-2660.
- Flint AJ, Tiganis T, Barford D, Tonks NK. 1997. Development of "substrate-trapping" mutants to identify physiological substrates of protein tyrosine phosphatases. *Proc Natl Acad Sci U S A* **94**: 1680-1685.
- Fonda ML. 1992. Purification and characterization of vitamin B6-phosphate phosphatase from human erythrocytes. *J Biol Chem* **267**: 15978-15983.
- Fowle H, Zhao Z, Grana X. 2019. PP2A holoenzymes, substrate specificity driving cellular functions and deregulation in cancer. *Adv Cancer Res* **144**: 55-93.
- Frankson R, Yu ZH, Bai Y, Li Q, Zhang RY, Zhang ZY. 2017. Therapeutic Targeting of Oncogenic Tyrosine Phosphatases. *Cancer Res* **77**: 5701-5705.
- Freschauf GK, Karimi-Busheri F, Ulaczyk-Lesanko A, Mereniuk TR, Ahrens A, Koshy JM, Rasouli-Nia A, Pasarj P, Holmes CF, Rininsland F et al. 2009. Identification of a small molecule inhibitor of the human DNA repair enzyme polynucleotide kinase/phosphatase. *Cancer Res* **69**: 7739-7746.

REFERENCES

- Fukushima N, Minami Y, Kakiuchi S, Kuwatsuka Y, Hayakawa F, Jamieson C, Kiyoi H, Naoe T. 2016. Small-molecule Hedgehog inhibitor attenuates the leukemia-initiation potential of acute myeloid leukemia cells. *Cancer Sci* **107**: 1422-1429.
- Gallipoli P, Giotopoulos G, Tzelepis K, Costa ASH, Vohra S, Medina-Perez P, Basheer F, Marando L, Di Lisio L, Dias JML et al. 2018. Glutaminolysis is a metabolic dependency in FLT3(ITD) acute myeloid leukemia unmasked by FLT3 tyrosine kinase inhibition. *Blood* **131**: 1639-1653.
- Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, Alnemri ES, Altucci L, Amelio I, Andrews DW et al. 2018. Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death Differ* **25**: 486-541.
- Gelens L, Qian J, Bollen M, Saurin AT. 2018. The Importance of Kinase-Phosphatase Integration: Lessons from Mitosis. *Trends Cell Biol* **28**: 6-21.
- Genovese G, Kahler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, Chambert K, Mick E, Neale BM, Fromer M et al. 2014. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* **371**: 2477-2487.
- Gilliland DG, Griffin JD. 2002. The roles of FLT3 in hematopoiesis and leukemia. *Blood* **100**: 1532-1542.
- Godfrey R, Arora D, Bauer R, Stopp S, Muller JP, Heinrich T, Bohmer SA, Dagnell M, Schnetzke U, Scholl S et al. 2012. Cell transformation by FLT3 ITD in acute myeloid leukemia involves oxidative inactivation of the tumor suppressor protein-tyrosine phosphatase DEP-1/ PTPRJ. *Blood* **119**: 4499-4511.
- Goyal P, Behring A, Kumar A, Siess W. 2009. Identifying and characterizing a novel protein kinase STK35L1 and deciphering its orthologs and close-homologs in vertebrates. *PLoS One* **4**: e6981.
- Goyal P, Behring A, Kumar A, Siess W. 2011. STK35L1 associates with nuclear actin and regulates cell cycle and migration of endothelial cells. *PLoS One* **6**: e16249.
- Grafone T, Palmisano M, Nicci C, Storti S. 2012. An overview on the role of FLT3-tyrosine kinase receptor in acute myeloid leukemia: biology and treatment. *Oncol Rev* **6**: e8.
- Greenbaum H, Galper BL, Decter DH, Eisenberg VH. 2021. Endometriosis and autoimmunity: Can autoantibodies be used as a non-invasive early diagnostic tool? *Autoimmun Rev* **20**: 102795.
- Gregory MA, Qi Y, Hann SR. 2003. Phosphorylation by glycogen synthase kinase-3 controls c-myc proteolysis and subnuclear localization. *J Biol Chem* **278**: 51606-51612.
- Grieselhuber NR, Mims AS. 2021. Novel Targeted Therapeutics in Acute Myeloid Leukemia: an Embarrassment of Riches. *Curr Hematol Malig Rep*.
- Grigsby SM, Friedman A, Chase J, Waas B, Ropa J, Serio J, Shen C, Muntean AG, Maillard I, Nikolovska-Coleska Z. 2021. Elucidating the Importance of DOT1L Recruitment in MLL-AF9 Leukemia and Hematopoiesis. *Cancers (Basel)* **13**.
- Grove CS, Vassiliou GS. 2014. Acute myeloid leukaemia: a paradigm for the clonal evolution of cancer? *Dis Model Mech* **7**: 941-951.
- Gu S, Chan WW, Mohi G, Rosenbaum J, Sayad A, Lu Z, Virtanen C, Li S, Neel BG, Van Etten RA. 2016. Distinct GAB2 signaling pathways are essential for myeloid and lymphoid transformation and leukemogenesis by BCR-ABL1. *Blood* **127**: 1803-1813.

REFERENCES

- Gu S, Sayad A, Chan G, Yang W, Lu Z, Virtanen C, Van Etten RA, Neel BG. 2018. SHP2 is required for BCR-ABL1-induced hematologic neoplasia. *Leukemia* **32**: 203-213.
- Han K, Jeng EE, Hess GT, Morgens DW, Li A, Bassik MC. 2017. Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nat Biotechnol* **35**: 463-474.
- Hannon GJ, Casso D, Beach D. 1994. KAP: a dual specificity phosphatase that interacts with cyclin-dependent kinases. *Proc Natl Acad Sci U S A* **91**: 1731-1735.
- Hantschel O. 2012. Structure, regulation, signaling, and targeting of abl kinases in cancer. *Genes Cancer* **3**: 436-446.
- Hao HX, Wang H, Liu C, Kovats S, Velazquez R, Lu H, Pant B, Shirley M, Meyer MJ, Pu M et al. 2019. Tumor Intrinsic Efficacy by SHP2 and RTK Inhibitors in KRAS-Mutant Cancers. *Mol Cancer Ther* **18**: 2368-2380.
- Hattori A, Tsunoda M, Konuma T, Kobayashi M, Nagy T, Glushka J, Tayyari F, McSkimming D, Kannan N, Tojo A et al. 2017. Cancer progression by reprogrammed BCAA metabolism in myeloid leukaemia. *Nature* **545**: 500-504.
- Hawkinson JE, Acosta-Burrueal M, Wood PL. 1996. The metabotropic glutamate receptor antagonist L-2-amino-3-phosphonopropionic acid inhibits phosphoserine phosphatase. *Eur J Pharmacol* **307**: 219-225.
- Hay N. 2016. Reprogramming glucose metabolism in cancer: can it be exploited for cancer therapy? *Nat Rev Cancer* **16**: 635-649.
- Helps NR, Luo X, Barker HM, Cohen PT. 2000. NIMA-related kinase 2 (Nek2), a cell-cycle-regulated protein kinase localized to centrosomes, is complexed to protein phosphatase 1. *Biochem J* **349**: 509-518.
- Hendel A, Bak RO, Clark JT, Kennedy AB, Ryan DE, Roy S, Steinfeld I, Lunstad BD, Kaiser RJ, Wilkens AB et al. 2015. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol* **33**: 985-989.
- Hensley CT, Wasti AT, DeBerardinis RJ. 2013. Glutamine and cancer: cell biology, physiology, and clinical opportunities. *J Clin Invest* **123**: 3678-3684.
- Hermiston ML, Zikherman J, Zhu JW. 2009. CD45, CD148, and Lyp/Pep: critical phosphatases regulating Src family kinase signaling networks in immune cells. *Immunol Rev* **228**: 288-311.
- Herold T, Rothenberg-Thurley M, Grunwald VV, Janke H, Goerlich D, Sauerland MC, Konstandin NP, Dufour A, Schneider S, Neusser M et al. 2020. Validation and refinement of the revised 2017 European LeukemiaNet genetic risk stratification of acute myeloid leukemia. *Leukemia* **34**: 3161-3172.
- Hiragami-Hamada K, Tani N, Nakayama J-i. 2018. Proteomic analysis of RNA-dependent chromatin association of nuclear proteins. *bioRxiv*: 391755.
- Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. 2015. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* **43**: D512-520.

REFERENCES

- Hosaka T, Biggs WH, 3rd, Tieu D, Boyer AD, Varki NM, Cavenee WK, Arden KC. 2004. Disruption of forkhead transcription factor (FOXO) family members in mice reveals their functional diversification. *Proc Natl Acad Sci U S A* **101**: 2975-2980.
- Hoy SM. 2019. Glasdegib: First Global Approval. *Drugs* **79**: 207-213.
- Hsieh JJ, Cheng EH, Korsmeyer SJ. 2003. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* **115**: 293-303.
- Hsu JJ, Dayaram T, Tovy A, De Braekeleer E, Jeong M, Wang F, Zhang J, Heffernan TP, Gera S, Kovacs JJ et al. 2018. PPM1D Mutations Drive Clonal Hematopoiesis in Response to Cytotoxic Chemotherapy. *Cell Stem Cell* **23**: 700-713 e706.
- Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O et al. 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**: 827-832.
- Hughes TP, Ross DM. 2016. Moving treatment-free remission into mainstream clinical practice in CML. *Blood* **128**: 17-23.
- Hunter T. 2000. Signaling--2000 and beyond. *Cell* **100**: 113-127.
- Hunter T. 2012. Why nature chose phosphate to modify proteins. *Philos Trans R Soc Lond B Biol Sci* **367**: 2513-2516.
- Ingham PW, Nakano Y, Seger C. 2011. Mechanisms and functions of Hedgehog signalling across the metazoa. *Nat Rev Genet* **12**: 393-406.
- Iqbal N, Iqbal N. 2014. Imatinib: a breakthrough of targeted therapy in cancer. *Chemother Res Pract* **2014**: 357027.
- Irie-Sasaki J, Sasaki T, Matsumoto W, Opavsky A, Cheng M, Welstead G, Griffiths E, Krawczyk C, Richardson CD, Aitken K et al. 2001. CD45 is a JAK phosphatase and negatively regulates cytokine receptor signalling. *Nature* **409**: 349-354.
- Jacque N, Ronchetti AM, Larrue C, Meunier G, Birsén R, Willems L, Saland E, Decroocq J, Maciel TT, Lambert M et al. 2015. Targeting glutaminolysis has antileukemic activity in acute myeloid leukemia and synergizes with BCL-2 inhibition. *Blood* **126**: 1346-1356.
- Jaiswal S, Ebert BL. 2019. Clonal hematopoiesis in human aging and disease. *Science* **366**.
- Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, Lindsley RC, Mermel CH, Burt N, Chavez A et al. 2014. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* **371**: 2488-2498.
- Jemc J, Rebay I. 2007. The eyes absent family of phosphotyrosine phosphatases: properties and roles in developmental regulation of transcription. *Annu Rev Biochem* **76**: 513-538.
- Jeong JJ, Gu X, Nie J, Sundaravel S, Liu H, Kuo WL, Bhagat TD, Pradhan K, Cao J, Nischal S et al. 2019. Cytokine-Regulated Phosphorylation and Activation of TET2 by JAK2 in Hematopoiesis. *Cancer Discov* **9**: 778-795.
- Jiang J, Srivastava S, Seim G, Pavlova NN, King B, Zou L, Zhang C, Zhong M, Feng H, Kapur R et al. 2019. Promoter demethylation of the asparagine synthetase gene is required for ATF4-dependent adaptation to asparagine depletion. *J Biol Chem* **294**: 18674-18684.

REFERENCES

- Jin J, Pawson T. 2012. Modular evolution of phosphorylation-based signalling systems. *Philos Trans R Soc Lond B Biol Sci* **367**: 2540-2555.
- Jin M, Jusiak B, Bai Z, Mardon G. 2013. Eyes absent tyrosine phosphatase activity is not required for *Drosophila* development or survival. *PLoS One* **8**: e58818.
- Jin S, Zhao H, Yi Y, Nakata Y, Kalota A, Gewirtz AM. 2010. c-Myb binds MLL through menin in human leukemia cells and is an important driver of MLL-associated leukemogenesis. *J Clin Invest* **120**: 593-606.
- Jones CL, Stevens BM, D'Alessandro A, Reisz JA, Culp-Hill R, Nemkov T, Pei S, Khan N, Adane B, Ye H et al. 2018. Inhibition of Amino Acid Metabolism Selectively Targets Human Leukemia Stem Cells. *Cancer Cell* **34**: 724-740 e724.
- Jones CL, Stevens BM, Pollyea DA, Culp-Hill R, Reisz JA, Nemkov T, Gehrke S, Gamboni F, Krug A, Winters A et al. 2020. Nicotinamide Metabolism Mediates Resistance to Venetoclax in Relapsed Acute Myeloid Leukemia Stem Cells. *Cell Stem Cell* **27**: 748-764 e744.
- Ju HQ, Zhan G, Huang A, Sun Y, Wen S, Yang J, Lu WH, Xu RH, Li J, Li Y et al. 2017. ITD mutation in FLT3 tyrosine kinase promotes Warburg effect and renders therapeutic sensitivity to glycolytic inhibition. *Leukemia* **31**: 2143-2150.
- Julien SG, Dube N, Hardy S, Tremblay ML. 2011. Inside the human cancer tyrosine phosphatome. *Nat Rev Cancer* **11**: 35-49.
- Jurcik J, Sivakova B, Cipakova I, Selicky T, Stupenova E, Jurcik M, Osadska M, Barath P, Cipak L. 2020. Phosphoproteomics Meets Chemical Genetics: Approaches for Global Mapping and Deciphering the Phosphoproteome. *Int J Mol Sci* **21**.
- Justice JFt, Morgan RW, Beemon KL. 2015. Common Viral Integration Sites Identified in Avian Leukosis Virus-Induced B-Cell Lymphomas. *mBio* **6**: e01863-01815.
- Kamps MP, Sefton BM. 1986. Neither arginine nor histidine can carry out the function of lysine-295 in the ATP-binding site of p60src. *Mol Cell Biol* **6**: 751-757.
- Kantarjian HM, Kadia TM, DiNardo CD, Welch MA, Ravandi F. 2021. Acute myeloid leukemia: Treatment and research outlook for 2021 and the MD Anderson approach. *Cancer* **127**: 1186-1207.
- Kashuba VI, Li J, Wang F, Senchenko VN, Protopopov A, Malyukova A, Kutsenko AS, Kadyrova E, Zabarovska VI, Muravenko OV et al. 2004. RBSP3 (HYA22) is a tumor suppressor gene implicated in major epithelial malignancies. *Proc Natl Acad Sci U S A* **101**: 4906-4911.
- Kats LM, Reschke M, Taulli R, Pozdnyakova O, Burgess K, Bhargava P, Straley K, Karnik R, Meissner A, Small D et al. 2014. Proto-oncogenic role of mutant IDH2 in leukemia initiation and maintenance. *Cell Stem Cell* **14**: 329-341.
- Kats LM, Vervoort SJ, Cole R, Rogers AJ, Gregory GP, Vidacs E, Li J, Nagaraja R, Yen KE, Johnstone RW. 2017. A pharmacogenomic approach validates AG-221 as an effective and on-target therapy in IDH2 mutant AML. *Leukemia* **31**: 1466-1470.
- Kauko O, Westermarck J. 2018. Non-genomic mechanisms of protein phosphatase 2A (PP2A) regulation in cancer. *Int J Biochem Cell Biol* **96**: 157-164.

REFERENCES

- Keller A, Nesvizhskii AI, Kolker E, Aebersold R. 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **74**: 5383-5392.
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845-858.
- Kelly LM, Gilliland DG. 2002. GENETICS OF MYELOID LEUKEMIAS. *Annual Review of Genomics and Human Genetics* **3**: 179-198.
- Kemmer D, Podowski RM, Arenillas D, Lim J, Hodges E, Roth P, Sonnhammer EL, Hoog C, Wasserman WW. 2006. NovelFam3000--uncharacterized human protein domains conserved across model organisms. *BMC Genomics* **7**: 48.
- Kennedy JA, Atenafu EG, Messner HA, Craddock KJ, Brandwein JM, Lipton JH, Minden MD, Schimmer AD, Schuh AC, Yee KW et al. 2013. Treatment outcomes following leukemic transformation in Philadelphia-negative myeloproliferative neoplasms. *Blood* **121**: 2725-2733.
- Khurana A, Shafer DA. 2019. MDM2 antagonists as a novel treatment option for acute myeloid leukemia: perspectives on the therapeutic potential of idasanutlin (RG7388). *Onco Targets Ther* **12**: 2903-2910.
- Kikushige Y, Yoshimoto G, Miyamoto T, Iino T, Mori Y, Iwasaki H, Niuro H, Takenaka K, Nagafuji K, Harada M et al. 2008. Human Flt3 is expressed at the hematopoietic stem cell and the granulocyte/macrophage progenitor stages to maintain cell survival. *J Immunol* **180**: 7358-7367.
- Kim ES. 2017. Midostaurin: First Global Approval. *Drugs* **77**: 1251-1259.
- Klaman LD, Boss O, Peroni OD, Kim JK, Martino JL, Zabolotny JM, Moghal N, Lubkin M, Kim YB, Sharpe AH et al. 2000. Increased energy expenditure, decreased adiposity, and tissue-specific insulin sensitivity in protein-tyrosine phosphatase 1B-deficient mice. *Mol Cell Biol* **20**: 5479-5489.
- Knott GJ, Doudna JA. 2018. CRISPR-Cas guides the future of genetic engineering. *Science* **361**: 866-869.
- Koretzky GA, Picus J, Thomas ML, Weiss A. 1990. Tyrosine phosphatase CD45 is essential for coupling T-cell antigen receptor to the phosphatidylinositol pathway. *Nature* **346**: 66-68.
- Kornev AP, Haste NM, Taylor SS, Eyck LF. 2006. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A* **103**: 17783-17788.
- Kosugi S, Hasebe M, Tomita M, Yanagawa H. 2009. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A* **106**: 10171-10176.
- Kota S, Hou S, Guerrant W, Madoux F, Troutman S, Fernandez-Vega V, Alekseeva N, Madala N, Scampavia L, Kissil J et al. 2018. A novel three-dimensional high-throughput screening approach identifies inducers of a mutant KRAS selective lethal phenotype. *Oncogene* **37**: 4372-4384.
- Kreitz J, Schonfeld C, Seibert M, Stolp V, Alshamleh I, Oellerich T, Steffen B, Schwalbe H, Schnutgen F, Kurrle N et al. 2019. Metabolic Plasticity of Acute Myeloid Leukemia. *Cells* **8**.

REFERENCES

- Kresinsky A, Bauer R, Schnoder TM, Berg T, Meyer D, Ast V, Konig R, Serve H, Heidel FH, Bohmer FD et al. 2018. Loss of DEP-1 (Ptpnj) promotes myeloproliferative disease in FLT3-ITD acute myeloid leukemia. *Haematologica* **103**: e505-e509.
- Kress TR, Sabo A, Amati B. 2015. MYC: connecting selective transcriptional control to global RNA production. *Nat Rev Cancer* **15**: 593-607.
- Krishnan N, Jeong DG, Jung SK, Ryu SE, Xiao A, Allis CD, Kim SJ, Tonks NK. 2009. Dephosphorylation of the C-terminal tyrosyl residue of the DNA damage-related histone H2A.X is mediated by the protein phosphatase eyes absent. *J Biol Chem* **284**: 16066-16070.
- Krishnan N, Koveal D, Miller DH, Xue B, Akshinthala SD, Kragelj J, Jensen MR, Gauss CM, Page R, Blackledge M et al. 2014. Targeting the disordered C terminus of PTP1B with an allosteric inhibitor. *Nat Chem Biol* **10**: 558-566.
- Krivtsov AV, Twomey D, Feng Z, Stubbs MC, Wang Y, Faber J, Levine JE, Wang J, Hahn WC, Gilliland DG et al. 2006. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* **442**: 818-822.
- Kronke J, Bullinger L, Teleanu V, Tschurtz F, Gaidzik VI, Kuhn MW, Rucker FG, Holzmann K, Paschka P, Kapp-Schworer S et al. 2013. Clonal evolution in relapsed NPM1-mutated acute myeloid leukemia. *Blood* **122**: 100-108.
- Krueger AB, Dehdashti SJ, Southall N, Marugan JJ, Ferrer M, Li X, Ford HL, Zheng W, Zhao R. 2013. Identification of a selective small-molecule inhibitor series targeting the eyes absent 2 (Eya2) phosphatase activity. *J Biomol Screen* **18**: 85-96.
- Krueger AB, Drasin DJ, Lea WA, Patrick AN, Patnaik S, Backos DS, Matheson CJ, Hu X, Barnaeva E, Holliday MJ et al. 2014. Allosteric inhibitors of the Eya2 phosphatase are selective and inhibit Eya2-mediated cell migration. *J Biol Chem* **289**: 16349-16361.
- Kung C, Pingel JT, Heikinheimo M, Klemola T, Varkila K, Yoo LI, Vuopala K, Poyhonen M, Uhari M, Rogers M et al. 2000. Mutations in the tyrosine phosphatase CD45 gene in a child with severe combined immunodeficiency disease. *Nat Med* **6**: 343-345.
- Kustatscher G, Wills KL, Furlan C, Rappsilber J. 2014. Chromatin enrichment for proteomics. *Nat Protoc* **9**: 2090-2099.
- Lagadinou ED, Sach A, Callahan K, Rossi RM, Neering SJ, Minhajuddin M, Ashton JM, Pei S, Grose V, O'Dwyer KM et al. 2013. BCL-2 inhibition targets oxidative phosphorylation and selectively eradicates quiescent human leukemia stem cells. *Cell Stem Cell* **12**: 329-341.
- LaMontagne KR, Jr., Flint AJ, Franza BR, Jr., Pandergast AM, Tonks NK. 1998. Protein tyrosine phosphatase 1B antagonizes signalling by oncoprotein tyrosine kinase p210 bcr-abl in vivo. *Mol Cell Biol* **18**: 2965-2975.
- Lancet JE, Uy GL, Cortes JE, Newell LF, Lin TL, Ritchie EK, Stuart RK, Strickland SA, Hogge D, Solomon SR et al. 2018. CPX-351 (cytarabine and daunorubicin) Liposome for Injection Versus Conventional Cytarabine Plus Daunorubicin in Older Patients With Newly Diagnosed Secondary Acute Myeloid Leukemia. *J Clin Oncol* **36**: 2684-2692.
- Lancet JE, Uy GL, Newell LF, Lin TL, Ritchie EK, Stuart RK, Strickland SA, Hogge D, Solomon SR, Bixby D et al. 2020. Five-year final results of a phase III study of CPX-351 versus 7+3 in older adults with newly diagnosed high-risk/secondary AML. *Journal of Clinical Oncology* **38**: 7510-7510.

REFERENCES

- Lazarevic VL. 2021. Acute myeloid leukemia in patients we judge as being older and/or unfit. *J Intern Med*.
- Lecompte O, Poch O, Laporte J. 2008. PtdIns5P regulation through evolution: roles in membrane trafficking? *Trends Biochem Sci* **33**: 453-460.
- Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J et al. 2010. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* **363**: 2424-2433.
- Li J, Rodriguez Y, Cheng C, Zeng L, Wong EYM, Xu CY, Zhou MM, Xu PX. 2017. EYA1's Conformation Specificity in Dephosphorylating Phosphothreonine in Myc and Its Activity on Myc Stabilization in Breast Cancer. *Mol Cell Biol* **37**.
- Li X, Chen F, Zhu Q, Ding B, Zhong Q, Huang K, Jiang X, Wang Z, Yin C, Zhu Y et al. 2016. Gli-1/PI3K/AKT/NF-kB pathway mediates resistance to radiation and is a target for reversion of responses in refractory acute myeloid leukemia cells. *Oncotarget* **7**: 33004-33015.
- Liau NPD, Laktyushin A, Morris R, Sandow JJ, Nicola NA, Kershaw NJ, Babon JJ. 2019. Enzymatic Characterization of Wild-Type and Mutant Janus Kinase 1. *Cancers (Basel)* **11**.
- Lim WA, Pawson T. 2010. Phosphotyrosine signaling: evolving a new cellular communication system. *Cell* **142**: 661-667.
- Lindsley RC, Saber W, Mar BG, Redd R, Wang T, Haagenson MD, Grauman PV, Hu ZH, Spellman SR, Lee SJ et al. 2017. Prognostic Mutations in Myelodysplastic Syndrome after Stem-Cell Transplantation. *N Engl J Med* **376**: 536-547.
- Liu S, Sun JP, Zhou B, Zhang ZY. 2006. Structural basis of docking interactions between ERK2 and MAP kinase phosphatase 3. *Proc Natl Acad Sci U S A* **103**: 5326-5331.
- Liu X, Yu R, Sun L, Garibotto G, Lin X, Wang Y, Thomas SS, Li R, Hu Z. 2017. The nuclear phosphatase SCP4 regulates FoxO transcription factors during muscle wasting in chronic kidney disease. *Kidney Int* **92**: 336-348.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.
- Lu X, Nguyen TA, Moon SH, Darlington Y, Sommer M, Donehower LA. 2008. The type 2C phosphatase Wip1: an oncogenic regulator of tumor suppressor and DNA damage response pathways. *Cancer Metastasis Rev* **27**: 123-135.
- Lu Z, Wang L, Dunaway-Mariano D, Allen KN. 2009. Structure-function analysis of 2-keto-3-deoxy-D-glycero-D-galactonononate-9-phosphate phosphatase defines specificity elements in type C0 haloalkanoate dehalogenase family members. *J Biol Chem* **284**: 1224-1233.
- Lucas CM, Harris RJ, Giannoudis A, Copland M, Slupsky JR, Clark RE. 2011. Cancerous inhibitor of PP2A (CIP2A) at diagnosis of chronic myeloid leukemia is a critical determinant of disease progression. *Blood* **117**: 6660-6668.
- Lucet IS, Fantino E, Styles M, Bamert R, Patel O, Broughton SE, Walter M, Burns CJ, Treutlein H, Wilks AF et al. 2006. The structural basis of Janus kinase 2 inhibition by a potent and specific pan-Janus kinase inhibitor. *Blood* **107**: 176-183.

REFERENCES

- Luo J, Solimini NL, Elledge SJ. 2009. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell* **136**: 823-837.
- Ma YN, Zhang X, Yu HC, Zhang JW. 2010. CTD small phosphatase like 2 (CTDSPL2) can increase epsilon- and gamma-globin gene expression in K562 cells and CD34+ cells derived from umbilical cord blood. *BMC Cell Biol* **11**: 75.
- Mackarehtschian K, Hardin JD, Moore KA, Boast S, Goff SP, Lemischka IR. 1995. Targeted disruption of the flk2/flt3 gene leads to deficiencies in primitive hematopoietic progenitors. *Immunity* **3**: 147-161.
- Mackay GM, Zheng L, van den Broek NJ, Gottlieb E. 2015. Analysis of Cell Metabolism Using LC-MS and Isotope Tracers. *Methods Enzymol* **561**: 171-196.
- Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD et al. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**: W636-W641.
- Malaise M, Steinbach D, Corbacioglu S. 2009. Clinical implications of c-Kit mutations in acute myelogenous leukemia. *Curr Hematol Malig Rep* **4**: 77-82.
- Malempati S, Tibbitts D, Cunningham M, Akkari Y, Olson S, Fan G, Sears RC. 2006. Aberrant stabilization of c-Myc protein in some lymphoblastic leukemias. *Leukemia* **20**: 1572-1581.
- Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. 2002. The protein kinase complement of the human genome. *Science* **298**: 1912-1934.
- Marcucci G, Geyer S, Laumann K, Zhao W, Bucci D, Uy GL, Blum W, Eisfeld AK, Pardee TS, Wang ES et al. 2020. Combination of dasatinib with chemotherapy in previously untreated core binding factor acute myeloid leukemia: CALGB 10801. *Blood Adv* **4**: 696-705.
- Martelli MP, Martino G, Cardinali V, Falini B, Martinelli G, Cerchione C. 2020. Enasidenib and ivosidenib in AML. *Minerva Med* **111**: 411-426.
- Matre P, Velez J, Jacamo R, Qi Y, Su X, Cai T, Chan SM, Lodi A, Sweeney SR, Ma H et al. 2016. Inhibiting glutaminase in acute myeloid leukemia: metabolic dependency of selected AML subtypes. *Oncotarget* **7**: 79722-79735.
- Matsumura I, Mizuki M, Kanakura Y. 2008. Roles for deregulated receptor tyrosine kinases and their downstream signaling molecules in hematologic malignancies. *Cancer Sci* **99**: 479-485.
- Matsuo R, Ochiai W, Nakashima K, Taga T. 2001. A new expression cloning strategy for isolation of substrate-specific kinases by using phosphorylation site-specific antibody. *J Immunol Methods* **247**: 141-151.
- Mazaloukas MD, Godoy-Ruiz R, Weber DJ, Zimmer DB, Honkanen RE, Wadzinski BE. 2014. Small G proteins Rac1 and Ras regulate serine/threonine protein phosphatase 5 (PP5).extracellular signal-regulated kinase (ERK) complexes involved in the feedback regulation of Raf1. *J Biol Chem* **289**: 4219-4232.
- McDonald ER, 3rd, de Weck A, Schlabach MR, Billy E, Mavrakis KJ, Hoffman GR, Belur D, Castelletti D, Frias E, Gampa K et al. 2017. Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening. *Cell* **170**: 577-592 e510.

REFERENCES

- McNerney ME, Godley LA, Le Beau MM. 2017. Therapy-related myeloid neoplasms: when genetics and environment collide. *Nat Rev Cancer* **17**: 513-527.
- Medeiros BC, Satram-Hoang S, Hurst D, Hoang KQ, Momin F, Reyes C. 2015. Big data analysis of treatment patterns and outcomes among elderly acute myeloid leukemia patients in the United States. *Ann Hematol* **94**: 1127-1138.
- Meng TC, Fukada T, Tonks NK. 2002. Reversible oxidation and inactivation of protein tyrosine phosphatases in vivo. *Mol Cell* **9**: 387-399.
- Merino D, Kelly GL, Lessene G, Wei AH, Roberts AW, Strasser A. 2018. BH3-Mimetic Drugs: Blazing the Trail for New Cancer Medicines. *Cancer Cell* **34**: 879-891.
- Meszaros B, Erdos G, Dosztanyi Z. 2018. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res* **46**: W329-W337.
- Metais JY, Doerfler PA, Mayuranathan T, Bauer DE, Fowler SC, Hsieh MM, Katta V, Keriwala S, Lazzarotto CR, Luk K et al. 2019. Genome editing of HBG1 and HBG2 to induce fetal hemoglobin. *Blood Adv* **3**: 3379-3392.
- Meyer C, Kowarz E, Hofmann J, Renneville A, Zuna J, Trka J, Ben Abdelali R, Macintyre E, De Braekeleer E, De Braekeleer M et al. 2009. New insights to the MLL recombinome of acute leukemias. *Leukemia* **23**: 1490-1499.
- Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, Dharia NV, Montgomery PG, Cowley GS, Pantel S et al. 2017. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* **49**: 1779-1784.
- Michowski W, Chick JM, Chu C, Kolodziejczyk A, Wang Y, Suski JM, Abraham B, Anders L, Day D, Dunkl LM et al. 2020. Cdk1 Controls Global Epigenetic Landscape in Embryonic Stem Cells. *Mol Cell* **78**: 459-476 e413.
- Miraki-Moud F, Ghazaly E, Ariza-McNaughton L, Hodby KA, Clear A, Anjos-Afonso F, Liapis K, Grantham M, Sohrabi F, Cavenagh J et al. 2015. Arginine deprivation using pegylated arginine deiminase has activity against primary acute myeloid leukemia cells in vivo. *Blood* **125**: 4060-4068.
- Miyamoto Y, Whiley PAF, Goh HY, Wong C, Higgins G, Tachibana T, McMenamin PG, Mayne L, Loveland KL. 2018. The STK35 locus contributes to normal gametogenesis and encodes a lncRNA responsive to oxidative stress. *Biol Open* **7**.
- Modi V, Dunbrack RL, Jr. 2019. Defining a new nomenclature for the structures of active and inactive kinases. *Proc Natl Acad Sci U S A* **116**: 6818-6827.
- Molica M, Mazzone C, Niscola P, de Fabritiis P. 2020. TP53 Mutations in Acute Myeloid Leukemia: Still a Daunting Challenge? *Front Oncol* **10**: 610820.
- Moorhead GB, De Wever V, Templeton G, Kerk D. 2009. Evolution of protein phosphatases in plants and animals. *Biochem J* **417**: 401-409.
- Morita K, He S, Nowak RP, Wang J, Zimmerman MW, Fu C, Durbin AD, Martel MW, Prutsch N, Gray NS et al. 2020. Allosteric Activators of Protein Phosphatase 2A Display Broad Antitumor Activity Mediated by Dephosphorylation of MYBL2. *Cell* **181**: 702-715 e720.

REFERENCES

- Morotti A, Panuzzo C, Fava C, Saglio G. 2014. Kinase-inhibitor-insensitive cancer stem cells in chronic myeloid leukemia. *Expert Opin Biol Ther* **14**: 287-299.
- Muller JP, Schmidt-Arras D. 2020. Novel Approaches to Target Mutant FLT3 Leukaemia. *Cancers (Basel)* **12**.
- Mulloy JC, Wunderlich M, Zheng Y, Wei J. 2008. Transforming human blood stem and progenitor cells: a new way forward in leukemia modeling. *Cell Cycle* **7**: 3314-3319.
- Mussai F, Egan S, Higginbotham-Jones J, Perry T, Beggs A, Odintsova E, Loke J, Pratt G, U KP, Lo A et al. 2015. Arginine dependence of acute myeloid leukemia blast proliferation: a novel therapeutic target. *Blood* **125**: 2386-2396.
- Nabet B, Roberts JM, Buckley DL, Paulk J, Dastjerdi S, Yang A, Leggett AL, Erb MA, Lawlor MA, Souza A et al. 2018. The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol* **14**: 431-441.
- Nabinger SC, Li XJ, Ramdas B, He Y, Zhang X, Zeng L, Richine B, Bowling JD, Fukuda S, Goenka S et al. 2013. The protein tyrosine phosphatase, Shp2, positively contributes to FLT3-ITD-induced hematopoietic progenitor hyperproliferation and malignant disease in vivo. *Leukemia* **27**: 398-408.
- Nasa I, Kettenbach AN. 2018. Coordination of Protein Kinase and Phosphoprotein Phosphatase Activities in Mitosis. *Front Cell Dev Biol* **6**: 30.
- Natsume T, Kiyomitsu T, Saga Y, Kanemaki MT. 2016. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep* **15**: 210-218.
- Nesti E, Corson GM, McCleskey M, Oyer JA, Mandel G. 2014. C-terminal domain small phosphatase 1 and MAP kinase reciprocally control REST stability and neuronal differentiation. *Proc Natl Acad Sci U S A* **111**: E3929-3936.
- Nesvizhskii AI, Keller A, Kolker E, Aebersold R. 2003. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* **75**: 4646-4658.
- Neviani P, Harb JG, Oaks JJ, Santhanam R, Walker CJ, Ellis JJ, Ferenchak G, Dorrance AM, Paisie CA, Eiring AM et al. 2013. PP2A-activating drugs selectively eradicate TKI-resistant chronic myeloid leukemic stem cells. *J Clin Invest* **123**: 4144-4157.
- Neviani P, Santhanam R, Oaks JJ, Eiring AM, Notari M, Blaser BW, Liu S, Trotta R, Muthusamy N, Gambacorti-Passerini C et al. 2007. FTY720, a new alternative for treating blast crisis chronic myelogenous leukemia and Philadelphia chromosome-positive acute lymphocytic leukemia. *J Clin Invest* **117**: 2408-2421.
- Neviani P, Santhanam R, Trotta R, Notari M, Blaser BW, Liu S, Mao H, Chang JS, Galietta A, Uttam A et al. 2005. The tumor suppressor PP2A is functionally inactivated in blast crisis CML through the inhibitory activity of the BCR/ABL-regulated SET protein. *Cancer Cell* **8**: 355-368.
- Nguyen Ba AN, Pogoutse A, Provart N, Moses AM. 2009. NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics* **10**: 202.
- Ni F, Yu WM, Li Z, Graham DK, Jin L, Kang S, Rossi MR, Li S, Broxmeyer HE, Qu CK. 2019. Critical role of ASCT2-mediated amino acid metabolism in promoting leukaemia development and progression. *Nat Metab* **1**: 390-403.

REFERENCES

- Nicklin P, Bergman P, Zhang B, Triantafellow E, Wang H, Nyfeler B, Yang H, Hild M, Kung C, Wilson C et al. 2009. Bidirectional transport of amino acids regulates mTOR and autophagy. *Cell* **136**: 521-534.
- Niemeyer CM, Flotho C. 2019. Juvenile myelomonocytic leukemia: who's the driver at the wheel? *Blood* **133**: 1060-1070.
- Ning ZQ, Li J, Arceci RJ. 2001. Signal transducer and activator of transcription 3 activation is required for Asp(816) mutant c-Kit-mediated cytokine-independent survival and proliferation in human leukemia cells. *Blood* **97**: 3559-3567.
- Nishi H, Shaytan A, Panchenko AR. 2014. Physicochemical mechanisms of protein regulation by phosphorylation. *Front Genet* **5**: 270.
- Nolen B, Taylor S, Ghosh G. 2004. Regulation of protein kinases; controlling activity through activation segment conformation. *Mol Cell* **15**: 661-675.
- Nowak DG, Katsenelson KC, Watrud KE, Chen M, Mathew G, D'Andrea VD, Lee MF, Swamynathan MM, Casanova-Salas I, Jibilian MC et al. 2019. The PHLPP2 phosphatase is a druggable driver of prostate cancer progression. *J Cell Biol* **218**: 1943-1957.
- Nowell PC. 1962. The minute chromosome (Ph1) in chronic granulocytic leukemia. *Blut* **8**: 65-66.
- O'Neil J, Look AT. 2007. Mechanisms of transcription factor deregulation in lymphoid cell transformation. *Oncogene* **26**: 6838-6849.
- Ochoa D, Jarnuczak AF, Vieitez C, Gehre M, Soucheray M, Mateus A, Kleefeldt AA, Hill A, Garcia-Alonso L, Stein F et al. 2020. The functional landscape of the human phosphoproteome. *Nat Biotechnol* **38**: 365-373.
- Ochocki JD, Khare S, Hess M, Ackerman D, Qiu B, Daisak JI, Worth AJ, Lin N, Lee P, Xie H et al. 2018. Arginase 2 Suppresses Renal Carcinoma Progression via Biosynthetic Cofactor Pyridoxal Phosphate Depletion and Increased Polyamine Toxicity. *Cell Metab* **27**: 1263-1280 e1266.
- Offman J, Opelz G, Doehler B, Cummins D, Halil O, Banner NR, Burke MM, Sullivan D, Macpherson P, Karran P. 2004. Defective DNA mismatch repair in acute myeloid leukemia/myelodysplastic syndrome after organ transplantation. *Blood* **104**: 822-828.
- Ohanian M, Rozovski U, Kanagal-Shamanna R, Abruzzo LV, Loghavi S, Kadia T, Futreal A, Bhalla K, Zuo Z, Huh YO et al. 2019. MYC protein expression is an important prognostic factor in acute myeloid leukemia. *Leuk Lymphoma* **60**: 37-48.
- Ok CY, Patel KP, Garcia-Manero G, Routbort MJ, Fu B, Tang G, Goswami M, Singh R, Kanagal-Shamanna R, Pierce SA et al. 2015. Mutational profiling of therapy-related myelodysplastic syndromes and acute myeloid leukemia by next generation sequencing, a comparison with de novo diseases. *Leuk Res* **39**: 348-354.
- Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, Su L, Xu G, Zhang Y. 2005. hDOT1L links histone methylation to leukemogenesis. *Cell* **121**: 167-178.
- Oliveria SF, Dell'Acqua ML, Sather WA. 2007. AKAP79/150 anchoring of calcineurin controls neuronal L-type Ca²⁺ channel activity and nuclear signaling. *Neuron* **55**: 261-275.
- Olsen JV, Blagoev B, Gnadt F, Macek B, Kumar C, Mortensen P, Mann M. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**: 635-648.

REFERENCES

- Oshima K, Zhao J, Perez-Duran P, Brown JA, Patino-Galindo JA, Chu T, Quinn A, Gunning T, Belver L, Ambesi-Impiombato A et al. 2020. Mutational and functional genetics mapping of chemotherapy resistance mechanisms in relapsed acute lymphoblastic leukemia. *Nat Cancer* **1**: 1113-1127.
- Ostergaard HL, Shackelford DA, Hurley TR, Johnson P, Hyman R, Sefton BM, Trowbridge IS. 1989. Expression of CD45 alters phosphorylation of the lck-encoded tyrosine protein kinase in murine lymphoma T-cell lines. *Proc Natl Acad Sci U S A* **86**: 8959-8963.
- Ostergaard P, Simpson MA, Connell FC, Steward CG, Brice G, Woollard WJ, Dafou D, Kilo T, Smithson S, Lunt P et al. 2011. Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nat Genet* **43**: 929-931.
- Ostman A, Yang Q, Tonks NK. 1994. Expression of DEP-1, a receptor-like protein-tyrosine-phosphatase, is enhanced with increasing cell density. *Proc Natl Acad Sci U S A* **91**: 9680-9684.
- Padua RAP, Sun Y, Marko I, Pitsawong W, Stiller JB, Otten R, Kern D. 2018. Mechanism of activating mutations and allosteric drug inhibition of the phosphatase SHP2. *Nat Commun* **9**: 4507.
- Pan R, Hogdal LJ, Benito JM, Bucci D, Han L, Borthakur G, Cortes J, DeAngelo DJ, Debose L, Mu H et al. 2014. Selective BCL-2 inhibition by ABT-199 causes on-target cell death in acute myeloid leukemia. *Cancer Discov* **4**: 362-375.
- Pandey RN, Rani R, Yeo EJ, Spencer M, Hu S, Lang RA, Hegde RS. 2010. The Eyes Absent phosphatase-transactivator proteins promote proliferation, transformation, migration, and invasion of tumor cells. *Oncogene* **29**: 3715-3722.
- Pandey RN, Wang TS, Tadjuidje E, McDonald MG, Rettie AE, Hegde RS. 2013. Structure-activity relationships of benzobromarone metabolites and derivatives as EYA inhibitory anti-angiogenic agents. *PLoS One* **8**: e84582.
- Park H, Jung SK, Yu KR, Kim JH, Kim YS, Ko JH, Park BC, Kim SJ. 2011. Structure-based virtual screening approach to the discovery of novel inhibitors of eyes absent 2 phosphatase with various metal chelating moieties. *Chem Biol Drug Des* **78**: 642-650.
- Park H, Lee HS, Ku B, Lee SR, Kim SJ. 2017. Two-track virtual screening approach to identify both competitive and allosteric inhibitors of human small C-terminal domain phosphatase 1. *J Comput Aided Mol Des* **31**: 743-753.
- Parsons JF, Lim K, Tempczyk A, Krajewski W, Eisenstein E, Herzberg O. 2002. From structure to function: YrbI from *Haemophilus influenzae* (HI1679) is a phosphatase. *Proteins* **46**: 393-404.
- Pastore F, Dufour A, Benthous T, Metzeler KH, Maharry KS, Schneider S, Ksienzyk B, Mellert G, Zellmeier E, Kakadia PM et al. 2014. Combined molecular and clinical prognostic index for relapse and survival in cytogenetically normal acute myeloid leukemia. *J Clin Oncol* **32**: 1586-1594.
- Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**: 417-419.
- Paulsson K, Horvat A, Strombeck B, Nilsson F, Heldrup J, Behrendtz M, Forestier E, Andersson A, Fioretos T, Johansson B. 2008. Mutations of FLT3, NRAS, KRAS, and PTPN11 are frequent and possibly mutually exclusive in high hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer* **47**: 26-33.

REFERENCES

- Pawson T, Nash P. 2000. Protein-protein interactions define specificity in signal transduction. *Genes Dev* **14**: 1027-1047.
- Pereira O, Teixeira A, Sampaio-Marques B, Castro I, Girao H, Ludovico P. 2018. Signalling mechanisms that regulate metabolic profile and autophagy of acute myeloid leukaemia cells. *J Cell Mol Med* **22**: 4807-4817.
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605-1612.
- Piao X, Paulson R, van der Geer P, Pawson T, Bernstein A. 1996. Oncogenic mutation in the Kit receptor tyrosine kinase alters substrate specificity and induces degradation of the protein tyrosine phosphatase SHP-1. *Proc Natl Acad Sci U S A* **93**: 14665-14669.
- Pignoni F, Hu B, Zavitz KH, Xiao J, Garrity PA, Zipursky SL. 1997. The eye-specification proteins So and Eya form a complex and regulate multiple steps in Drosophila eye development. *Cell* **91**: 881-891.
- Pingel JT, Thomas ML. 1989. Evidence that the leukocyte-common antigen is required for antigen-induced T lymphocyte proliferation. *Cell* **58**: 1055-1065.
- Placke T, Faber K, Nonami A, Putwain SL, Salih HR, Heidel FH, Kramer A, Root DE, Barbie DA, Krivtsov AV et al. 2014. Requirement for CDK6 in MLL-rearranged acute myeloid leukemia. *Blood* **124**: 13-23.
- Poon RY, Hunter T. 1995. Dephosphorylation of Cdk2 Thr160 by the cyclin-dependent kinase-interacting phosphatase KAP in the absence of cyclin. *Science* **270**: 90-93.
- Possemato R, Marks KM, Shaul YD, Pacold ME, Kim D, Birsoy K, Sethumadhavan S, Woo HK, Jang HG, Jha AK et al. 2011. Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476**: 346-350.
- Potter N, Miraki-Moud F, Ermini L, Titley I, Vijayaraghavan G, Papaemmanuil E, Campbell P, Gribben J, Taussig D, Greaves M. 2019. Single cell analysis of clonal architecture in acute myeloid leukaemia. *Leukemia* **33**: 1113-1123.
- Pulverer BJ, Fisher C, Vousden K, Littlewood T, Evan G, Woodgett JR. 1994. Site-specific modulation of c-Myc cotransformation by residues phosphorylated in vivo. *Oncogene* **9**: 59-70.
- Qian H, Ji C, Zhao S, Chen J, Jiang M, Zhang Y, Yan M, Zheng D, Sun Y, Xie Y et al. 2007. Expression and characterization of HSPC129, a RNA polymerase II C-terminal domain phosphatase. *Mol Cell Biochem* **303**: 183-188.
- Qu CK, Yu WM, Azzarelli B, Cooper S, Broxmeyer HE, Feng GS. 1998. Biased suppression of hematopoiesis and multiple developmental defects in chimeric mice containing Shp-2 mutant cells. *Mol Cell Biol* **18**: 6075-6082.
- Raffel S, Falcone M, Kneisel N, Hansson J, Wang W, Lutz C, Bullinger L, Poschet G, Nonnenmacher Y, Barnert A et al. 2017. BCAT1 restricts alphaKG levels in AML stem cells leading to IDHmut-like DNA hypermethylation. *Nature* **551**: 384-388.
- Rauen KA. 2013. The RASopathies. *Annu Rev Genomics Hum Genet* **14**: 355-369.

REFERENCES

- Reece KM, Mazalouskas MD, Wadzinski BE. 2009. The Balpha and Bdelta regulatory subunits of PP2A are necessary for assembly of the CaMKIV/PP2A signaling complex. *Biochem Biophys Res Commun* **386**: 582-587.
- Remenyi A, Good MC, Lim WA. 2006. Docking interactions in protein kinase and phosphatase networks. *Curr Opin Struct Biol* **16**: 676-685.
- Ren L, Chen X, Luechapanichkul R, Selner NG, Meyer TM, Wavreille AS, Chan R, Iorio C, Zhou X, Neel BG et al. 2011. Substrate specificity of protein tyrosine phosphatases 1B, RPTPalpha, SHP-1, and SHP-2. *Biochemistry* **50**: 2339-2356.
- Roberts KG, Smith AM, McDougall F, Carpenter H, Horan M, Neviani P, Powell JA, Thomas D, Guthridge MA, Perrotti D et al. 2010. Essential requirement for PP2A inhibition by the oncogenic receptor c-KIT suggests PP2A reactivation as a strategy to treat c-KIT+ cancers. *Cancer Res* **70**: 5438-5447.
- Robin TP, Smith A, McKinsey E, Reaves L, Jedlicka P, Ford HL. 2012. EWS/FLI1 regulates EYA3 in Ewing sarcoma via modulation of miRNA-708, resulting in increased cell survival and chemoresistance. *Mol Cancer Res* **10**: 1098-1108.
- Roboz GJ, Montesinos P, Selleslag D, Wei A, Jang JH, Falantes J, Voso MT, Sayar H, Porkka K, Marlton P et al. 2016. Design of the randomized, Phase III, QUAZAR AML Maintenance trial of CC-486 (oral azacitidine) maintenance therapy in acute myeloid leukemia. *Future Oncol* **12**: 293-302.
- Rossiter NJ, Huggler KS, Adelman CH, Keys HR, Soens RW, Sabatini DM, Cantor JR. 2020. CRISPR screens in physiologic medium reveal conditionally essential genes in human cells. *bioRxiv*: 2020.2008.2031.275107.
- Rothe K, Porter V, Jiang X. 2019. Current Outlook on Autophagy in Human Leukemia: Foe in Cancer Stem Cells and Drug Resistance, Friend in New Therapeutic Interventions. *Int J Mol Sci* **20**.
- Rowan FC, Richards M, Bibby RA, Thompson A, Bayliss R, Blagg J. 2013. Insights into Aurora-A kinase activation using unnatural amino acids incorporated by chemical modification. *ACS Chem Biol* **8**: 2184-2191.
- Ryan MB, Fede de la Cruz F, Phat S, Myers DT, Wong E, Shahzade HA, Hong CB, Corcoran RB. 2020. Vertical Pathway Inhibition Overcomes Adaptive Feedback Resistance to KRAS(G12C) Inhibition. *Clin Cancer Res* **26**: 1633-1643.
- Saint-Paul L, Nguyen CH, Buffiere A, Pais de Barros JP, Hammann A, Landras-Guetta C, Filomenko R, Chretien ML, Johnson P, Bastie JN et al. 2016. CD45 phosphatase is crucial for human and murine acute myeloid leukemia maintenance through its localization in lipid rafts. *Oncotarget* **7**: 64785-64797.
- Santos A, Wernersson R, Jensen LJ. 2015. Cyclebase 3.0: a multi-organism database on cell-cycle regulation and phenotypes. *Nucleic Acids Res* **43**: D1140-1144.
- Sarosiek KA, Letai A. 2016. Directly targeting the mitochondrial pathway of apoptosis for cancer therapy using BH3 mimetics - recent successes, current challenges and future promise. *FEBS J* **283**: 3523-3533.
- Sasaki M, Knobbe CB, Munger JC, Lind EF, Brenner D, Brustle A, Harris IS, Holmes R, Wakeham A, Haight J et al. 2012. IDH1(R132H) mutation increases murine haematopoietic progenitors and alters epigenetics. *Nature* **488**: 656-659.

REFERENCES

- Schmidt-Arras D, Bohmer FD. 2020. Mislocalisation of Activated Receptor Tyrosine Kinases - Challenges for Cancer Therapy. *Trends Mol Med* **26**: 833-847.
- Schnittger S, Bacher U, Haferlach C, Alpermann T, Kern W, Haferlach T. 2012. Diversity of the juxtamembrane and TKD1 mutations (exons 13-15) in the FLT3 gene with regards to mutant load, sequence, length, localization, and correlation with biological data. *Genes Chromosomes Cancer* **51**: 910-924.
- Schonfeld SJ, Gilbert ES, Dores GM, Lynch CF, Hodgson DC, Hall P, Storm H, Andersen A, Pukkala E, Holowaty E et al. 2006. Acute myeloid leukemia following Hodgkin lymphoma: a population-based study of 35,511 patients. *J Natl Cancer Inst* **98**: 215-218.
- Schuler A, Schwieger M, Engelmann A, Weber K, Horn S, Muller U, Arnold MA, Olson EN, Stocking C. 2008. The MADS transcription factor Mef2c is a pivotal modulator of myeloid cell fate. *Blood* **111**: 4532-4541.
- Schulze M, Fedorchenko O, Zink TG, Knobbe-Thomsen CB, Kraus S, Schwinn S, Beilhack A, Reifenberger G, Monoranu CM, Siren AL et al. 2016. Chronophin is a glial tumor modifier involved in the regulation of glioblastoma growth and invasiveness. *Oncogene* **35**: 3163-3177.
- Schwieger M, Schuler A, Forster M, Engelmann A, Arnold MA, Delwel R, Valk PJ, Lohler J, Slany RK, Olson EN et al. 2009. Homing and invasiveness of MLL/ENL leukemic cells is regulated by MEF2C. *Blood* **114**: 2476-2488.
- Seavitt JR, White LS, Murphy KM, Loh DY, Perlmutter RM, Thomas ML. 1999. Expression of the p56(Lck) Y505F mutation in CD45-deficient mice rescues thymocyte development. *Mol Cell Biol* **19**: 4200-4208.
- Seifried A, Schultz J, Gohla A. 2013. Human HAD phosphatases: structure, mechanism, and roles in health and disease. *FEBS J* **280**: 549-571.
- Selner NG, Luechapanichkul R, Chen X, Neel BG, Zhang ZY, Knapp S, Bell CE, Pei D. 2014. Diverse levels of sequence selectivity and catalytic efficiency of protein-tyrosine phosphatases. *Biochemistry* **53**: 397-412.
- Seo J, Fortunato ES, 3rd, Suh JM, Stenesen D, Tang W, Parks EJ, Adams CM, Townes T, Graff JM. 2009. Atf4 regulates obesity, glucose homeostasis, and energy expenditure. *Diabetes* **58**: 2565-2573.
- Shah NP. 2008. Advanced CML: therapeutic options for patients in accelerated and blast phases. *J Natl Compr Canc Netw* **6 Suppl 2**: S31-S36.
- Shallis RM, Wang R, Davidoff A, Ma X, Zeidan AM. 2019. Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. *Blood Rev* **36**: 70-87.
- Shallis RM, Weiss JJ, Deziel NC, Gore SD. 2020. Challenging the concept of de novo acute myeloid leukemia: Environmental and occupational leukemogens hiding in our midst. *Blood Rev*: 100760.
- Sharif T, Martell E, Dai C, Singh SK, Gujar S. 2019. Regulation of the proline regulatory axis and autophagy modulates stemness in TP73/p73 deficient cancer stem-like cells. *Autophagy* **15**: 934-936.
- Shen C, Ipsaro JJ, Shi J, Milazzo JP, Wang E, Roe JS, Suzuki Y, Pappin DJ, Joshua-Tor L, Vakoc CR. 2015. NSD3-Short Is an Adaptor Protein that Couples BRD4 to the CHD8 Chromatin Remodeler. *Mol Cell* **60**: 847-859.

REFERENCES

- Shen T, Sun C, Zhang Z, Xu N, Duan X, Feng XH, Lin X. 2014. Specific control of BMP signaling and mesenchymal differentiation by cytoplasmic phosphatase PPM1H. *Cell Res* **24**: 727-741.
- Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB, Vakoc CR. 2015. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat Biotechnol* **33**: 661-667.
- Shtivelman E, Lifshitz B, Gale RP, Canaani E. 1985. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature* **315**: 550-554.
- Sitnicka E, Bryder D, Theilgaard-Monch K, Buza-Vidas N, Adolfsson J, Jacobsen SE. 2002. Key role of flt3 ligand in regulation of the common lymphoid progenitor but not in maintenance of the hematopoietic stem cell pool. *Immunity* **17**: 463-472.
- Smith AM, Dun MD, Lee EM, Harrison C, Kahl R, Flanagan H, Panicker N, Mashkani B, Don AS, Morris J et al. 2016. Activation of protein phosphatase 2A in FLT3+ acute myeloid leukemia cells enhances the cytotoxicity of FLT3 tyrosine kinase inhibitors. *Oncotarget* **7**: 47465-47478.
- Smith CC, Wang Q, Chin CS, Salerno S, Damon LE, Levis MJ, Perl AE, Travers KJ, Wang S, Hunt JP et al. 2012. Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature* **485**: 260-263.
- Smith ML, Cavenagh JD, Lister TA, Fitzgibbon J. 2004. Mutation of CEBPA in familial acute myeloid leukemia. *N Engl J Med* **351**: 2403-2407.
- Song H, Hanlon N, Brown NR, Noble ME, Johnson LN, Barford D. 2001. Phosphoprotein-protein interactions revealed by the crystal structure of kinase-associated phosphatase in complex with phosphoCDK2. *Mol Cell* **7**: 615-626.
- Song WJ, Sullivan MG, Legare RD, Hutchings S, Tan X, Kufirin D, Ratajczak J, Resende IC, Haworth C, Hock R et al. 1999. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet* **23**: 166-175.
- Spencer DH, Russler-Germain DA, Ketkar S, Helton NM, Lamprecht TL, Fulton RS, Fronick CC, O'Laughlin M, Heath SE, Shinawi M et al. 2017. CpG Island Hypermethylation Mediated by DNMT3A Is a Consequence of AML Progression. *Cell* **168**: 801-816 e813.
- Steensma DP. 2018. Clinical consequences of clonal hematopoiesis of indeterminate potential. *Blood Adv* **2**: 3404-3410.
- Su YA, Lee MM, Hutter CM, Meltzer PS. 1997. Characterization of a highly conserved gene (OS4) amplified with CDK4 in human sarcomas. *Oncogene* **15**: 1289-1294.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**: 15545-15550.
- Sun J, Pedersen M, Ronnstrand L. 2009. The D816V mutation of c-Kit circumvents a requirement for Src family kinases in c-Kit signal transduction. *J Biol Chem* **284**: 11039-11047.
- Szomolay B, Shahrezaei V. 2012. Bell-shaped and ultrasensitive dose-response in phosphorylation-dephosphorylation cycles: the role of kinase-phosphatase complex formation. *BMC Syst Biol* **6**: 26.

REFERENCES

- Taatjes DJ, Sobel BE, Budd RC. 2008. Morphological and cytochemical determination of cell death by apoptosis. *Histochem Cell Biol* **129**: 33-43.
- Tadjuidje E, Wang TS, Pandey RN, Sumanas S, Lang RA, Hegde RS. 2012. The EYA tyrosine phosphatase activity is pro-angiogenic and is inhibited by benzbromarone. *PLoS One* **7**: e34806.
- Takahashi S. 2011. Downstream molecular pathways of FLT3 in the pathogenesis of acute myeloid leukemia: biology and therapeutic implications. *J Hematol Oncol* **4**: 13.
- Takeda H, Jenkins NA, Copeland NG. 2021. Identification of cancer driver genes using Sleeping Beauty transposon mutagenesis. *Cancer Sci*.
- Tarlock K, Alonzo TA, Wang YC, Gerbing RB, Ries R, Loken MR, Pardo L, Hylkema T, Joaquin J, Sarukkai L et al. 2019. Functional Properties of KIT Mutations Are Associated with Differential Clinical Outcomes and Response to Targeted Therapeutics in CBF Acute Myeloid Leukemia. *Clin Cancer Res* **25**: 5038-5048.
- Tarumoto Y, Lin S, Wang J, Milazzo JP, Xu Y, Lu B, Yang Z, Wei Y, Polyanskaya S, Wunderlich M et al. 2020. Salt-inducible kinase inhibition suppresses acute myeloid leukemia progression in vivo. *Blood* **135**: 56-70.
- Tarumoto Y, Lu B, Somerville TDD, Huang YH, Milazzo JP, Wu XS, Klingbeil O, El Demerdash O, Shi J, Vakoc CR. 2018. LKB1, Salt-Inducible Kinases, and MEF2C Are Linked Dependencies in Acute Myeloid Leukemia. *Mol Cell* **69**: 1017-1027 e1016.
- Ten Eyck LF, Taylor SS, Kornev AP. 2008. Conserved spatial patterns across the protein kinase family. *Biochim Biophys Acta* **1784**: 238-243.
- Thomas ML, Barclay AN, Gagnon J, Williams AF. 1985. Evidence from cDNA clones that the rat leukocyte-common antigen (T200) spans the lipid bilayer and contains a cytoplasmic domain of 80,000 Mr. *Cell* **41**: 83-93.
- Thomas ML, Brown EJ. 1999. Positive and negative regulation of Src-family membrane kinases by CD45. *Immunol Today* **20**: 406-411.
- Tian B, Yang J, Brasier AR. 2012. Two-step cross-linking for analysis of protein-chromatin interactions. *Methods Mol Biol* **809**: 105-120.
- Timm T, Balusamy K, Li X, Biernat J, Mandelkow E, Mandelkow EM. 2008. Glycogen synthase kinase (GSK) 3beta directly phosphorylates Serine 212 in the regulatory loop and inhibits microtubule affinity-regulating kinase (MARK) 2. *J Biol Chem* **283**: 18873-18882.
- Tonks NK, Charbonneau H, Diltz CD, Fischer EH, Walsh KA. 1988a. Demonstration that the leukocyte common antigen CD45 is a protein tyrosine phosphatase. *Biochemistry* **27**: 8695-8701.
- Tonks NK, Diltz CD, Fischer EH. 1988b. Purification of the major protein-tyrosine-phosphatases of human placenta. *J Biol Chem* **263**: 6722-6730.
- Tonks NK. 2003. PTP1B: from the sidelines to the front lines! *FEBS Lett* **546**: 140-148.
- Tonks NK. 2005. Redox redux: revisiting PTPs and the control of cell signaling. *Cell* **121**: 667-670.

REFERENCES

- Tonks NK. 2013. Protein tyrosine phosphatases--from housekeeping enzymes to master regulators of signal transduction. *FEBS J* **280**: 346-378.
- Tootle TL, Silver SJ, Davies EL, Newman V, Latek RR, Mills IA, Selengut JD, Parlikar BE, Rebay I. 2003. The transcription factor Eyes absent is a protein tyrosine phosphatase. *Nature* **426**: 299-302.
- Topka S, Vijai J, Walsh MF, Jacobs L, Maria A, Villano D, Gaddam P, Wu G, McGee RB, Quinn E et al. 2015. Germline ETV6 Mutations Confer Susceptibility to Acute Lymphoblastic Leukemia and Thrombocytopenia. *PLoS Genet* **11**: e1005262.
- Torrence ME, MacArthur MR, Hosios AM, Valvezan AJ, Asara JM, Mitchell JR, Manning BD. 2021. The mTORC1-mediated activation of ATF4 promotes protein and glutathione synthesis downstream of growth signals. *Elife* **10**.
- Tsai CH, Hou HA, Tang JL, Liu CY, Lin CC, Chou WC, Tseng MH, Chiang YC, Kuo YY, Liu MC et al. 2016. Genetic alterations and their clinical implications in older patients with acute myeloid leukemia. *Leukemia* **30**: 1485-1492.
- Tweedie S, Braschi B, Gray K, Jones TEM, Seal RL, Yates B, Bruford EA. 2021. [Genenames.org](https://www.genenames.org/): the HGNC and VGNC resources in 2021. *Nucleic Acids Res* **49**: D939-D946.
- Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A et al. 2015. Proteomics. Tissue-based map of the human proteome. *Science* **347**: 1260419.
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Res* **40**: e115.
- Uras IZ, Sexl V, Kollmann K. 2020. CDK6 Inhibition: A Novel Approach in AML Management. *Int J Mol Sci* **21**.
- Uren AG, Kool J, Berns A, van Lohuizen M. 2005. Retroviral insertional mutagenesis: past, present and future. *Oncogene* **24**: 7656-7672.
- Vainonen JP, Momeny M, Westermarck J. 2021. Druggable cancer phosphatases. *Sci Transl Med* **13**.
- Valencia-Sama I, Ladumor Y, Kee L, Adderley T, Christopher G, Robinson CM, Kano Y, Ohh M, Irwin MS. 2020. NRAS Status Determines Sensitivity to SHP2 Inhibitor Combination Therapies Targeting the RAS-MAPK Pathway in Neuroblastoma. *Cancer Res* **80**: 3413-3423.
- van Galen P, Hovestadt V, Wadsworth Ii MH, Hughes TK, Griffin GK, Battaglia S, Verga JA, Stephansky J, Pastika TJ, Lombardi Story J et al. 2019. Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity. *Cell* **176**: 1265-1281 e1224.
- Venugopal S, Mascarenhas J. 2020. Novel therapeutics in myeloproliferative neoplasms. *J Hematol Oncol* **13**: 162.
- von Kriegsheim A, Pitt A, Grindlay GJ, Kolch W, Dhillon AS. 2006. Regulation of the Raf-MEK-ERK pathway by protein phosphatase 5. *Nat Cell Biol* **8**: 1011-1016.
- Walkinshaw DR, Weist R, Kim GW, You L, Xiao L, Nie J, Li CS, Zhao S, Xu M, Yang XJ. 2013. The tumor suppressor kinase LKB1 activates the downstream kinases SIK2 and SIK3 to stimulate nuclear export of class IIa histone deacetylases. *J Biol Chem* **288**: 9345-9362.

REFERENCES

- Walter RB, Appelbaum FR, Estey EH, Bernstein ID. 2012. Acute myeloid leukemia stem cells and CD33-targeted immunotherapy. *Blood* **119**: 6198-6208.
- Walz C, Ahmed W, Lazarides K, Betancur M, Patel N, Hennighausen L, Zaleskas VM, Van Etten RA. 2012. Essential role for Stat5a/b in myeloproliferative neoplasms induced by BCR-ABL1 and JAK2(V617F) in mice. *Blood* **119**: 3550-3560.
- Wang BJ, Tang W, Zhang P, Wei Q. 2011. Tyr306 near the C-terminus of protein phosphatase-1 affects enzyme stability and inhibitor binding. *IUBMB Life* **63**: 574-581.
- Wang M, Yang C, Zhang L, Schaar DG. 2017. Molecular Mutations and Their Cooccurrences in Cytogenetically Normal Acute Myeloid Leukemia. *Stem Cells Int* **2017**: 6962379.
- Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM. 2015. Identification and characterization of essential genes in the human genome. *Science* **350**: 1096-1101.
- Wang W, Liao P, Shen M, Chen T, Chen Y, Li Y, Lin X, Ge X, Wang P. 2016a. SCP1 regulates c-Myc stability and functions through dephosphorylating c-Myc Ser62. *Oncogene* **35**: 491-500.
- Wang Y, Tadjuidje E, Pandey RN, Stefater JA, 3rd, Smith LE, Lang RA, Hegde RS. 2016. The Eyes Absent Proteins in Developmental and Pathological Angiogenesis. *Am J Pathol* **186**: 568-578.
- Wang Y, Tadjuidje E, Pandey RN, Stefater JA, 3rd, Smith LE, Lang RA, Hegde RS. 2016b. The Eyes Absent Proteins in Developmental and Pathological Angiogenesis. *Am J Pathol* **186**: 568-578.
- Wani S, Sugita A, Ohkuma Y, Hirose Y. 2016. Human SCP4 is a chromatin-associated CTD phosphatase and exhibits the dynamic translocation during erythroid differentiation. *J Biochem* **160**: 111-120.
- Warburg O. 1956. On the origin of cancer cells. *Science* **123**: 309-314.
- Watson AS, Riffelmacher T, Stranks A, Williams O, De Boer J, Cain K, MacFarlane M, McGouran J, Kessler B, Khandwala S et al. 2015. Autophagy limits proliferation and glycolytic metabolism in acute myeloid leukemia. *Cell Death Discov* **1**.
- Wei J, Wunderlich M, Fox C, Alvarez S, Cigudosa JC, Wilhelm JS, Zheng Y, Cancelas JA, Gu Y, Jansen M et al. 2008. Microenvironment determines lineage fate in a human model of MLL-AF9 leukemia. *Cancer Cell* **13**: 483-495.
- Wei Y, Iyer SV, Costa ASH, Yang Z, Kramer M, Adelman ER, Klingbeil O, Demerdash OE, Polyanskaya S, Chang K et al. 2020. *In vivo* genetic screen identifies a SLC5A3-dependent myo-inositol auxotrophy in acute myeloid leukemia. *bioRxiv*: 2020.2012.2022.424018.
- Weis K. 2003. Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. *Cell* **112**: 441-451.
- Wen Z, Zhong Z, Darnell JE, Jr. 1995. Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. *Cell* **82**: 241-250.
- Westphal RS, Anderson KA, Means AR, Wadzinski BE. 1998. A signaling complex of Ca²⁺-calmodulin-dependent protein kinase IV and protein phosphatase 2A. *Science* **280**: 1258-1261.

REFERENCES

- Westphal RS, Tavalin SJ, Lin JW, Alto NM, Fraser ID, Langeberg LK, Sheng M, Scott JD. 1999. Regulation of NMDA receptors by an associated phosphatase-kinase signaling complex. *Science* **285**: 93-96.
- Willems L, Jacque N, Jacquelin A, Neveux N, Maciel TT, Lambert M, Schmitt A, Poulain L, Green AS, Uzunov M et al. 2013. Inhibiting glutamine uptake represents an attractive new strategy for treating acute myeloid leukemia. *Blood* **122**: 3521-3532.
- Wolf I, Rubinfeld H, Yoon S, Marmor G, Hanoch T, Seger R. 2001. Involvement of the activation loop of ERK in the detachment from cytosolic anchoring. *J Biol Chem* **276**: 24490-24497.
- Wong AS, Choi GC, Cui CH, Pregernig G, Milani P, Adam M, Perli SD, Kazer SW, Gaillard A, Hermann M et al. 2016. Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. *Proc Natl Acad Sci U S A* **113**: 2544-2549.
- Wood V, Lock A, Harris MA, Rutherford K, Bahler J, Oliver SG. 2019. Hidden in plain sight: what remains to be discovered in the eukaryotic proteome? *Open Biol* **9**: 180241.
- Wu J, Woodard RW. 2003. Escherichia coli YrbI is 3-deoxy-D-manno-octulosonate 8-phosphate phosphatase. *J Biol Chem* **278**: 18117-18123.
- Wu Z, Liu J, Hu S, Zhu Y, Li S. 2018. Serine/Threonine Kinase 35, a Target Gene of STAT3, Regulates the Proliferation and Apoptosis of Osteosarcoma Cells. *Cell Physiol Biochem* **45**: 808-818.
- Xiao S, Li S, Yuan Z, Zhou L. 2020. Pyrroline-5-carboxylate reductase 1 (PYCR1) upregulation contributes to gastric cancer progression and indicates poor survival outcome. *Ann Transl Med* **8**: 937.
- Xie T, Peng W, Yan C, Wu J, Gong X, Shi Y. 2013. Structural insights into RIP3-mediated necroptotic signaling. *Cell Rep* **5**: 70-78.
- Xie T, Saleh T, Rossi P, Kalodimos CG. 2020. Conformational states dynamically populated by a kinase determine its function. *Science* **370**.
- Xiong W, Dabbouseh NM, Rebay I. 2009. Interactions with the Abelson tyrosine kinase reveal compartmentalization of eyes absent function between nucleus and cytoplasm. *Dev Cell* **16**: 271-279.
- Xu Y, Milazzo JP, Somerville TDD, Tarumoto Y, Huang YH, Ostrander EL, Wilkinson JE, Challen GA, Vakoc CR. 2018. A TFIID-SAGA Perturbation that Targets MYB and Suppresses Acute Myeloid Leukemia. *Cancer Cell* **33**: 13-28 e18.
- Yang H, Zhu J, Wang G, Liu H, Zhou Y, Qian J. 2020. STK35 Is Ubiquitinated by NEDD4L and Promotes Glycolysis and Inhibits Apoptosis Through Regulating the AKT Signaling Pathway, Influencing Chemoresistance of Colorectal Cancer. *Front Cell Dev Biol* **8**: 582695.
- Yang Z, Wu XS, Wei Y, Polyanskaya SA, Iyer SV, Jung M, Lach FP, Adelman ER, Klingbeil O, Milazzo JP et al. 2021. Transcriptional silencing of ALDH2 confers a dependency on Fanconi anemia proteins in acute myeloid leukemia. *Cancer Discov*.
- Ye H, Adane B, Khan N, Alexeev E, Nusbacher N, Minhajuddin M, Stevens BM, Winters AC, Lin X, Ashton JM et al. 2018. Subversion of Systemic Glucose Metabolism as a Mechanism to Support the Growth of Leukemia Cells. *Cancer Cell* **34**: 659-673 e656.

REFERENCES

- Yeh E, Cunningham M, Arnold H, Chasse D, Monteith T, Ivaldi G, Hahn WC, Stukenberg PT, Shenolikar S, Uchida T et al. 2004. A signalling pathway controlling c-Myc degradation that impacts oncogenic transformation of human cells. *Nat Cell Biol* **6**: 308-318.
- Yeo M, Lee SK, Lee B, Ruiz EC, Pfaff SL, Gill GN. 2005. Small CTD phosphatases function in silencing neuronal gene expression. *Science* **307**: 596-600.
- Yoshimoto G, Miyamoto T, Jabbarzadeh-Tabrizi S, Iino T, Rocnik JL, Kikushige Y, Mori Y, Shima T, Iwasaki H, Takenaka K et al. 2009. FLT3-ITD up-regulates MCL-1 to promote survival of stem cells in acute myeloid leukemia via FLT3-ITD-specific STAT5 activation. *Blood* **114**: 5034-5043.
- Yuan B, Cheng L, Chiang HC, Xu X, Han Y, Su H, Wang L, Zhang B, Lin J, Li X et al. 2014. A phosphotyrosine switch determines the antitumor activity of ERbeta. *J Clin Invest* **124**: 3378-3390.
- Yun S, Sharma R, Chan O, Vincelette ND, Sallman DA, Sweet K, Padron E, Komrokji R, Lancet JE, Abraham I et al. 2019. Prognostic significance of MYC oncoprotein expression on survival outcome in patients with acute myeloid leukemia with myelodysplasia related changes (AML-MRC). *Leuk Res* **84**: 106194.
- Zahreddine HA, Culjkovic-Kraljacic B, Assouline S, Gendron P, Romeo AA, Morris SJ, Cormack G, Jaquith JB, Cerchietti L, Cocolakis E et al. 2014. The sonic hedgehog factor GLI1 imparts drug resistance through inducible glucuronidation. *Nature* **511**: 90-93.
- Zhang L, Yang N, Huang J, Buckanovich RJ, Liang S, Barchetti A, Vezzani C, O'Brien-Jenkins A, Wang J, Ward MR et al. 2005. Transcriptional coactivator *Drosophila* eyes absent homologue 2 is up-regulated in epithelial ovarian cancer and promotes tumor growth. *Cancer Res* **65**: 925-932.
- Zhang L, Zhou H, Li X, Vartuli RL, Rowse M, Xing Y, Rudra P, Ghosh D, Zhao R, Ford HL. 2018. Eya3 partners with PP2A to induce c-Myc stabilization and tumor progression. *Nat Commun* **9**: 1047.
- Zhang M, Cho EJ, Burstein G, Siegel D, Zhang Y. 2011. Selective inactivation of a human neuronal silencing phosphatase by a small molecule inhibitor. *ACS Chem Biol* **6**: 511-519.
- Zhang QH, Ye M, Wu XY, Ren SX, Zhao M, Zhao CJ, Fu G, Shen Y, Fan HY, Lu G et al. 2000. Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells. *Genome Res* **10**: 1546-1560.
- Zhang S, Broxmeyer HE. 1999. p85 subunit of PI3 kinase does not bind to human Flt3 receptor, but associates with SHP2, SHIP, and a tyrosine-phosphorylated 100-kDa protein in Flt3 ligand-stimulated hematopoietic cells. *Biochem Biophys Res Commun* **254**: 440-445.
- Zhang S, H E Broxmeyer HE. 2000. Flt3 ligand induces tyrosine phosphorylation of gab1 and gab2 and their association with shp-2, grb2, and PI3 kinase. *Biochem Biophys Res Commun* **277**: 195-199.
- Zhang Y, Kim Y, Genoud N, Gao J, Kelly JW, Pfaff SL, Gill GN, Dixon JE, Noel JP. 2006. Determinants for dephosphorylation of the RNA polymerase II C-terminal domain by Scp1. *Mol Cell* **24**: 759-770.
- Zhao C, Chen A, Jamieson CH, Fereshteh M, Abrahamsson A, Blum J, Kwon HY, Kim J, Chute JP, Rizzieri D et al. 2009. Hedgehog signalling is essential for maintenance of cancer stem cells in myeloid leukaemia. *Nature* **458**: 776-779.

REFERENCES

- Zhao Y, Xiao M, Sun B, Zhang Z, Shen T, Duan X, Yu PB, Feng XH, Lin X. 2014. C-terminal domain (CTD) small phosphatase-like 2 modulates the canonical bone morphogenetic protein (BMP) signaling and mesenchymal differentiation via Smad dephosphorylation. *J Biol Chem* **289**: 26441-26450.
- Zhong X, Prinz A, Steger J, Garcia-Cuellar MP, Radsak M, Bentaher A, Slany RK. 2018. HoxA9 transforms murine myeloid cells by a feedback loop driving expression of key oncogenes and cell cycle control genes. *Blood Adv* **2**: 3137-3148.
- Zhou B, Wu L, Shen K, Zhang J, Lawrence DS, Zhang ZY. 2001. Multiple regions of MAP kinase phosphatase 3 are involved in its recognition and activation by ERK2. *J Biol Chem* **276**: 6506-6515.
- Zhou B, Zhang J, Liu S, Reddy S, Wang F, Zhang ZY. 2006. Mapping ERK2-MKP3 binding interfaces by hydrogen/deuterium exchange mass spectrometry. *J Biol Chem* **281**: 38834-38844.
- Zhou T, Medeiros LJ, Hu S. 2018. Chronic Myeloid Leukemia: Beyond BCR-ABL1. *Curr Hematol Malig Rep* **13**: 435-445.
- Zuber J, McJunkin K, Fellmann C, Dow LE, Taylor MJ, Hannon GJ, Lowe SW. 2011. Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nat Biotechnol* **29**: 79-83.
- Zuber J, McJunkin K, Fellmann C, Dow LE, Taylor MJ, Hannon GJ, Lowe SW. 2011a. Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nat Biotechnol* **29**: 79-83.
- Zuber J, Radtke I, Pardee TS, Zhao Z, Rappaport AR, Luo W, McCurrach ME, Yang MM, Dolan ME, Kogan SC et al. 2009. Mouse models of human AML accurately predict chemotherapy response. *Genes Dev* **23**: 877-889.
- Zuber J, Rappaport AR, Luo W, Wang E, Chen C, Vaseva AV, Shi J, Weissmueller S, Fellmann C, Taylor MJ et al. 2011b. An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance. *Genes Dev* **25**: 1628-1640.
- Zuber J, Shi J, Wang E, Rappaport AR, Herrmann H, Sison EA, Magoon D, Qi J, Blatt K, Wunderlich M et al. 2011c. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**: 524-528.

APPENDICES

**Appendix 1. Categorization of human phosphatases targeted by
CRISPR-Cas9 genetic screens in this study**

Metacategory	# of genes	HUGO Group Name	# of genes	Genes
Acid phosphatases	7	Acid phosphatases	7	<i>ACP1, ACP2, ACP3, ACP4, ACP5, ACP6, PXYLPI</i>
Alkaline phosphatases	4	Alkaline phosphatases	4	<i>ALPG, ALPI, ALPL, ALPP</i>
Dual specificity phosphatases (DUSP)	40	Dual specificity phosphatases (DUSP)	40	<i>CDC14A, CDC14B, DUSP1, DUSP10, DUSP11, DUSP12, DUSP13, DUSP14, DUSP15, DUSP16, DUSP18, DUSP19, DUSP2, DUSP21, DUSP22, DUSP23, DUSP26, DUSP28, DUSP29, DUSP3, DUSP4, DUSP5, DUSP6, DUSP7, DUSP8, DUSP9, EPM2A, PTEN, PTP4A3, PTPDC1, PTPMT1, RINGTT, SSH1, SSH2, SSH3, STYX, STYXL1, TNS1, TNS2, TPTE</i>
HAD Asp-based protein phosphatases	13	HAD Asp-based protein phosphatases	13	<i>CTDP1, CTDSP1, CTDSP2, CTDSPL, CTDSPL2, EYA1, EYA2, EYA3, EYA4, MDP1, PDXP, TIMM50, UBLCP1</i>
HAD Asp-based non-protein phosphatases	21	5'-nucleotidases	7	<i>NT5C, NT5C1A, NT5C1B, NT5C2, NT5C3A, NT5E, NT5M</i>
		HAD Asp-based non-protein phosphatases	11	<i>ENOPH1, EPHX2, LHPP, NANP, PHOSPHO1, PHOSPHO2, PMM1, PMM2, PNKP, PSPH, PUDP</i>
		Lipid phosphatases	3	<i>LPIN1, LPIN2, LPIN3</i>
Lipid phosphatases	46	Lipid phosphatases	46	<i>BPNT1, DOLPP1, FIG4, IMPA1, IMPA2, INPP1, INPP4B, INPP5A, INPP5B, INPP5D, INPP5E, INPP5F, INPP5J, INPP5K, INPPL1, MINPP1, MTM1, MTMR1, MTMR10, MTMR11, MTMR12, MTMR13, MTMR14, MTMR2, MTMR3, MTMR4, MTMR5, MTMR6, MTMR7, MTMR8, MTMR9, OCRL, PIP4P1, PIP4P2, PLPP1, PLPP2, PLPP3, PLPP4, PLPP5, PLPP6, PLPP7, SACM1L, SGPPP1, SGPPP2, SYNJ1, SYNJ2</i>
Other protein phosphatases	5	CDC25 phosphatases	3	<i>CDC25A, CDC25B, CDC25C</i>
		Protein phosphatases	2	<i>PHPT1, UBASH3B</i>
Serine/threonine phosphatases	32	Serine/threonine phosphatases	32	<i>PGAM5, SSU72, ILKAP, PDP1, PDP2, PHLPP1, PHLPP2, PPEF1, PPEF2, PPM1A, PPM1B, PPM1D, PPM1E, PPM1F, PPM1G, PPM1H, PPM1J, PPM1K, PPM1L, PPM1M, PPM1N, PPP1CA, PPP1CB, PPP1CC, PPP2CA, PPP2CB, PPP3CA, PPP3CB, PPP3CC, PPP4C, PPP5C, PPP6C</i>
Sugar phosphatases	13	Sugar phosphatases	13	<i>BPGM, FBP1, FBP2, G6PC1, G6PC2, G6PC3, PFKFB1, PFKFB2, PFKFB3, PFKFB4, PGAM2, PGAM4, TIGAR</i>
Protein tyrosine phosphatases	36	Protein tyrosine phosphatases	36	<i>PTPN1, PTPN11, PTPN12, PTPN13, PTPN14, PTPN18, PTPN2, PTPN21, PTPN22, PTPN23, PTPN3, PTPN4, PTPN5, PTPN6, PTPN7, PTPN9, PTPRA, PTPRB, PTPRC, PTPRD, PTPRE, PTPRF, PTPRG, PTPRH, PTPRJ, PTPRK, PTPRM, PTPRN, PTPRN2, PTPRO, PTPRQ, PTPRR, PTPRS, PTPRT, PTPRU, PTPRZ1</i>

Appendix 2. Average log₂ FC in phosphatase domain-focused screen (1/7)

Disease	AML	CML	SCLC	PDAC	RMS	RMS						
RRID	CVCL_2119	CVCL_0095	CVCL_1609	CVCL_0001	CVCL_1844	CVCL_0007	CVCL_0589	CVCL_0004	CVCL_1453	CVCL_0152	CVCL_0041	CVCL_1649
Cell_line	MOLM-13	SEM	NOMO-1	HEL	OCI-AML-3	U-937	Kasumi-1	K-562	NCI-H1048	AsPC-1	Rh30	RD
ACP1	-0.04	0.14	-0.01	0.48	-0.04	0.36	0.36	0.44	-0.28	0.12	0.86	0.10
ACP2	0.26	-0.25	-0.23	0.19	-0.19	-0.16	-0.02	-0.42	-0.26	0.04	-0.18	-0.08
ACP5	0.22	-0.28	-0.18	-0.23	-0.33	-0.16	0.13	-0.39	-0.67	-0.12	0.89	0.24
ACP6	0.35	0.06	0.41	0.10	-0.22	0.36	0.42	0.12	0.28	0.16	-0.26	0.15
ACPP	0.12	0.15	-0.01	0.73	0.09	0.12	0.27	0.10	0.19	0.18	-0.04	0.42
ACPT	-0.50	-0.08	-0.33	0.70	-0.13	0.00	-0.10	0.29	0.29	0.06	0.20	-0.06
ALPI	-0.21	-0.09	-0.15	0.22	-0.39	-0.02	-0.16	-0.31	-0.81	-0.11	-0.57	-0.29
ALPL	-0.20	-0.12	-0.13	0.44	0.26	0.15	-0.15	-0.21	0.04	0.28	0.30	-0.58
ALPP	0.10	0.16	0.53	0.27	0.22	0.15	0.23	0.03	-0.14	0.20	0.05	0.03
ALPPL2	0.29	0.19	0.02	0.66	0.40	0.43	-0.02	-0.02	0.25	-0.04	-0.20	0.22
BPGM	0.23	0.20	0.27	0.11	0.05	0.66	0.46	0.39	0.62	0.65	0.43	0.44
BPNT1	-0.19	0.05	0.01	-0.81	-0.40	-0.14	0.14	-0.15	0.37	-0.30	-0.17	-0.25
CDC14A	-0.03	-0.02	-0.29	0.40	-0.26	0.17	-0.26	0.09	0.12	-0.27	-2.29	-0.43
CDC14B	0.09	0.12	0.02	0.43	-0.09	-0.03	-0.10	-0.21	0.08	0.17	0.24	-0.27
CDC25A	-0.12	-0.41	-0.80	-1.20	-0.95	-0.63	-0.88	-1.41	-0.41	-0.08	-0.73	-0.19
CDC25B	-0.17	-0.36	-0.79	-0.10	-0.69	-0.34	-0.41	-0.01	-0.20	-0.77	-3.29	-0.13
CDC25C	-0.13	0.31	-0.58	-0.20	0.05	-0.02	-0.34	-0.52	0.01	-0.50	-0.04	0.42
CTDP1	-2.60	-1.85	-1.82	-1.42	-1.99	-3.95	-1.07	-1.08	-3.45	-4.41	-5.25	-5.16
CTDSP1	-0.12	-0.34	-0.48	-0.63	-0.24	-0.10	-0.15	-0.05	0.24	-0.12	0.21	-0.82
CTDSP2	0.11	-0.24	-0.25	-0.18	-0.48	-0.06	-0.35	-0.18	-0.06	-0.12	-3.37	0.09
CTDSPL	-0.14	-0.06	-0.33	-0.58	-0.44	0.04	-0.07	-0.44	0.11	0.01	-0.49	-0.85
CTDSPL2	-4.86	-1.69	-2.51	-2.64	-1.40	-0.44	-0.60	-0.98	-0.27	-1.40	1.01	-0.69
DOLPP1	-0.96	-0.09	-1.17	0.02	-0.49	-0.91	-0.57	-0.04	-0.15	-2.27	-2.40	-3.36
DUPD1	0.33	0.22	-0.22	-0.06	0.13	0.24	0.00	-0.07	-0.23	0.52	-0.86	-0.51
DUSP1	0.06	-0.15	0.07	0.40	-0.11	0.26	0.12	0.02	0.20	-0.10	0.21	0.36
DUSP10	0.34	-0.05	0.47	-0.10	-0.13	0.10	-0.02	0.30	0.10	0.12	0.46	0.45
DUSP11	0.00	-0.02	0.20	0.35	0.22	0.09	0.15	0.08	0.17	0.31	-2.26	0.15
DUSP12	-1.81	-0.40	-0.91	-0.49	-0.56	-0.12	-0.08	-0.31	-0.69	-0.55	0.26	0.41
DUSP13	0.02	0.21	-0.22	0.21	0.06	-0.33	-0.56	-0.47	-0.06	0.23	-0.24	-0.27
DUSP14	0.30	-0.04	0.22	-0.16	-0.01	-0.21	0.12	-0.16	0.15	-0.09	-0.31	0.02
DUSP15	-0.10	0.20	0.28	-0.23	0.14	-0.19	0.18	0.07	-0.10	-0.01	-0.28	0.08

Appendix 2. Average log₂ FC in phosphatase domain-focused screen (2/7)

Disease RRID	AML CVCL_2119	AML CVCL_0095	AML CVCL_1609	AML CVCL_0001	AML CVCL_1844	AML CVCL_0007	AML CVCL_0589	CML CVCL_0004	SCLC CVCL_1453	PDAC CVCL_0152	RMS CVCL_0041	RMS CVCL_1649
Cell_line	MOLM-13	SEM	NOMO-1	HEL	OCI-AML-3	U-937	Kasumi-1	K-562	NCI-H1048	AsPC-1	Rh30	RD
DUSP16	-0.11	0.19	0.36	0.58	0.07	0.15	0.03	-0.01	-0.02	0.07	0.51	0.27
DUSP18	0.27	0.13	-0.23	-0.06	0.05	0.06	-0.25	0.11	-0.19	0.22	0.04	-0.41
DUSP19	-0.06	-0.16	0.36	-0.03	-0.04	-0.26	-0.21	-0.24	0.12	-0.02	0.28	-0.48
DUSP2	-0.43	-0.04	-0.41	0.34	-0.17	-0.04	-1.02	-0.02	0.19	-0.22	-0.24	-0.18
DUSP21	0.26	0.42	-0.06	-0.10	0.13	0.18	0.22	-0.20	-0.49	0.10	0.56	0.28
DUSP22	-0.79	-0.55	-0.27	0.36	-0.21	-0.17	-0.27	-0.25	-0.15	-0.03	-0.39	0.39
DUSP23	0.05	0.01	-0.14	0.30	0.18	-0.17	-0.30	0.25	-0.41	0.31	0.45	-0.23
DUSP26	-0.44	-0.53	0.04	-0.11	-0.02	0.07	-0.23	-0.13	0.03	0.11	0.52	-0.57
DUSP28	0.06	0.20	0.13	-0.12	0.17	0.19	0.26	0.03	0.22	0.17	-0.41	-1.64
DUSP3	0.06	0.30	-0.29	-0.24	0.06	0.08	-0.06	-0.34	0.10	-0.33	-0.47	0.01
DUSP4	-0.55	-0.20	-0.56	-0.10	-0.31	0.02	-0.15	0.05	-0.04	-1.27	0.35	-0.15
DUSP5	0.12	0.12	-0.03	0.14	-0.07	-0.17	-0.11	-0.06	0.26	0.23	0.52	-0.85
DUSP6	-0.10	0.11	0.15	0.47	0.00	0.24	-0.23	0.06	0.08	0.42	-0.10	-0.03
DUSP7	-0.15	0.15	-0.25	-0.15	-0.03	-0.04	-0.25	-0.15	0.07	0.27	0.36	-0.30
DUSP8	0.08	-0.42	-0.10	0.39	0.04	0.16	0.07	-0.17	-0.75	0.13	0.40	0.07
DUSP9	-0.10	0.11	0.22	0.09	-0.03	0.19	0.13	-0.06	-0.11	0.06	0.06	0.18
ENOPH1	-0.26	-0.20	-0.02	0.21	0.11	-0.14	-0.09	0.02	-0.09	-0.24	-0.08	-0.06
EPHX2	-0.21	-0.03	0.00	0.56	0.06	-0.04	0.23	0.23	0.11	0.34	-0.13	-0.20
EPM2A	-0.07	0.19	-0.09	0.19	-0.22	0.08	-0.22	-0.17	-0.74	0.25	-0.74	0.03
EYA1	-0.59	0.05	-0.18	0.63	-0.02	0.19	0.15	0.30	0.28	0.00	-0.32	0.06
EYA2	0.04	0.10	-0.04	0.15	0.09	0.20	0.29	0.02	0.35	-0.06	0.06	-0.16
EYA3	-0.18	-0.04	0.18	-0.21	0.18	-0.35	0.25	-0.47	0.16	-0.27	0.35	0.25
EYA4	0.44	0.01	-0.01	0.46	0.28	0.08	-0.06	-0.50	0.01	-0.06	0.21	0.66
FBP1	0.17	0.10	-0.05	-0.46	-0.03	0.08	-0.18	-0.23	-0.62	-0.09	-0.63	-0.53
FBP2	-0.24	-0.03	0.07	0.17	0.17	0.16	-0.33	-0.09	-1.46	0.17	-0.03	0.04
FIG4	-0.43	-0.14	0.05	0.38	-0.07	0.09	-0.09	-0.04	-0.13	-0.33	-1.10	-0.27
G6PC	0.19	0.12	-0.40	-0.07	-0.15	-0.24	-0.05	-0.60	0.07	-0.04	-0.86	-0.44
G6PC2	0.15	-0.08	-0.15	-0.29	-0.12	0.04	-0.05	-0.42	-0.13	0.12	-0.05	0.78
G6PC3	-0.12	-0.16	-0.49	-0.52	-0.31	-0.18	-0.05	-0.18	-0.29	0.09	-0.29	-0.05
ILKAP	0.36	-0.44	-0.30	0.09	0.13	-0.13	0.10	0.30	-0.55	-0.03	-0.04	-0.23
IMPA1	-1.06	-0.12	-0.28	-0.06	-0.05	0.00	0.42	0.26	0.03	0.06	-0.18	0.38

Appendix 2. Average log₂ FC in phosphatase domain-focused screen (3/7)

Disease	AML	CML	SCLC	PDAC	RMS	RMS						
RRID	CVCL_2119	CVCL_0095	CVCL_1609	CVCL_0001	CVCL_1844	CVCL_0007	CVCL_0589	CVCL_0004	CVCL_1453	CVCL_0152	CVCL_0041	CVCL_1649
Cell_line	MOLM-13	SEM	NOMO-1	HEL	OCI-AML-3	U-937	Kasumi-1	K-562	NCI-H1048	AsPC-1	Rh30	RD
IMPA2	0.17	0.07	0.13	0.66	0.33	0.64	0.17	0.14	0.54	0.65	0.65	0.31
INPP1	0.19	0.31	0.11	-0.17	0.10	-0.03	0.08	0.08	0.34	-0.19	-0.03	0.10
INPP4B	-0.09	0.34	-0.01	0.46	0.19	0.27	0.01	0.07	0.49	0.09	0.19	0.44
INPP5A	-0.03	-0.19	0.25	0.43	-0.34	0.11	0.04	0.04	-0.02	0.08	0.35	-0.45
INPP5B	0.03	-0.03	-0.25	-0.12	-0.16	-0.19	-0.05	-0.10	0.16	-0.12	-0.08	-0.23
INPP5D	0.55	0.06	0.07	0.86	1.24	0.45	0.34	0.17	-0.17	0.10	0.14	-0.24
INPP5E	-0.46	0.01	-0.68	0.27	0.15	0.15	-0.31	0.06	0.06	-0.09	0.55	0.18
INPP5F	-0.26	0.01	0.29	0.19	-0.12	0.05	0.08	0.26	-0.25	0.00	0.22	-0.29
INPP5J	0.42	-0.19	0.05	0.36	0.10	0.14	-0.04	0.09	0.44	0.04	-0.32	-0.04
INPP5K	-0.17	-0.18	-0.31	0.58	-0.09	0.02	0.12	-0.21	-0.18	0.11	-0.43	-0.56
INPPL1	-0.05	-0.28	0.07	0.41	-0.01	0.17	0.16	-1.32	0.55	0.17	-1.27	0.05
LHPP	0.53	0.40	-0.38	0.13	-0.07	-0.06	0.13	-0.40	-0.73	0.13	0.53	-0.02
LPIN1	-0.14	-0.04	-0.29	-0.15	-0.17	0.04	0.10	0.05	0.02	-0.20	-0.09	0.07
LPIN2	0.33	0.00	0.16	0.73	-0.08	-0.16	0.02	-0.13	0.63	0.23	-0.15	-0.01
LPIN3	0.16	0.18	-0.30	0.56	0.02	0.16	-0.26	0.13	-0.02	0.38	0.13	-0.28
MDP1	-0.15	0.11	0.15	-0.02	-0.21	-0.13	0.06	0.17	-0.11	-0.05	-0.11	0.14
MINPP1	0.16	-0.19	-0.48	0.61	0.13	0.23	0.09	-0.16	0.17	0.68	0.56	0.20
MTM1	0.17	0.09	-0.03	0.15	0.11	0.00	0.20	0.44	-0.30	-0.03	-0.16	0.60
MTMR1	0.41	0.03	0.05	0.47	0.48	0.13	0.09	0.45	0.17	0.38	-0.39	-0.18
MTMR10	0.07	0.00	-0.01	0.16	0.01	0.12	0.18	0.24	0.26	0.20	0.31	0.31
MTMR11	-0.02	-0.14	0.11	-0.08	-0.21	0.05	0.08	-0.03	-0.15	-0.06	0.08	0.09
MTMR12	0.14	0.00	-0.45	-0.05	-0.52	0.14	0.10	0.07	0.11	0.17	-0.16	0.00
MTMR14	-0.18	-0.29	-0.14	0.19	0.04	-0.12	-0.03	-0.39	0.29	0.06	0.52	0.21
MTMR2	0.20	0.24	0.04	0.46	0.17	0.06	0.45	0.01	0.14	-0.01	-0.60	-0.02
MTMR3	-0.06	0.07	-0.28	0.12	-0.10	0.09	-0.30	0.08	0.17	0.25	-0.35	-0.73
MTMR4	-0.09	0.00	0.11	-0.24	-0.25	-0.28	-0.37	-0.29	-0.64	0.12	0.32	-1.07
MTMR6	0.07	-0.27	0.15	0.14	0.23	0.03	0.12	0.22	-0.18	0.11	-0.50	-0.14
MTMR7	-0.42	0.13	0.03	0.10	-0.43	0.24	-0.01	0.25	0.10	0.25	0.68	0.20
MTMR8	0.08	0.07	-0.20	0.48	0.23	0.25	0.05	0.18	0.04	0.03	-1.10	0.03
MTMR9	-0.38	-0.18	0.00	0.08	-0.41	-0.10	0.03	-0.44	0.16	0.05	0.02	-0.04
NANP	0.12	0.31	0.46	0.60	0.06	0.34	0.35	-0.16	-0.08	-0.74	0.39	0.33

Appendix 2. Average log₂ FC in phosphatase domain-focused screen (4/7)

Disease	AML	CML	SCLC	PDAC	RMS	RMS						
RRID	CVCL_2119	CVCL_0095	CVCL_1609	CVCL_0001	CVCL_1844	CVCL_0007	CVCL_0589	CVCL_0004	CVCL_1453	CVCL_0152	CVCL_0041	CVCL_1649
Cell_line	MOLM-13	SEM	NOMO-1	HEL	OCI-AML-3	U-937	Kasumi-1	K-562	NCI-H1048	AsPC-1	Rh30	RD
NT5C	-0.56	0.03	0.29	0.44	-0.23	-0.06	-0.54	-0.46	-0.58	0.42	-0.43	-0.57
NT5C1A	-0.04	-0.05	-0.20	0.57	-0.06	0.21	-0.25	-0.30	-0.22	-0.29	-0.28	-0.09
NT5C1B	0.19	0.34	0.50	0.46	0.17	0.21	0.04	0.07	-0.07	-0.05	-0.33	0.52
NT5C2	0.34	0.01	-0.23	-0.08	0.06	-0.18	-0.12	-0.06	-0.17	-0.32	0.39	-0.38
NT5C3A	-0.10	0.06	0.24	0.37	-0.02	0.16	0.17	0.13	-0.05	0.28	0.29	-0.01
NT5E	-0.19	0.00	0.04	-0.65	0.04	-0.03	-0.29	0.50	0.08	0.17	-0.18	0.13
NT5M	0.01	-0.10	-0.07	0.74	-0.04	0.27	0.08	0.09	-0.46	0.10	-0.65	0.03
OCRL	-0.41	-0.34	-0.45	-0.01	-0.09	0.01	0.34	-0.27	-0.53	-0.50	0.20	0.46
PDP1	-0.64	-0.01	0.09	0.36	0.01	0.14	-0.02	0.17	0.15	-0.22	0.57	0.23
PDP2	0.02	0.16	-0.31	0.06	-0.16	0.12	-0.24	0.06	0.02	-0.32	-0.70	0.28
PDXP	-0.58	0.27	-0.67	0.07	-0.17	0.13	-0.28	-0.23	-0.01	-0.44	-0.31	-0.70
PFKFB1	0.44	-0.06	0.23	-0.06	0.15	0.02	0.13	0.15	0.40	-0.17	0.36	0.40
PFKFB2	0.17	0.07	-0.13	0.42	-0.13	-0.01	0.02	0.25	0.28	0.45	-0.02	0.14
PFKFB3	0.27	-0.40	-0.20	0.02	-0.05	-0.34	-0.30	0.26	0.14	-0.24	0.64	-0.13
PFKFB4	0.12	0.15	-0.84	0.12	-0.06	-0.35	-0.06	-0.26	0.05	-0.04	0.32	-1.89
PGAM2	-0.28	0.12	-0.01	0.24	-0.26	0.05	-0.59	-0.12	-0.19	0.11	-0.24	0.28
PGAM4	0.01	-0.41	0.08	0.14	-0.37	0.09	0.14	0.07	-0.02	-0.02	-0.33	-0.22
PGAM5	-0.26	-0.37	0.03	-0.90	-0.02	-0.14	-0.24	-0.65	-0.02	-0.52	0.23	-0.16
PHLPP1	0.25	-0.99	-0.13	0.00	-0.38	-0.01	-0.31	0.06	0.09	-0.06	0.28	0.15
PHLPP2	-0.26	-0.24	-0.28	-0.52	0.04	0.07	0.01	0.08	0.03	-0.02	-0.10	0.26
PHOSPHO1	-0.07	0.06	-0.86	0.52	-0.10	-0.05	-0.59	-0.02	0.23	0.50	-1.50	0.21
PHOSPHO2	0.56	0.21	-0.22	0.03	0.20	-0.04	0.05	-0.31	0.52	0.36	-0.28	0.62
PHPT1	0.07	0.39	-0.78	0.41	-0.12	0.09	-0.18	0.24	-0.15	-0.17	0.39	-0.44
PLPP1	-0.15	0.17	-0.13	0.26	0.20	-0.04	0.01	0.01	0.13	0.28	-0.44	0.16
PLPP2	-0.40	-0.18	-0.27	0.03	-0.17	0.05	0.24	-0.10	-0.74	0.01	0.19	0.09
PLPP3	0.13	0.13	0.27	0.28	-0.11	0.13	-0.19	-0.21	0.37	0.25	-0.51	-1.06
PLPP4	0.29	-0.01	-0.05	0.43	0.21	0.21	0.14	0.28	0.37	0.15	-0.32	0.40
PLPP5	-0.75	0.09	-0.33	-0.03	-0.18	0.15	0.10	-0.40	0.33	0.12	0.21	0.06
PLPP6	-0.26	-0.05	0.27	-3.62	0.18	0.22	0.08	0.17	-0.16	0.22	0.27	-0.38
PLPP7	0.13	0.10	-0.04	-0.38	0.31	0.00	-0.30	0.03	-0.03	-0.03	-0.12	-0.86
PMM1	-0.40	-0.05	-0.57	0.12	0.03	0.05	-0.11	-0.19	-0.73	0.08	0.26	-0.62

Appendix 2. Average log₂ FC in phosphatase domain-focused screen (5/7)

Disease	AML	CML	SCLC	PDAC	RMS	RMS						
RRID	CVCL_2119	CVCL_0095	CVCL_1609	CVCL_0001	CVCL_1844	CVCL_0007	CVCL_0589	CVCL_0004	CVCL_1453	CVCL_0152	CVCL_0041	CVCL_1649
Cell_line	MOLM-13	SEM	NOMO-1	HEL	OCI-AML-3	U-937	Kasumi-1	K-562	NCI-H1048	AsPC-1	Rh30	RD
PMM2	-1.17	-3.11	-1.56	-4.08	-2.24	-0.58	-1.89	-0.15	-1.39	-3.51	-0.30	-0.27
PNKP	-3.70	-2.06	-2.89	-1.64	-2.21	-1.12	-1.14	-1.43	-3.87	-3.81	-2.18	-3.12
PPEF1	0.32	0.00	0.21	-0.10	0.22	0.00	0.02	-0.32	-0.57	0.34	0.55	0.06
PPEF2	0.03	0.03	0.20	0.53	0.10	0.19	-0.17	0.41	0.00	0.21	-0.34	0.04
PPM1A	-0.34	-0.14	-0.50	-0.07	0.63	0.21	0.37	0.01	0.02	-0.52	0.55	-0.49
PPM1B	-0.32	0.08	0.05	-0.59	0.31	-0.14	-0.09	0.09	-0.35	0.06	-0.11	-0.06
PPM1D	-3.48	-0.18	-0.73	-0.26	-2.80	0.01	-0.27	-0.32	0.40	-0.04	-0.40	0.00
PPM1E	-0.10	0.21	0.07	0.07	-0.19	-0.15	-0.12	-0.22	-0.06	-0.23	0.12	0.34
PPM1F	-0.01	-0.13	0.31	0.22	0.09	-0.04	-0.03	-1.41	-0.24	-0.46	-0.26	-0.11
PPM1G	-1.33	0.23	-0.61	-1.27	-0.20	-0.39	-0.09	-0.50	0.03	-0.21	-0.83	1.38
PPM1H	0.09	0.10	-0.12	-0.26	-0.03	0.12	-0.08	-0.17	-0.75	0.09	-1.06	-0.93
PPM1J	-0.59	-0.07	-0.75	0.78	-0.23	0.01	-0.58	-0.65	-0.46	-0.11	-0.91	0.00
PPM1K	-0.15	-0.04	-0.04	0.27	-0.11	0.08	0.11	-0.02	0.29	0.34	0.03	0.50
PPM1L	-0.07	-0.01	-0.27	-0.03	-0.17	0.19	-0.18	-0.37	0.36	-0.01	0.01	-0.03
PPM1M	-0.04	0.22	-0.56	0.04	-0.17	-0.01	-0.32	0.06	-0.27	-0.02	-0.41	-0.41
PPM1N	-0.36	-0.32	0.14	0.22	0.08	-0.10	0.01	0.19	0.18	0.02	-0.35	0.15
PPP1CA	-0.35	-1.24	-0.42	-0.37	-1.51	-0.19	-0.02	-0.85	-0.18	-0.18	-0.54	-0.55
PPP1CB	-2.45	-2.41	-2.47	-3.31	-2.90	-2.52	-1.56	-1.47	-3.23	-3.81	-4.04	-4.73
PPP1CC	0.13	0.08	0.10	0.24	-2.88	0.30	-0.07	-0.52	0.00	0.24	-0.05	-0.51
PPP2CA	-7.19	-5.63	-3.39	-2.67	-5.99	-3.85	-2.68	-0.85	-4.62	-0.57	-6.28	-1.64
PPP2CB	-0.61	-0.02	-0.23	0.00	0.02	0.21	0.06	-0.10	0.20	0.16	0.42	0.23
PPP3CA	-0.30	0.42	0.39	0.13	-0.08	0.10	-0.15	-0.12	0.06	0.35	-0.70	0.05
PPP3CB	0.00	0.18	0.41	-0.06	0.04	0.31	-0.28	0.16	0.06	-0.15	-0.41	0.04
PPP3CC	-0.84	0.30	-0.26	0.79	-0.18	0.16	0.20	0.31	0.61	0.16	0.56	0.07
PPP4C	-5.71	-3.37	-3.08	-4.12	-3.17	-4.66	-2.44	-1.88	-5.68	-5.02	-3.24	-7.31
PPP5C	-0.35	-0.31	0.09	-0.08	0.16	-0.07	-0.31	-0.17	-0.03	0.13	0.33	0.01
PPP6C	-2.66	-1.32	-1.48	-2.19	-0.15	-1.40	-0.82	-0.34	-1.31	-1.82	-2.13	-1.92
PSPH	-0.55	-0.13	-0.23	-0.14	-0.80	-0.34	-0.04	-0.59	0.15	0.21	-0.14	0.19
PTEN	0.26	0.60	0.85	1.05	1.00	0.35	0.27	0.17	-0.48	0.66	0.10	0.93
PTP4A3	-0.82	-0.46	-0.58	0.28	-0.06	-0.14	-0.65	-0.27	0.14	-0.22	-0.75	-0.52
PTPDC1	0.02	0.18	0.16	0.58	0.21	0.28	0.12	0.15	-0.11	0.08	0.58	-0.42

Appendix 2. Average log₂ FC in phosphatase domain-focused screen (6/7)

Disease	AML	CML	SCLC	PDAC	RMS	RMS						
RRID	CVCL_2119	CVCL_0095	CVCL_1609	CVCL_0001	CVCL_1844	CVCL_0007	CVCL_0589	CVCL_0004	CVCL_1453	CVCL_0152	CVCL_0041	CVCL_1649
Cell_line	MOLM-13	SEM	NOMO-1	HEL	OCI-AML-3	U-937	Kasumi-1	K-562	NCI-H1048	AsPC-1	Rh30	RD
PTPMT1	-7.73	-3.45	-3.77	-4.61	-0.85	-4.43	-0.86	-3.99	-4.81	-2.71	-4.82	-5.14
PTPN1	1.20	0.72	0.45	0.69	0.28	0.24	0.20	-1.40	-0.07	-0.40	-0.34	-0.31
PTPN11	-3.44	-2.87	-0.31	-3.54	-0.08	-4.26	-2.82	-2.33	-3.67	-0.84	-2.29	-0.04
PTPN12	0.04	-0.35	-0.45	0.32	0.14	0.31	0.19	-0.11	-0.16	1.48	0.63	-0.25
PTPN13	0.24	0.13	0.05	0.07	0.04	0.29	0.23	0.29	0.44	0.27	0.04	0.09
PTPN14	-0.09	0.16	0.28	-0.04	0.01	0.25	0.16	-0.46	0.11	0.36	0.40	0.12
PTPN18	-0.09	-0.04	-0.19	0.41	0.09	-0.05	-0.31	-0.04	0.69	-0.43	-0.51	-0.51
PTPN2	0.48	0.14	0.06	0.18	0.27	0.09	0.13	-0.36	0.34	0.16	0.13	0.04
PTPN21	0.39	0.12	-0.07	0.46	0.20	0.15	-0.05	0.17	0.05	0.21	0.13	0.24
PTPN22	0.18	-0.15	-0.49	0.04	0.23	-0.12	0.16	-0.73	-0.37	0.15	-0.60	0.33
PTPN23	-1.55	-2.20	-1.88	-1.27	-1.43	-1.73	-1.50	-0.70	-2.38	-1.27	-2.39	-1.63
PTPN3	0.15	0.17	0.07	0.09	0.16	0.15	-0.09	0.09	0.08	-0.01	0.43	-0.78
PTPN4	0.00	0.54	0.15	0.23	0.07	-0.09	-0.18	-0.22	0.17	-0.02	-1.49	0.15
PTPN5	0.33	-0.35	-0.17	0.74	0.03	0.23	-0.08	0.18	0.11	0.23	-1.50	-0.29
PTPN6	0.09	0.21	-0.08	0.40	-0.09	0.03	-0.10	-0.28	0.04	0.02	-0.22	0.22
PTPN7	-0.70	0.00	-1.91	0.06	-4.51	-0.12	-0.13	-0.37	-0.18	-0.16	0.32	0.01
PTPN9	0.33	-0.05	-0.75	-0.08	-0.18	-0.28	-0.10	-0.14	0.23	0.09	0.68	-1.23
PTPRA	-0.58	0.11	-0.32	-0.29	0.01	-0.35	0.11	-0.24	-0.12	-0.14	-0.26	-0.14
PTPRB	0.02	0.16	0.11	0.14	0.12	0.11	0.02	0.11	0.22	0.33	-0.06	-0.47
PTPRC	0.06	0.09	0.01	0.08	-0.07	0.04	0.45	-0.26	-0.49	0.00	-0.13	-0.03
PTPRD	0.20	0.03	0.27	-1.41	0.03	0.35	0.21	0.09	0.32	0.52	-0.14	0.01
PTPRE	0.35	-0.09	-0.17	0.37	0.18	0.44	-0.11	-0.17	-0.07	0.06	-0.41	-0.12
PTPRF	0.03	-0.08	-0.34	-0.28	0.03	-0.05	-0.04	0.02	0.14	-0.19	0.06	-0.24
PTPRG	0.11	0.27	-0.03	0.15	-0.12	0.08	0.42	0.12	-0.08	0.17	0.07	0.15
PTPRH	-0.28	-0.06	-0.07	0.19	0.17	-0.06	0.02	-0.23	-0.13	-0.11	-0.16	0.36
PTPRJ	0.07	0.66	-0.52	-0.02	-0.24	-0.21	-0.24	-0.35	0.08	0.29	-0.20	-0.13
PTPRK	-0.07	0.23	0.29	0.55	0.02	0.18	0.03	0.23	0.12	0.46	0.19	-0.15
PTPRM	0.07	0.13	0.15	0.34	0.12	0.18	0.26	0.11	0.25	0.22	0.28	0.07
PTPRN	-0.21	0.08	-0.49	-0.15	0.04	-0.06	0.03	0.17	0.10	-0.29	0.21	-0.19
PTPRN2	-0.15	0.19	0.39	0.04	-0.09	0.17	-0.08	-0.02	-0.06	0.04	0.35	0.38
PTPRO	0.06	-0.19	0.40	0.19	0.28	0.17	0.38	-0.32	-0.20	0.14	0.54	1.11

Appendix 2. Average log₂ FC in phosphatase domain-focused screen (7/7)

Disease RRID	AML CVCL_2119	AML CVCL_0095	AML CVCL_1609	AML CVCL_0001	AML CVCL_1844	AML CVCL_0007	AML CVCL_0589	CML CVCL_0004	SCLC CVCL_1453	PDAC CVCL_0152	RMS CVCL_0041	RMS CVCL_1649
Cell_line	MOLM-13	SEM	NOMO-1	HEL	OCI-AML-3	U-937	Kasumi-1	K-562	NCI-H1048	AsPC-1	Rh30	RD
PTPRQ	0.08	0.18	0.07	0.30	0.13	0.28	0.02	-0.28	-0.12	0.02	0.48	0.11
PTPRR	0.15	0.47	0.38	0.58	0.07	0.19	-0.04	-0.14	-0.12	0.02	-1.22	-0.27
PTPRS	0.07	-0.35	0.09	0.21	0.07	0.00	-0.09	-0.11	0.08	0.16	0.04	0.23
PTPRT	0.05	-0.05	0.10	0.04	0.05	0.13	-0.05	0.13	0.18	-0.03	-0.51	-0.24
PTPRU	-0.18	-0.09	0.05	0.31	0.03	0.05	0.02	-0.11	-0.04	0.12	0.18	-0.18
PTPRZ1	0.09	-0.04	0.19	0.53	-0.08	0.10	0.38	0.32	0.16	-0.35	0.57	-0.44
PUDP	0.22	-0.22	0.16	0.28	0.36	0.25	0.41	0.07	0.31	-0.01	0.91	0.27
PXYLP1	0.05	-0.03	-0.29	0.93	-0.41	-0.03	0.22	-0.50	0.00	0.21	0.07	0.07
RNGTT	-3.04	-3.34	-2.42	-2.92	-2.89	-3.35	-2.08	-1.35	-3.72	-4.41	-3.36	-4.28
SACM1L	-3.59	-2.22	-3.58	-2.74	-4.56	-3.10	-2.20	-1.87	-6.26	-5.02	-3.99	-4.86
SBF1	0.08	0.02	-0.21	0.37	0.09	0.13	-0.02	-0.03	-0.01	-0.09	-0.27	-0.54
SBF2	-0.05	-0.07	0.12	0.27	0.00	0.08	0.12	-0.15	0.40	0.11	0.04	-0.24
SGPP1	0.36	0.12	0.33	0.72	-0.29	0.22	-0.22	0.03	-0.16	0.27	0.28	0.08
SGPP2	0.16	-0.34	-0.11	0.05	0.01	0.04	0.10	0.11	-0.09	-0.10	0.01	-0.01
SSH1	0.17	0.07	0.09	0.08	0.05	-0.02	-0.26	-0.26	-0.17	0.27	-0.38	-0.09
SSH2	0.04	-0.24	-0.01	-0.24	0.12	-0.17	0.31	-0.19	-3.28	-0.42	0.61	-0.22
SSH3	-0.12	-0.04	0.08	-0.14	-0.02	0.04	0.06	-0.21	-0.27	0.06	-0.29	-0.23
SSU72	-5.20	-4.64	-2.41	-3.98	-4.17	-6.48	-2.34	-1.58	-6.81	-6.04	-6.53	-6.43
STYX	0.02	-0.21	-0.54	-0.23	-0.11	0.07	-0.24	-0.19	0.20	-0.08	0.19	-0.10
STYXL1	-0.40	-0.12	-1.20	0.09	-0.03	0.22	-0.09	-0.11	0.16	0.14	0.01	-0.42
SYNJ1	-0.36	-0.12	-0.18	-0.06	0.12	0.18	-0.23	0.02	-0.17	0.10	0.28	0.26
SYNJ2	-0.08	0.04	-0.23	-0.05	-0.24	0.03	0.08	-0.20	0.12	0.45	0.44	-0.12
TIGAR	0.49	-0.08	0.37	0.45	-0.08	0.19	-0.05	0.34	-0.08	0.36	0.11	0.23
TIMM50	-7.29	-3.78	-3.08	-3.25	-3.16	-4.43	-2.24	-2.55	-8.61	-6.51	-8.74	-7.47
TMEM55A	-0.95	0.23	-0.53	0.12	0.09	0.32	-0.04	0.17	0.25	-0.22	-0.19	0.42
TMEM55B	-0.31	-0.33	-0.31	0.00	-0.01	-0.29	-0.44	0.06	-0.14	0.15	0.09	0.09
TNS1	0.17	0.23	0.24	-0.05	0.01	0.16	-0.07	0.10	0.28	-0.02	0.20	-0.40
TNS2	0.17	0.01	-0.29	-0.29	0.03	-0.02	-0.26	-0.35	-0.07	-0.31	-0.09	-0.05
TPTE	-0.35	-0.09	0.16	0.17	-0.23	0.15	0.54	-0.02	0.68	0.36	0.25	0.37
UBASH3B	0.03	-0.04	-0.35	-0.15	0.17	0.27	0.01	0.45	0.33	0.20	-0.04	0.18
UBLCP1	0.06	0.19	0.44	0.43	0.42	0.18	0.27	0.13	0.32	0.12	0.39	0.45

**Appendix 3. Average log₂FC in SCP4 CRISPR exon
scanning screen in MOLM-13**

Cut Site	Cut Codon	sgRNA Sequence	log ₂ FC
53	17	CTCTGGCAGTGCGTTGTGTT	0.48
66	22	AACGCACTGCCAGAGCAAAG	-2.02
70	23	TGAATATTTCTCTTTGCTC	-0.88
79	26	AGCAAAGAGGAAATATTCAG	0.26
101	33	GTTGATGATAGCCTGCCTTC	-0.47
104	34	GATGATAGCCTGCCTTCAGG	-1.53
107	35	GTTTTTCTCCTCCTGAAGGC	0.83
111	37	GATGGTTTTTCTCCTCCTGA	0.44
129	43	AAGCCGGTTTCATTCTTCGA	0.34
131	43	AAACCATCGAAGAATGAAAC	0.51
145	48	TTTAATTGAAGACAACAAGC	0.62
164	54	TCAATTAATAAATTTATTAA	2.75
204	68	AGAGAGAAAATCCTTCAAAA	-0.77
210	70	TCAATTCTACTCCGTTTTGA	2.21
260	86	ATCACGTCAACACCAAGAGC	-0.42
267	89	TTAGGTTTTTCTCCTGCTCT	0.77
285	95	ACTCGAGATATCTGTTTGTT	0.74
297	99	AACAGATATCTCGAGTAAGA	-0.23
314	104	AGACGGAAAAGTCAAGTAAA	-0.01
362	120	AATCAACATGTAAAACAAAA	2.28
386	128	TTAGAAGATAATCCTTCCTC	0.34
393	131	CTTGAGGACTGCCAGAGGA	0.56
397	132	AGTCCTTGGAGGACTGCCAG	0.74
402	134	CTTCCTCTGGCAGTCCTCCA	-1.47
408	136	CCCAACAAAGTAGTCCTTGG	0.84
413	137	CCTCCAAGGACTACTTTGTT	0.52
414	138	GTCCCAACAAAGTAGTCCT	1.34
415	138	TCCTCCAAGGACTACTTTGT	1.80
417	139	CTCCAAGGACTACTTTGTTG	1.33

APPENDICES

Cut Site	Cut Codon	sgRNA Sequence	log ₂ FC
436	145	GTTGAAGACAGGTGAAAATA	1.18
444	148	GGTGAAAAAAGTTGAAGAC	0.30
488	162	ACGTCAGGATCAGATTCTCC	0.90
493	164	AGGATCAGATTCTCCAGGAC	-0.23
499	166	AGATTCTCCAGGACAGGCTG	0.15
501	167	TCAGCTTCCACAGCCTGTCC	-0.36
535	178	AATAGTAAAACAACCTTGATA	0.52
541	180	ACAACCTTGATATGGAACAGG	1.05
541	180	AAAACAACCTTGATATGGAAC	1.07
571	190	ATTAGTTGATGTAGTAGTAC	1.26
578	192	AGTACTACTACATCAACTAA	0.56
629	209	GTGAGACCATCACTAAACAA	0.94
633	211	TCTAAACCATTGTTTAGTGA	0.76
699	233	TAACCACTATCTGGAGTTAC	0.08
708	236	GCTGATGAATAACCACTATC	-0.13
721	240	TGGTTATTCATCAGCCCACG	-0.11
727	242	TTATTCATCAGCCCACGCGG	-2.04
731	243	CTTCATAGGTGGCCTCCGCG	-1.32
733	244	TTCATAGGTGGCCTCCGCGT	-0.68
745	248	TACTTCCCAGTCTTCTTCAT	-1.88
745	248	TTCCCAGTCTTCTTCATAGG	-0.67
745	248	GGCCACCTATGAAGAAGACT	0.29
747	249	AGGCCACCTATGAAGAAGAC	0.19
797	265	GTTGTTCTTCTGTCAGTGGC	-0.06
801	267	AGTTGTTCTTCTGTCAGTGG	-3.02
804	268	TTAGTTGTTCTTCTGTCAG	2.16
810	270	TGACAGAAGAACAACATAAT	-3.88
834	278	CTTGTTTTCAACGGAAGAGC	-2.03
840	280	GGTGTGCTTCTTGTTTTCAA	-1.27
850	283	GTTGAAAACAAGAAGCACAC	-3.61
945	315	ATGACATCTTGAAAAGGAC	-0.78

APPENDICES

Cut Site	Cut Codon	sgRNA Sequence	log ₂ FC
950	316	GATAAATGACATCTTGGAAA	-3.02
989	329	ATTTCTTCTTTGCCAGTAGA	0.78
991	330	GAGATTAAGACCATTTTTCA	-1.78
990	330	TGAGATTAAGACCATTTTTTC	1.49
993	331	TCCAGGAATCCCTGAAAAA	-3.04
997	332	ACCATTTTTTCAGGGAATTCC	-3.36
1004	334	GCAAAGAAGAAATATACACA	1.50
1005	335	CAAAGAAGAAATATACACAA	2.46
1051	350	TCTTTTTACTGCTTCTAAGA	0.78
1099	366	ACTAGACCCTAAAAAGCAAC	-1.42
1124	374	CACAATGTTACCGAAAAGC	-2.71
1130	376	GTACACAAACACAATGTTCA	-2.81
1142	380	GAACATTGTGTTTGTGTACA	0.94
1156	385	TGTACAAGGAAACTATATAA	-5.40
1260	420	ATGGAATCCCTATAGAAAGT	-3.91
1260	420	TCCATAAACCAACTTTCTAT	-3.79
1262	420	CCATAAACCAACTTTCTATA	-2.41
1267	422	CCCTATAGAAAGTTGGTTTA	-1.06
1301	433	CCAGGAATGGAATCAATTTT	1.20
1309	436	CCTAAAATTGATTCCATTCC	-2.85
1356	452	CGAAATCTGTCTCTGATGTG	-1.06

Appendix 4. Genes downregulated upon SCP4 knockout

sgSCP4_dn_MOLM13_spolyans				
ACOT7	CHRNA6	HSD11B1	NPM2	SERF1A
ADAM11	COCH	HSD11B2	NR1I2	SFXN2
ADAMTS13	COPZ2	HYAL3	NRARP	SH2D2A
ADSSL1	CPNE5	IGFBP2	NRTN	SHISA2
ALDH5A1	CPNE7	IGLL1	OBSL1	SKA3
ALS2CL	CSPG5	IL2RB	OIP5	SKIDA1
ANGPT1	CST7	IQSEC3	ONECUT2	SLC16A13
ANK2	CTDSPL2	IRX3	OPRL1	SLC22A31
ANKRD34A	CTSG	IRX5	ORM2	SLC25A10
ANO7	CUX2	ITGA11	OSBP2	SLC29A1
ANXA8	DDN	JCAD	OSM	SLC29A4
ARHGAP23	DLEC1	KAZALD1	PAC SIN3	SLC2A5
ARHGEF39	DLG5	KCNQ4	PALM	SLC46A1
ASIC1	DMPK	KDELC2	PAQR4	SLCO4A1
ATP1A3	DMWD	KHK	PIF1	SMIM10
ATP6V1C2	DOCK6	KIRREL2	PIMREG	SMKR1
BSN	DOK4	KLHL35	PKMYT1	SPRY4
C19orf57	DRP2	KRT86	PLS1	SYNPO2
C21orf58	DZIP1L	LBHD1	PLXND1	TEDC2
C2orf48	EMID1	LDHD	PNMA6A	TERT
C3orf18	EMILIN1	LGALS12	PPM1E	TLN2
CACNB3	EPDR1	LOC112268013	PRRT4	TMEM229B
CAMSAP2	ETV4	LRP11	PTK2	TMEM45A
CASKIN2	FAM57B	LZTS1	PTPRN2	TNXB
CBSL	FAM69B	MAP2K6	PTPRS	TP53I11
CBX2	FAM83H	MAPK11	PYCR3	TRIM7
CCDC136	FAM86B1	MESP1	RAB26	TRNP1
CCDC142	FCGBP	MINOS1-NBL1	RAD54L	TSPAN32
CCDC182	FHL1	MLC1	RASIP1	TTC30A
CCDC85C	FIBCD1	MORN1	RCOR2	TUBB4A
CCND1	FLRT1	MS4A3	RHEX	TUSC1
CDC14B	GAL	MT1G	RHOV	TYRO3
CDC45	GATA2	MT1H	RIBC2	ULBP2
CDCA7	GATM	MTSS1L	RIMBP3C	VAT1L
CDH24	GINS1	MYH6	ROBO3	ZBTB16
CDK18	GOLGA7B	MYL5	ROS1	ZBTB47
CDT1	GPR153	NFIX	RPL22L1	ZDHHC11
CELSR2	GPT	NOTUM	RTN4R	ZNF367
CENPM	HES7	NOV	SEMA3F	ZNF469
CHRNA5	HIP1R	NOXA1	SEPT5	ZP3

Appendix 5. Genes upregulated upon SCP4 knockout

sgSCP4_up_MOLM13_spolyans				
A2M	CST3	GSN	MILR1	SAMD8
ABHD4	CTSZ	HBEGF	MPEG1	SAMHD1
ABTB1	CXCL8	HECA	MS4A6A	SCPEP1
ACPP	CYBB	HIST1H1C	MTMR11	SDCBP
ADA2	CYP27A1	HMOX1	MTURN	SEMA6B
ADAP2	CYSLTR1	HTR7	MVP	SHTN1
ADCY6	CYTH4	IFI30	MYBPH	SIGLEC7
AFDN	CYTIP	IGSF6	MYO1E	SIPA1L2
ALOX5	DAPP1	IL10RA	MYOF	SKAP2
AOAH	DPYD	IL13RA1	NAIP	SLAMF7
APOC1	DSC2	IL1A	NATD1	SLC25A20
ARHGEF10L	EEPDI1	IL1B	NBPF19	SLC37A2
ARHGEF3	EPB41L3	IPCEF1	NCF1	SLC43A2
ARPC4-TTLL3	EVI2A	IQCIN	NEFM	SLCO3A1
ARRDC3	EVI2B	ISG20	NFXL1	SLFN5
ARTN	FAM20C	ITGB7	NLRC4	SNX18
BCL3	FAM210B	JAK2	NTNG1	SORT1
BMF	FAM214B	JAML	OR52K2	SQSTM1
BTG2	FAM49A	KCNE3	PDK4	SRC
C1orf162	FCAR	KCTD12	PELI2	STX11
C1orf54	FCGR2B	KCTD7	PHLDA1	SULF2
C6orf226	FCGR3A	KIAA0513	PIGR	TCEAL4
CA2	FCGRT	KIAA1551	PIK3IP1	TFEC
CALCOCO1	FCMR	KLHDC8B	PITPNM1	TGFBR2
CAMK2D	FGL2	LACTB	PKD2L1	THEMIS2
CCDC144NL	FIG4	LIPA	PLD3	TLR1
CCM2	FILIP1L	LOC101059949	PLEKHO1	TLR2
CD101	FRY	LOC102725035	PLIN2	TLR8
CD109	FTH1	LOC112268354	PLK2	TMEM176B
CD180	FUCA1	LOC441155	PPP1R15A	TNFAIP2
CD36	GABARAP	LOC643802	PRDM1	TP53INP1
CD82	GABARAPL1	LPAR6	PRKCA	TSHZ3
CD84	GAPT	LPXN	PSAP	TYROBP
CDKN1A	GAS7	LRRK2	PTPRO	VSIG4
CHST15	GDF15	LSP1	RAB43	VSIR
CLEC10A	GLIPR1	LYZ	RBM47	YPEL3
CNTN4	GPR137B	MAP3K3	RCBTB2	YPEL5
COLQ	GPR141	MCOLN1	RHOC	ZFYVE16
CPEB4	GPR183	METTL7A	RTL5	ZNF257
CREBRF	GPR34	MFS14C	S100A6	ZNF366

Appendix 6. sgRNA sequences

sgRNA ID	Sequence
SCP4_e9	ACATCTGAGACATTCGTTCC
SCP4_e7	TCTAAAACACTAGGGAGAATTC
SCP4_e12	TCTACAAGCTTCTCCAGGAA
PDIK1L_e2	CTAATACGGGAGGTAGGCCG
STK35_e2	CGCGCCCGATCTCCGCCAAC
hROSA	AAGATGGGCGGGAGTCTTCG
PCNA_e2	GGACTCGTCCCACGTCTCTT
MYC_e3	GCTCCTCTGCTTGGACGGAC
Rosa26_(mouse)	GAAGATGGGCGGGAGTCTTC
Rpa3_(mouse)	ACGGGCCGGTCGATATACTG

**Appendix 7. Primers for the generation of CRISPR-resistant constructs
and introduction of the point mutations**

Name	Sequence	Comment
SCP4_e7_fwd	CAAGAAGCACACCGGAGTTTTTCGTTGGTCCTTGAC	CRISPR_resistant
SCP4_e7_rev	GTGTTTCATCCAAGTCAAGGACCAACGAAAACCTC	CRISPR_resistant
PDIK1L_e2_fwd	GCCAAAGTACGATCTTATTAGAGAAGTTGGAAGAGGTAGTTACGGTGTG	CRISPR_resistant
PDIK1L_e2_rev	CAACACCGTAACCTCTTCCAACCTCTCTAATAAGATCGTACTTTGGC	CRISPR_resistant
STK35_e2_fwd	CCGCGTTACAGTTTGTGGCGGAGATCG	CRISPR_resistant
STK35_e2_rev	CGATCTCCGCCAACAACTGTAACGCGG	CRISPR_resistant
SCP4236_466_dir	GATAGTGGTTATTTCATCAGCCACGCG	Truncation
SCP4_D293A_fwd	GGAGTTTTCGTTGGTCCTTGCCCTGGATGAAACACTAGTGCATTG	Catalytic_mutant
SCP4_D293A_rev	CAATGCACTAGTGTTCATCCAAGGCAAGGACCAACGAAAACCTCC	Catalytic_mutant
SCP4_D295A_fwd	GGAGTTTTCGTTGGTCCTTGACTGGCTGAAACACTAGTGCATTG	Catalytic_mutant
SCP4_D295A_rev	CAATGCACTAGTGTTCAGCCAAGTCAAGGACCAACGAAAACCTCC	Catalytic_mutant
SCP4_D293A_D295A_fwd	GGAGTTTTCGTTGGTCCTTGCCCTGGCTGAAACACTAGTGCATTG	Catalytic_mutant
SCP4_D293A_D295A_rev	CAATGCACTAGTGTTCAGCCAAGGCAAGGACCAACGAAAACCTCC	Catalytic_mutant
PDIK1L_K37H_fwd	CGGGTGGCAGTGCACAAAATTCGATGTCAC	Catalytic_mutant
PDIK1L_K37H_rev	GTGACATCGAATTTTGTGCACTGCCACCCG	Catalytic_mutant
STK35_K231H_fwd	GGTGGCGGTCCACAAGATCCGCTGC	Catalytic_mutant
STK35_K231H_rev	GCAGCGGATCTTGTGGACCGCCACC	Catalytic_mutant
PDIK1L_S194D_fwd	GGTGGCTGATTTTGGTCTAGATAAAGTTTGTTCAGCCTCTG	Phosphomimetic
PDIK1L_S194D_rev	CAGAGGCTGAACAACTTTATCTAGACCAAAATCAGCCACC	Phosphomimetic
STK35_S385D_fwd	GCCGACTTTGGACTAGACAAGTCTGTGCTG	Phosphomimetic
STK35_S385D_rev	CAGCACAGACCTTGTCTAGTCCAAAGTCGGC	Phosphomimetic
PDIK1L_S216D_fwd	GTGTTTCCTTGACACAGCATGTGGAACAGATTTTAC	Phosphomimetic
PDIK1L_S216D_rev	GTAAAAATCTGTTCCACATGCTGTGTCAAGGAAACAC	Phosphomimetic
PDIK1L_T217E_fwd	GTGTTTCCTTTCCGAGCATGTGGAACAGATTTTAC	Phosphomimetic
PDIK1L_T217E_rev	GTAAAAATCTGTTCCACATGCTTCCGAAAGGAAACAC	Phosphomimetic
PDIK1L_T221E_fwd	GTGTTTCCTTTCCACAGCATGTGGAGAAGATTTTAC	Phosphomimetic
PDIK1L_T221E_rev	GTAAAAATCTTCCACATGCTGTGGAAGGAAACAC	Phosphomimetic
STK35_S413D_fwd	GTAAGTGGCTGGACTCAGCCTGC	Phosphomimetic
STK35_S413D_rev	GCAGGCTGAGTCCAGCCAGTAC	Phosphomimetic
STK35_S414E_fwd	GTAAGTGGCTGTCCGAAAGCCTGC	Phosphomimetic
STK35_S414E_rev	GCAGGCTTCCGACAGCCAGTAC	Phosphomimetic
STK35_S418E_fwd	GCCTGCGGTGAGGACTTCTACATG	Phosphomimetic
STK35_S418E_rev	CATGTAGAAGTCTCACCAGCCAGGC	Phosphomimetic
PDIK1L_S194A_fwd	GGTGGCTGATTTTGGTCTAGCTAAAGTTTGTTCAGCCTCTG	Phosphomutant
PDIK1L_S194A_rev	CAGAGGCTGAACAACTTTAGCTAGACCAAAATCAGCCACC	Phosphomutant
STK35_S385A_fwd	GCCGACTTTGGACTAGCCAAGTCTGTGCTG	Phosphomutant
STK35_S385A_rev	CAGCACAGACCTTGGCTAGTCCAAAGTCGGC	Phosphomutant
PDIK1L_S216A_fwd	GTGTTTCCTTGCCACAGCATGTGGAACAGATTTTAC	Phosphomutant
PDIK1L_S216A_rev	GTAAAAATCTGTTCCACATGCTGTGGCAAGGAAACAC	Phosphomutant
PDIK1L_T217A_fwd	GTGTTTCCTTTCCGAGCATGTGGAACAGATTTTAC	Phosphomutant
PDIK1L_T217A_rev	GTAAAAATCTGTTCCACATGCTGCGGAAAGGAAACAC	Phosphomutant
PDIK1L_T221A_fwd	GTGTTTCCTTTCCACAGCATGTGGAGCAGATTTTAC	Phosphomutant
PDIK1L_T221A_rev	GTAAAAATCTGCTCCACATGCTGTGGAAAGGAAACAC	Phosphomutant
STK35_S413A_fwd	GTAAGTGGCTGGCCTCAGCCTGC	Phosphomutant
STK35_S413A_rev	GCAGGCTGAGGCCAGCCAGTAC	Phosphomutant

APPENDICES

STK35_S414A_fwd	GTACTGGCTGTCCGCAGCCTGC	Phosphomutant
STK35_S414A_rev	GCAGGCTGCGGACAGCCAGTAC	Phosphomutant
STK35_S418A_fwd	GCCTGCGGTGCGGACTTCTACATG	Phosphomutant
STK35_S418A_rev	CATGTAGAAGTCCGCACCGCAGGC	Phosphomutant

**Appendix 8. sgRNAs; media and cytokines; and antibodies used in
flow cytometry panels in CD34⁺ experiments**

Antibodies used in flow cytometry panels				
Lineage	Antibody	Clone	Manufacturer	Catalog #
Erythroid	BV421 Anti-Human CD49d	9F10	Biologend	304322
	FITC Mouse Anti-Human CD235a	GA-R2 (HIR2) (RUO)	BD Pharmingen™	559943
	APC Anti-Human Band3	New York Blood Center	Gift from X. An	
Myeloid	PECy7 CD11b	ICRF44	Biologend	301322
	BV421 CD33	clone HIM3-4	BD Bioscience	744350
Megakaryocyte	PECy7 CD41a	clone HIP8	BD Bioscience	561424
	BV421 Anti-Human CD42b	Clone HIP1	BD Bioscience	740075

Media and cytokines				
Medium	Component	Manufacturer	Catalog #	Final Concentration
CD34 ⁺ maintenance	StemSpan™ SFEM (base)	Stemcell Technologies	9650	
	Human Stem Cell Factor (hSCF)	R&D systems	rhSCF, 255-SC/CF	100 ng/mL
	Thrombopoietin (TPO)	R&D systems	rhTPO # 288-TP/CF	100 ng/mL
	FLT3-Ligand (FLT3-L)	R&D systems	rhFlt-3 # 3088-FK/CF	100 ng/mL
	Penicillin-Streptomycin	Thermo Fisher Scientific	15070063	Penicillin: 50 U/mL Streptomycin: 50 µg/mL
Myeloid differentiation	SFEM II	STEMCELL Technologies	9655	1X
	StemSpan™ Myeloid Expansion Supplement (100X)	STEMCELL Technologies	2693	1X
Megakaryocyte differentiation	SFEM II	STEMCELL Technologies	9655	1X
	StemSpan™ Megakaryocyte Expansion Supplement (100X)	STEMCELL Technologies	2696	1X
Erythroid differentiation (Common to all phases)	IMDM (base)	Thermo Fisher Scientific	12440061	
	Human Male AB Plasma	SeraCare	1810-0001	2%
	Human AB Serum	Atlanta Biologicals	S40110	3%
	Heparin	Sagent Pharmaceuticals	NDC # 25021-401-02	3 UI/mL
	Insulin	Lilly	Humulin® R U-100, NDC # 0002-8215-17	10 µg/mL
	EPO	Amgen	EPOGEN®, NDC # 55513-144-01	3 UI/mL

APPENDICES

Media and cytokines				
Medium	Component	Manufacturer	Catalog #	Final Concentration
	Penicillin-Streptomycin	Thermo Fisher Scientific	15070063	Penicillin 50 U/ mL Streptomycin 50 µg/mL
Erythroid differentiation (phase I): Add the following to common components	Human Holo-Transferrin	Millipore Sigma	T0665	200 µg/mL
	Human Stem Cell Factor (hSCF)	R&D systems	rhSCF, 255-SC/CF	10 ng/mL
	Human IL-3	R&D systems	rhIL-3, 203-IL/CF	1 ng/mL
Erythroid Differentiation (phase II): Add the following to common components:	Human Holo-Transferrin	Millipore Sigma	T0665	200 µg/mL
	Human Stem Cell Factor (hSCF)	R&D systems	rhSCF, 255-SC/CF	10 ng/mL
Erythroid differentiation (phase III): Add the following to common components	Human Holo-Transferrin	Millipore Sigma	T0665	1 mg/mL

The sequences of sgRNAs and primers		
Assay	Name	Sequence (5' to 3')
Single guide RNAs	e' refers to the exon # that is targeted by each sgRNA. Underlined sequence corresponds to target.	
	sgSCP4_e7	*U*C*U*AAAACUAGGGAGAAU <u>UCGUUUUAGAGCUAGAAAUA</u> GCAAGUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGC*U* U*U*U
	sgSCP4_e12	*U*C*U*ACAAGCUUCUCCAGG <u>AAGUUUUAGAGCUAGAAAUA</u> GCAAGUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGC*U* U*U*U
	sgMyc-e3	*G*C*U*CCUCUGCUUGGACGG <u>ACGUUUUAGAGCUAGAAAUA</u> GCAAGUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGC*U* U*U*U
	* 2'-O-methyl modification and phosphorothioate internucleotide linkage. Chemically modified versions of sgRNA-1 contained 2'-O-methyl modifications and phosphorothioate linkages in the first and last 3 nucleotides. gRNAs were synthesized by Trilink.	

APPENDICES

The sequences of sgRNAs and primers		
Assay	Name	Sequence (5' to 3')
		Underlined sequence corresponds to sgRNA + PAM
	SCP4e7_fwd	AGAAGCACAC <u>CCGGAATTCTC</u>
	SCP4e7_rev	TCACCCCTTTATTCTCTACCTC
	SCP4e12_fwd	<u>TCCATTCTGGAGAAGCTGTAG</u>
Indels by RT-qPCR	SCP4e12_rev	TCAGTACACACAAAGCAAAGACAT
	MYCe3_1_fwd	AT <u>CCTGTC</u> CCGICCAAGCAGA
	MYCe3_1_rev	ACTGTTCTCGTCGTTTCCGC
	MYCe3_2_fwd	TACAT <u>CCTGTC</u> CCGICCAAGC
	MYCe3_2_rev	CGCACAAGAGTTCCTAGCT
	chr10_fwd (negative control)	TAAATTTGGTTCAGGTGGTTGA
	chr10_rev (negative control)	ACCAAACTCTTCCCTCAGC
Indels by Surveyor	SCP4e7_Surv_fwd	AGGTATTTGAAACTTTTAATGGGGC
	SCP4e7_Surv_rev	AAGAAAAGGGTTATGGTCTTGCATA
	SCP4e12_Surv_fwd	TGAATGGGAGATATGGTCACTGTAT
	SCP4e12_Surv_rev	CAGTACACACAAAGCAAAGACATTT