



Published in final edited form as:

Cell Rep. 2022 January 11; 38(2): 110233. doi:10.1016/j.celrep.2021.110233.

## SCP4-STK35/PDIK1L complex is a dual phospho-catalytic signaling dependency in acute myeloid leukemia

Sofya A. Polyanskaya<sup>1</sup>, Rosamaria Y. Moreno<sup>2</sup>, Bin Lu<sup>1</sup>, Ruopeng Feng<sup>3</sup>, Yu Yao<sup>3</sup>, Seema Irani<sup>2</sup>, Olaf Klingbeil<sup>1</sup>, Zhaolin Yang<sup>1</sup>, Yiliang Wei<sup>1</sup>, Osama E. Demerdash<sup>1</sup>, Lukas A. Benjamin<sup>1</sup>, Mitchell J. Weiss<sup>3</sup>, Yan Jessie Zhang<sup>2</sup>, Christopher R. Vakoc<sup>1,4,\*</sup>

<sup>1</sup>Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

<sup>2</sup>Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA

<sup>3</sup>Department of Hematology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA

<sup>4</sup>Lead contact

### SUMMARY

Acute myeloid leukemia (AML) cells rely on phospho-signaling pathways to gain unlimited proliferation potential. Here, we use domain-focused CRISPR screening and identify the nuclear phosphatase SCP4 as a dependency in AML, yet this enzyme is dispensable in normal hematopoietic progenitor cells. Using CRISPR exon scanning and gene complementation assays, we show that the catalytic function of SCP4 is essential in AML. Through mass spectrometry analysis of affinity-purified complexes, we identify the kinase paralogs STK35 and PDIK1L as binding partners and substrates of the SCP4 phosphatase domain. We show that STK35 and PDIK1L function catalytically and redundantly in the same pathway as SCP4 to maintain AML proliferation and to support amino acid biosynthesis and transport. We provide evidence that SCP4 regulates STK35/PDIK1L through two distinct mechanisms: catalytic removal of inhibitory phosphorylation and by promoting kinase stability. Our findings reveal a phosphatase-kinase signaling complex that supports the pathogenesis of AML.

### Graphical Abstract

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

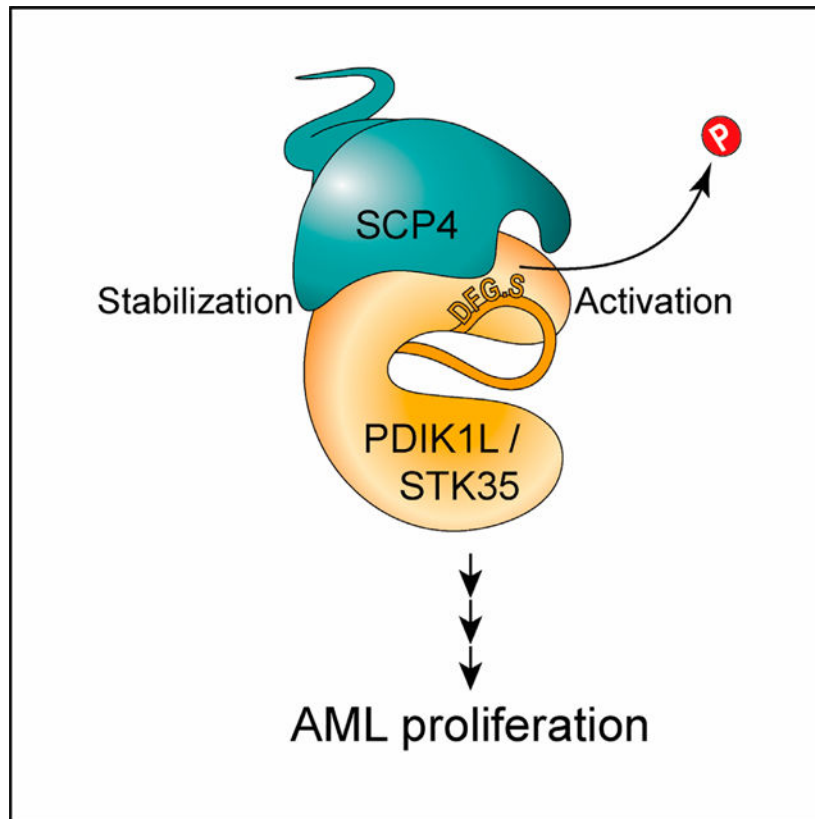
\*Correspondence: vakoc@cshl.edu.

#### AUTHORS CONTRIBUTIONS

S.A.P. and C.R.V. designed experiments and analyzed results. S.A.P. performed and analyzed experiments. R.Y.M., S.I., and Y.J.Z. purified and measured phosphatase activity of SCP4. B.L. performed and analyzed SCP4 CRISPR exon scans. S.A.P., R.F., and Y.Y. performed editing in human CD34<sup>+</sup> cells. S.A.P., O.K., and O.E.D. analyzed RNA-seq data. O.K. provided bi-cistronic vectors for double knockout. Z.Y. performed and analyzed cell-cycle arrest and apoptosis experiments. Y.W. performed and analyzed mouse xenograft experiments. L.A.B. performed several competition assays. S.A.P. and C.R.V. wrote the manuscript.

#### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2021.110233>.



### In brief

Polyanskaya et al. use unbiased genetic screening and biochemical strategies and find a phospho-signaling complex composed of a phosphatase SCP4 and the kinases STK35/PDIK1L. The catalytic activities of these proteins are critical in acute myeloid leukemia cells but dispensable in normal blood cells

## INTRODUCTION

Acute myeloid leukemia (AML) is an aggressive hematopoietic malignancy caused by the impaired differentiation and unrestrained proliferation of myeloid progenitor cells (Grove and Vassiliou, 2014). While genetically heterogeneous, most cases of AML harbor a mutation that deregulates signaling downstream of growth factor receptors (Bullinger et al., 2017). For example, activating mutations in the receptor tyrosine kinase FLT3 are present in one-third of AML cases, and inhibitors of this kinase are an approved therapy in this group of patients (Daver et al., 2019; Short et al., 2019; Kim, 2017; Dhillon, 2019). An example of a phosphatase target in leukemia is the oncoprotein SHP2, which is a positive regulator of RAS/MAP kinase signaling and can be targeted by allosteric small-molecule inhibitors (Elson, 2018; Yuan et al., 2020; Vainonen et al., 2021). Based on the actionability of kinase and phosphatase targets, a major goal in AML research is to discover and validate acquired dependencies on phospho-catalytic signaling proteins using functional genomics.

SCP4 (encoded by *CTDSPL2*) is one of the least-characterized human phosphatases, with yet-to-be-determined roles in human cancer. SCP4 belongs to the haloacid dehalogenase (HAD) family of proteins, which employ conserved active site aspartates within a DxDx(V/T) motif to remove phosphates from protein or non-protein substrates (Seifried et al., 2013). SCP4 localizes to the nucleus, where it can dephosphorylate several putative substrates, including transcription factors and RNA polymerase II (Zhao et al., 2014, 2018; Wani et al., 2016; Liu et al., 2017). SCP4<sup>-/-</sup> mice have been characterized and found to complete embryonic development, but these animals die shortly after birth due to defective gluconeogenesis (Cao et al., 2018). This phenotype was linked to the deregulation of FoxO transcription factors, suggesting that a critical function of SCP4 is to regulate metabolism (Cao et al., 2018). In chicken models of cancer, the *CTDSPL2* ortholog is a common integration site for Avian-leukosis-virus-induced B cell lymphomas (Winans et al., 2017). Despite belonging to a potentially druggable family of proteins (Krueger et al., 2014), it is unclear whether any human cancers rely on SCP4 for disease initiation or progression.

Another set of poorly characterized signaling proteins are the kinase paralogs STK35 and PDIK1L, which share 69% sequence identity. STK35 and PDIK1L localize to the nucleus, but the function of these enzymes in this cellular compartment is unclear. A small number of studies suggest that STK35 can regulate cell proliferation, apoptosis, and metabolism, albeit via unclear molecular mechanisms (Wu et al., 2018; Yang et al., 2020). In addition, the expression of STK35/PDIK1L correlates with disease progression and prognosis in several tissue contexts (Wu et al., 2018; Yang et al., 2020; Greenbaum et al., 2021). Nonetheless, the signaling function and biochemical interactions of STK35/PDIK1L remain unknown.

In this study, we used unbiased genetic and biochemical approaches to show that SCP4 forms a dual catalytic signaling complex with STK35/PDIK1L to support the pathogenesis of AML. Within this complex, SCP4 is critical for maintaining STK35/PDIK1L stability and also acts to remove inhibitory phosphorylation from the kinase activation loop. A downstream output of this phospho-catalytic complex in AML is to maintain the expression of genes involved in amino acid biosynthesis and transport. Taken together, our study identifies a nuclear SCP4-STK35/PDIK1L signaling complex and reveals its critical role in the pathogenesis of AML.

## RESULTS

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

We previously described a domain-focused CRISPR-Cas9 genetic screening strategy that can nominate catalytic dependencies in cancer by targeting insertion or deletion (indel) mutagenesis to domain-encoding exons (Shi et al., 2015). Here, we applied this method to the human phosphatase protein family in search of AML-specific dependencies. For this purpose, we designed and cloned a pooled library of 1,961 single-guide RNAs (sgRNAs) targeting exons that encode 217 known or putative phosphatase domains, including phosphatases with protein and non-protein substrates (Figure 1A).

We used this library to perform negative-selection “dropout” screens in eight leukemia and four solid tumor cell lines (Figures 1B, 1C, and S1A). The performance of spike-in negative and positive control sgRNAs in this library validated the overall quality of these screens (Figure S1B). This screen revealed that sgRNAs targeting 187 of the phosphatase domains showed no significant impairment of cell fitness in any of the cell lines tested, whereas 11 phosphatases were pan-essential across the entire panel of cell lines (Figures 1D and S1A; Table S1). The remainder showed varying levels of specificity, which included the validation of known context-specific phosphatase dependencies: SHP2 was only essential in cell lines that lacked activating mutations of RAS (Ahmed et al., 2019), PPM1D was only required in TP53 wild-type lines (Goloudina et al., 2016), and PTP1B was only a dependency in a BCR-ABL1-positive leukemia cell lines (Alvira et al., 2011) (Figures S1C–S1E).

Next, we ranked all of the phosphatases based on their relative essentiality in AML versus non-AML contexts, which nominated SCP4 as the top-ranking AML-biased dependency (Figures 1C and S1A). 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. The closest paralogs of SCP4 were not essential in AML (Figure 1C). Since no prior study has implicated SCP4 in human cancer, we investigated further the essential function of this phosphatase in AML.

### 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 and negative control sgRNAs in a total of 23 human cancer cell lines: 14 leukemia, 4 pancreatic cancer, 3 rhabdomyosarcoma, and 2 lung cancer lines (Figure 2A). In these experiments, ten leukemia cell lines demonstrated a clear dependency on SCP4, while four of the leukemia lines and all of the solid tumor cell lines were insensitive to SCP4 sgRNAs. The sensitive cell lines included engineered human MLL-AF9/Nras<sup>G12D</sup> and MLL-AF9/FLT3<sup>ITD</sup> 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. Western blot analysis confirmed that the differential SCP4 requirement was not because of differences in genome editing efficiency (Figures 2B–2E). Rescue experiments using an sgRNA-resistant cDNA confirmed that the growth arrest caused by SCP4 sgRNAs occurred via an on-target mechanism (Figures S2A–S2C). Using flow cytometry-based bromodeoxyuridine (BrdU) incorporation measurements and annexin V staining, we found that SCP4 inactivation in MOLM-13 cells led to a G1/G0-arrest and to the induction of apoptosis (Figures 2F, S2D, and S2E). We also validated that targeting SCP4 attenuated the growth of MOLM-13 cells when injected via tail vein and expanded in immune-deficient mice (Figures 2G and 2H). As additional validation, we verified that Scp4 was essential in an engineered murine AML cell line, MLL-AF9/Nras<sup>G12D</sup> (RN2) (Zuber et al., 2011), an effect that was rescued by the expression of the human SCP4 cDNA (Figure S2F). However, immortalized NIH3T3 murine fibroblasts were unaffected by SCP4 sgRNAs (Figure S2G), suggesting that the AML-biased dependency on SCP4 is conserved across species.

To evaluate the SCP4 requirement in non-transformed hematopoietic cells, we performed electroporation of Cas9:gRNA ribonucleoprotein complexes into human CD34<sup>+</sup> hematopoietic stem and progenitor cells (HSPCs) (Metais et al., 2019). After electroporation, successful inactivation of SCP4 was verified by western blotting and by measuring the presence of *CTDSPL2* indel mutations (Figures 2I–2J and S2H–S2L). 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 promote myeloid, erythroid, or megakaryocytic differentiation (Figures 2J, S2K, and S2L). In addition, flow cytometry analysis and methylcellulose plating experiments showed that SCP4-deficient HSPCs showed no defect in differentiation in these three lineages (Figures 2K and S2M–S2O). These findings are consistent with the survival of SCP4<sup>-/-</sup> 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 an acquired dependency in AML.

### The catalytic phosphatase activity of SCP4 is essential in AML

The C-terminal phosphatase domain of SCP4 (amino acids 236–442) is highly conserved among vertebrates, whereas amino acids 1–236 are less conserved and are predicted to be intrinsically disordered (Figures 3A and 3B) (Chang et al., 2018; Meszaros et al., 2018). In an initial effort to map the functionally important regions of SCP4 in AML, we performed CRISPR exon scanning (Shi et al., 2015). We constructed a pooled library of 85 sgRNAs tiling along each of the *CTDSPL2* exons and used this for negative selection CRISPR-Cas9 screening in MOLM-13 cells. By quantifying sgRNA depletion along the length of the gene, we observed a pattern in which sgRNAs targeting exons encoding amino acids 1–235 resulted in a minimal growth arrest phenotype, whereas sgRNAs targeting the 236–466 segment resulted in more striking negative selection (Figure 3C; Table S2). This result suggests that indel mutations produce a higher proportion of null alleles in the C-terminal segment of SCP4, likely due to the tendency for in-frame alleles to abrogate SCP4 functions when occurring at the catalytic phosphatase domain, yet in-frame alleles are likely to be functional when occurring at the SCP4 1–235 region (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 (Figure S3A), which explains why an AML dependency was likely to be underestimated 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 endogenous SCP4. Similar to results obtained with the wild-type protein, the SCP4<sup>236–466</sup> truncation was sufficient to rescue the knockout of endogenous *CTDSPL2*, verifying that the N-terminal half of SCP4 is dispensable in AML (Figures 3D–3F). In addition, we cloned point mutations of SCP4 that changed aspartate residues in the putative DxDx(V/T) catalytic motif to alanine. Despite similar expression levels to the wild-type protein, SCP4<sup>D293A</sup>, SCP4<sup>D295A</sup>, and SCP4<sup>D293A/D295A</sup> mutants were unable to support MOLM-13 growth (Figures 3G and 3H). Collectively, these genetic experiments suggest that the catalytic phosphatase function of SCP4 is essential in AML.

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

In accord with prior studies (Kemmer et al., 2006; Zhao et al., 2014; Kang et al., 2016), we found that SCP4 was localized to the nucleus of AML cells (Figure 4A). However, our efforts at performing SCP4 chromatin immunoprecipitation sequencing (ChIP-seq) in MOLM-13 cells failed to identify evidence of sequence-specific DNA occupancy (data not shown). Therefore, we turned to biochemical approaches to further understand the mechanism of SCP4 dependency. We stably expressed a FLAG-tagged SCP4 via lentivirus in MOLM-13 cells at levels comparable to endogenous SCP4 (Figure 4B). Importantly, this epitope-tagged protein localized to the nucleus and was able to rescue the knockout of endogenous *CTDSPL2* (Figures S4A, S2B, and S2C). Thereafter, we optimized an anti-FLAG affinity purification procedure that efficiently recovered FLAG-SCP4 complexes and removed endogenous SCP4 (Figure 4B). In these experiments, we compared wild-type SCP4 to SCP4<sup>D293A</sup> with the intent that any protein-protein interaction abrogated by this loss-of-function mutation would have a higher likelihood of functional significance.

Mass spectrometry (MS) analysis of purified FLAG-SCP4 was performed to identify the constituents of its protein complex. A comparison of 2 independent biological replicates revealed 59 protein candidates (Figure 4C). Among the most abundant SCP4-associated proteins- were several members of the importin protein family, which function to shuttle protein cargo into the nucleus (Cagatay and Chook, 2018) (Figure S4B). Our mapping experiments showed that the importin interaction occurred via the N-terminal segment 1–236, which contains predicted importin-binding nuclear localization sequences (Figures S4C–S4E). This result suggests that a function of the SCP4 N-terminal segment is to promote nuclear localization of the full-length protein. Interestingly, the low molecular weight of the SCP4<sup>236–466</sup> fragment (~38 kDa) would be expected to bypass the importin requirement for passage through the nuclear pore (Weis, 2003). This would explain the presence of SCP4<sup>236–466</sup> throughout the nucleus and cytoplasm and may account for the ability of this protein to complement the knockout of endogenous *CTDSPL2* (Figures S4A and S4E).

Among the SCP4-associated proteins identified in this analysis, the paralog kinases STK35 and PDIK1L were remarkable, as they were only detected in association with wild-type SCP4 but not with SCP4<sup>D293A</sup> (Figures 4D and 4E). As no commercially available immune reagents were able to detect these two poorly studied proteins, we cloned hemagglutinin (HA)-tagged proteins for further biochemical analysis. Upon transient transfection in HEK293T cells or using stable expression in MOLM-13, we found that SCP4<sup>WT</sup> and SCP4<sup>236–466</sup> each efficiently associated with HA-STK35 and HA-PDIK1L, whereas this interaction was diminished with the SCP4<sup>D293A</sup> protein (Figures 4F–4H). This interaction was also confirmed using reciprocal IP of HA-tagged STK35/PDIK1L with endogenous SCP4 (Figures S4F–4H). Moreover, an MS analysis of immunopurified 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 (Figure S4I). These biochemical experiments suggest that the kinases STK35 and PDIK1L exist in a stable complex with the catalytically active SCP4 phosphatase domain.

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

We previously performed kinase-domain-focused CRISPR screening in human cancer cell lines, which 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 (Figures 4D, S4B, and S6A), we hypothesized that functional redundancy might conceal their essentiality in this context. We investigated this by transducing human cancer cell lines with STK35 and PDIK1L sgRNAs, alone and in combination, and performing competition-based proliferation assays. These experiments revealed that MOLM-13 cells arrested their proliferation in response to the dual targeting of STK35 and PDIK1L, but the cells were less sensitive to targeting each gene individually (Figure 5A). In contrast, K562 cells did not show any phenotype following co-transduction with STK35/PDIK1L sgRNAs (Figure 5B). The proliferation-arrest phenotype of STK35/PDIK1L double knockout MOLM-13 cells was also fully rescued by expressing an sgRNA-resistant STK35 or PDIK1L cDNA (Figure 5C). Western blotting experiments confirmed the on-target editing of the lentivirally expressed wild-type cDNA but not of the engineered sgRNA-resistant cDNA (Figure 5D). By expanding this dual-targeting experiment to more leukemia lines using a bi-cistronic sgRNA vector, we found that the redundant STK35/PDIK1L dependency correlated with SCP4 dependency. For example, SCP4 and STK35/PDIK1L were not required in THP-1, U937, and K562 cells compared with the four sensitive AML cell lines (Figure 5E). These experiments demonstrate that STK35 and PDIK1L are redundant dependencies in AML and suggest that their function is linked to the SCP4 requirement in this context.

The analysis of AML patients' data from The Cancer Genome Atlas (TCGA) database (2013) revealed no significant differences in overall survival and most other clinical characteristics of the patients selected for the highest and the lowest expression of SCP4, PDIK1L, or STK35. We noticed the only significant clinical difference was in the French-American-British (FAB) classification (Bennett et al., 1985). SCP4- and PDIK1L-high AML samples were significantly associated with the M0 FAB status (minimally differentiated AML), whereas SCP4- and PDIK1L-low AML samples were enriched with the samples corresponding to the highly differentiated M5 FAB (monocytic leukemia) (Figures S5A and S5B). STK35-high samples were significantly associated with less differentiated M1 FAB subtypes, whereas STK35-low contained more M4 FAB samples (Figure S5C). These findings suggest that the expression of SCP4/STK35/PDIK1L is linked to the differentiation state of AML.

We next investigated whether the catalytic kinase function of STK35/PDIK1L was essential for AML cell proliferation. Sequence alignment identified the conserved ATP-binding lysine residue present in other catalytically active kinases (Cherry and Williams, 2004; Lucet et al., 2006) (Figure 5F). Guided by prior studies (Kamps and Sefton, 1986), we replaced this residue with histidine in our sgRNA-resistant STK35 and PDIK1L cDNA constructs and performed competition-based proliferation assays evaluating whether these mutant cDNAs can rescue the lethality of dual targeting of STK35/PDIK1L. Despite being expressed, albeit slightly below the level of wild-type proteins, STK35<sup>K231H</sup> and PDIK1L<sup>K37H</sup> behaved as null alleles in these assays, suggesting that the catalytic function of these kinase paralogs is essential in AML (Figures 5G and 5H).

To further investigate whether the SCP4 and STK35/PDIK1L co-dependency in AML reflects these proteins functioning in a common genetic pathway, we performed an early time point RNA-seq analysis following acute genetic knockout of SCP4 or STK35/PDIK1L. By comparing the global transcriptional changes upon these perturbations with data obtained following knock out of other essential genes in MOLM-13, we evaluated whether SCP4 and STK35/PDIK1L were functionally linked to each other. RNA was collected on day 5 for deep sequencing analysis following lentiviral infection with SCP4 or dual STK35/PDIK1L sgRNAs, and the results were compared with datasets produced following knock out of the essential genes TAF12, MYB, SIK3, and MEF2C in this same cell line (Xu et al., 2018; Tarumoto et al., 2018). Global correlation analysis of these datasets in all pairwise combinations recapitulated the TAF12-MYB and SIK3-MEF2C connectivity described previously (Xu et al., 2018; Tarumoto et al., 2018) (Figure 5I). Importantly, the transcriptional changes observed in SCP4- and STK35/PDIK1L-deficient MOLM-13 cells were closely correlated with one another when compared with these other knockouts. The transcriptome and the co-dependency correlation suggest that the physical interaction between SCP4 and STK35/PDIK1L reflects the functioning of these signaling molecules in a common genetic pathway.

### Biochemical evidence that SCP4 functions upstream of STK35/PDIK1L

We next investigated the mechanism that links SCP4 to STK35/PDIK1L. Using western blotting and cell fractionation, we found that knock out of endogenous SCP4 led to a marked reduction in the levels of both STK35 and PDIK1L (Figures 6A and 6B). These changes in STK35/PDIK1L did not occur at the mRNA level, which is indicative of an alteration of the protein (Figure S6A). In contrast, a double knock out of STK35/PDIK1L did not influence the SCP4 protein (Figure S6B). 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 proteomic databases for STK35/PDIK1L phosphorylation sites (Figure 6C) (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 loop of both kinases. Phosphorylation within the activation loop is a well-known mechanism to regulate kinase activity (Timm et al., 2008; Djouder et al., 2010; McCartney et al., 2016), which prompted us to investigate whether SCP4 removes phosphorylation from this segment. For this purpose, we synthesized four phosphoserine- or phosphothreonine-containing peptides representing the activation loop's phosphorylated residues and optimized an expression and purification strategy for the recombinant SCP4 phosphatase domain (residue 156–466), which was well-folded and exhibited phosphatase activity using para-nitrophenyl assays (Figures 6D and 6E; data not shown). Next, we quantified the phosphatase activity of SCP4 against each phosphopeptide using the malachite green assay to monitor phosphate release. We found that peptides corresponding to phosphoserines 194 and 216 of PDIK1L (S385 and S413 of STK35) were efficiently dephosphorylated by SCP4, whereas no activity was detected for the other two phosphopeptides (Figure 6D). The kinetic



parameters of these substrates were fitted into steady-state kinetics and  $k_{cat}/K_m$ -calculated as  $12.63 \pm 0.79$  and  $45.96 \pm 3.06$  mM<sup>-1</sup> min<sup>-1</sup> for phosphopeptides containing pS194 and pS216, respectively (Figures 6F and 6G).

We next performed gene complementation experiments to evaluate how alanine or phosphomimetic aspartate substitutions of these two phosphorylated serines of STK35/PDIK1L influence AML proliferation. While all of these mutant alleles tested were expressed at comparable levels to wild-type STK35/PDIK1L, the PDIK1L<sup>S194D</sup> and STK35<sup>S385D</sup> proteins were completely defective in supporting AML proliferation (Figures 6H, 6I, and 6C). In contrast, alanine substitutions of these same residues were fully functional, suggesting that mutations that mimic phosphorylation inhibit the function of these kinases. Notably, the PDIK1L<sup>S194</sup> and STK35<sup>S385</sup> 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). Collectively, these data support a model in which SCP4 removes inhibitory phosphorylation from the activation loop of STK35/PDIK1L, which may allow for kinase activation.

### SCP4-STK35/PDIK1L sustains a gene expression program of amino acid biosynthesis and transport

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 and STK35/PDIK1L double knockout in MOLM-13 cells (Figure 7A). Ontology analysis of these genes revealed a strong enrichment for amino acid biosynthesis and transport functions (Figure 7B). For example, all of 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 (Figure 7C). To evaluate whether these mRNA changes lead to corresponding changes in cellular amino acid levels, we performed MS to measure metabolites following SCP4 or STK35/PDIK1L inactivation in MOLM-13 cells. While the level of several metabolites remained unchanged in these knockout cells, many amino acids were present at reduced levels in cells deficient in either SCP4 or STK35/PDIK1L (Figure 7D). One of the most deficient amino acids was proline, which could be related to the decrease in *PYCR1* expression, encoding the rate-limiting enzyme for proline biosynthesis (Xiao et al., 2020) (Figure S7A). In accord with prior evidence describing aberrant amino acid metabolism in leukemia stem cells (LSCs) (Jones et al., 2018, 2020), we found that targeting of SCP4 led to diminished expression of transcriptional signatures of LSCs (Figure S7B) (Somervaille et al., 2009).

Since amino acid biosynthesis and transport are known to be deregulated to support the rapid growth of AML cells, at least in part, through maintaining the LSCs population (Jones et al., 2018; Van Galen et al., 2018; Bjelosevic et al., 2021; Kreitz et al., 2019; Raffel et al., 2020), these results lead us to propose that SCP4-STK35/PDIK1L signaling supports the aberrant metabolic state of this malignancy (Figure 7E).

## DISCUSSION

The main conclusion of this study is that the nuclear phosphatase SCP4 is a context-specific dependency in AML. Using biochemical and genetic approaches, we found that SCP4 executes this function by associating with the kinase paralogs STK35 and PDIK1L. While several phosphatases are known to regulate kinase activities by the removal of phosphorylation on the activation loop (Takekawa et al., 1998; Salmeen et al., 2000; Cheng et al., 2000; Hayward et al., 2019), relatively few examples exist of docking of phosphatase and kinase proteins to form a stable complex. One prominent example of such an interaction is the phosphatase KAP and its interaction with its kinase substrate, CDK2 (Hannon et al., 1994; Poon and Hunter, 1995). X-ray crystallography revealed that the binding interface between these two proteins involves extensive interactions remote from the active sites of both proteins, which position CDK2 for the removal of activating phosphorylation from its activation loop (Song et al., 2001). As another example, the phosphatase MKP3 stably associates with its kinase substrate ERK2 via a docking surface distinct from the catalytic active sites of each enzyme (Zhang et al., 2003; Liu et al., 2006). In each of these examples, the docking interaction leads to allosteric modulation of catalytic kinase and/or phosphatase activity. We speculate that a similar docking interaction and potential for allostery exists between SCP4 and STK35/PDIK1L, which may lead to the stabilization of the kinase substrate and the positioning of the kinase substrates for de-phosphorylation.

One paradoxical result in our study is the strength of the association between the catalytic domains of SCP4 and STK35/PDIK1L, yet this interaction is weakened by a single point mutation of the SCP4 active site. Since the interactions of enzymes with their substrates are often transient (Fahs et al., 2016), it is unlikely that the strong binding interaction we detect between SCP4 and STK35/PDIK1L is entirely due to an active site interaction with a phosphorylated substrate. While our experiments indicate that the activation loop of STK35/PDIK1L is a substrate of SCP4, we speculate that the binding interaction between SCP4 and STK35/PDIK1L includes a docking mechanism involving non-catalytic surfaces of these two proteins. One possibility is that the SCP4<sup>D293A</sup> protein becomes locked in an inactive conformation, which simultaneously impairs both catalysis and docking with STK35/PDIK1L. Future structural studies of the SCP4-STK35/PDIK1L complex coupled with deep mutagenesis could shed further light on the specific binding surface that links these two catalytic domains.

Our biochemical assays and mutational analyses implicate functionally important activation loop serines of STK35 (S385) and PDIK1L (S194) as substrates of SCP4. These serines lie at the +2 position relative to the DFG motif, a region of the activation loop that binds to a magnesium ion that interacts directly with ATP. Substantial evidence suggests that the DFG motif can adopt distinct open or closed conformations to regulate kinase activity (Kornev et al., 2006; Xie et al., 2020). While activation loop phosphorylation is known to promote kinase activity from several downstream locations, we hypothesize that incorporation of phosphorylation at the DFG+2 serine residue of STK35/PDIK1L locks the activation loop in inactive conformation that is incompatible with the precise ATP-magnesium coordination. This is consistent with the requirement for the 3-turn formation

between the DFG phenylalanine and the DFG+2 residue, as reported for other active kinases (Kornev et al., 2006; Xie et al., 2013).

The neonatal lethality phenotype of SCP4<sup>-/-</sup> mice can be rescued by an injection of 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 regulate FoxO transcription factors in hepatocytes, which, in turn, influences the expression of critical genes in the gluconeogenesis pathway (e.g., *PCK1* and *G6PC*). In AML, we find that SCP4 regulates instead the expression of a distinct group of metabolic genes that control amino acid biosynthesis and transport. Thus, it appears that distinct transcriptional outputs exist downstream of SCP4 in different cell types.

One of the questions remaining for further studies is why some leukemia cell lines exhibited a strong dependency on SCP4 and the paralog kinases, whereas others did not. The phenotype was not dependent on the expression of either of the genes and did not correlate with common mutations associated with carcinogenesis. A 2018 study from Pereira et al. showed that different AML cell lines had different metabolic profiles (Pereira et al., 2018). The differential dependency on SCP4 and STK35/PDIK1L may correlate with different metabolic requirements of these cell lines.

Another key question raised by our study is whether a specific transcription factor might be regulated by SCP4 or by STK35/PDIK1L to influence amino acid biosynthesis/transport genes. Interestingly, we found that many SCP4-regulated genes involved in metabolism are also target genes of the stress-induced transcription factor ATF4 (Harding et al., 2003). In the future, it would be interesting to determine whether SCP4 or STK35/PDIK1L directly modify the phosphorylation state of this protein to influence metabolism.

### Limitations of study

While our data indicate that SCP4 essentiality is context-specific, additional investigation is required to define whether a therapeutic window exists in AML when targeting this phosphatase. An appropriate follow-up to this study would include experiments targeting SCP4 *in vivo* in transplanted CD34<sup>+</sup> cells and in primary human AML cells *in vitro* and *in vivo*. Generation of genetically engineered mouse models would also be instrumental in the further development of SCP4 as a therapeutic target.

While our study defines several functions of SCP4 in AML, the precise downstream pathways that account for the SCP4 dependency in AML remain to be discovered. Genetic rescue (or epistasis) experiments would be necessary in this endeavor, which would include manipulating the levels of the two kinases or the metabolite levels to determine whether this alleviates SCP4 dependency.

Our *in vitro* biochemical experiments suggested that SCP4 removes the inhibitory phosphorylation site from the kinases. More experiments are needed to definitively prove that SCP4 dephosphorylates PDIK1L and STK35 *in vivo*. These studies would require specific antibodies to the phosphosites of interest and the endogenous proteins to determine phosphorylation levels upon SCP4 knockout and overexpression.

## STAR★METHODS

### RESOURCE AVAILABILITY

**Lead contact**—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Christopher Vakoc (vakoc@cshl.edu).

**Materials availability**—Reagents are available upon request from the lead contact.

#### Data and code availability

- All RNA-Seq data used in this study is publicly available. Raw and processed sequencing data for RNA-seq have been deposited to NCBI Gene Expression Omnibus with the accession numbers listed in the key resources table.

All the proteomics/metabolomics data used in this study is publicly available with the accession numbers listed in the key resources table.

- This study does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

**Cell lines**—Human cell lines MOLM-13, NOMO-1, MV4–11, ML-2, HEL, SET-2, THP-1, 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<sup>G12D</sup> AML) (Zuber et al., 2011) 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; Wunderlich et al., 2013) derivatives MA9-NRAS<sup>G12D</sup> and MA9-FLT3<sup>ITD</sup> (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% CO<sub>2</sub> and were regularly tested mycoplasma negative.

**Mice**—Female immunodeficient mice (NSG) at the age of 8 weeks were used for injection. All animals were purchased from The Jackson Laboratory (Stock No: 005557). 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 (Wei et al., 2021). These cells were infected with lentivirus encoding GFP-linked sgRNAs targeting either ROSA locus (negative control) or CTDSPL2 (SCP4). 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 6 sublethally (2.5Gy) irradiated mice (3 mice for sgROSA, 3 mice for sgSCP4) 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.

All animal procedures and studies were conducted in accordance with the Institutional Animal Care and Use Committee (IACUC) at Cold Spring Harbor Laboratory.

## METHOD DETAILS

**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 CDS. LRG2.1-Blast was generated by the introduction of the blasticidin resistance gene after GFP CDS with a P2A linker. In this study, LRG2.1, LRG2.1-Blast, or LRCherry2.1 were used for introducing sgRNA into human cell lines, while 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 Table S3 (Table S3 sgRNAs sequences, CRISPR-resistant constructs, and primers. Related to all Figures).

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 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 mutants used in this study can be found in the Table S3 (Table S3 sgRNAs sequences, CRISPR-resistant constructs, and primers. Related to all Figures).

**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. Six to thirty independent 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> (El Demerdash, unpublished) 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).

**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<sup>TM</sup>; 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 × 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.

**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<sup>TM</sup> 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. F548S). The final libraries were quantified using Bioanalyzer Agilent High Sensitivity DNA Kit (Agilent 5067-4626) 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 Table S1 (Table S1. Average  $\log_2$  FC in phosphatase domain-focused screen. Related to Figures 1 and S1). The SCP4 CRISPR exon scanning screening data is provided in the Table S2 (Table S2. Average  $\log_2$  fold-change in SCP4 CRISPR exon scanning screen in MOLM-13. Related to Figures 3 and S3).

**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.

**Western Blot**—For knock-out 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 2-mercaptoethanol, 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 2-mercaptoethanol.

The samples were separated by SDS-PAGE (NuPAGE 4–12% Bis-Tris Protein Gels, ThermoFisher), followed by transfer to nitro-cellulose 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 (ECL) 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),  $\beta$ -Actin-HRP (Sigma A3854-200UL; 1:50,000).

**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).

**Editing and differentiation of human peripheral blood CD34<sup>+</sup> cells**—Circulating G-CSF-mobilized human CD34<sup>+</sup> cells were obtained from two de-identified healthy donors (KeyBiologics). Enrichment of CD34<sup>+</sup> cells was performed by immunomagnetic bead selection using an CliniMACS instrument (Miltenyi Biotec). After enrichment, the CD34<sup>+</sup> cell fraction was 95% of the total, with <0.2% CD3<sup>+</sup> cells and <2% CD19<sup>+</sup> 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<sup>+</sup> cells was performed as described before (Metais et al., 2019). Briefly, cryopreserved CD34<sup>+</sup> cells were thawed and pre-stimulated for 48 hours in the maintenance medium (Table S4). 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  $\mu$ L of 10 mM HEPES (Thermo Fisher Scientific, catalog # 15630080), 150 mM NaCl (Thermo Fisher Scientific, catalog # 9759) for 35 minutes. 1 million CD34<sup>+</sup> cells resuspended in 50  $\mu$ L of T buffer were mixed with RNP complex in a final volume of 100  $\mu$ 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 (Table S4) 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 Table S4. 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 CT values. The percentage of editing in CD34<sup>+</sup> cells electroporated with sgRNAs vs. Cas9 alone was approximated as  $1 - 2^{(-CT)}$ . For the myeloid lineage, Surveyor nuclease assay was performed with primers in Table S4, according to the manufacturer protocol (Integrated DNA Technologies, Cat. No. 706020) (Table S4. sgRNAs; media and cytokines; and antibodies used in flow cytometry panels in CD34<sup>+</sup> experiments. Related to Figures 2 and S2).

For the CFC assay, ~800 CD34<sup>+</sup> 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 Table S4 (Table S4 sgRNAs; media and cytokines; and antibodies used in flow cytometry panels in CD34<sup>+</sup> experiments. Related to Figures 2 and S2).

#### **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 (WB), 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 MgCl<sub>2</sub>; 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  $11,000 \times g$  at  $4^{\circ}\text{C}$ . 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  $\text{MgCl}_2$ ; 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  $4^{\circ}\text{C}$  for 10 minutes at  $21,000 \times g$ . The supernatant was diluted to 150 mM NaCl using No-salt dilution buffer (20 mM HEPES pH 7.9; 1.5 mM  $\text{MgCl}_2$ ; 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<sup>TM</sup> 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  $4^{\circ}\text{C}$  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 WB the beads were boiled for 10 minutes in 1X Laemmli Sample Buffer (Bio-Rad, Cat. No. 1610737) containing 2-mercaptoethanol. For MS, the beads were send on ice to the Proteomics and Metabolomics Facility Center for Biotechnology at the University of Nebraska–Lincoln.

**Protein identification by mass spectrometry (MS)**—Bead samples were made up in 60 $\mu\text{L}$  1x non-reducing LDS sample buffer and incubated at  $95^{\circ}\text{C}$  for 10 min prior to loading 50 $\mu\text{L}$  of sample onto a Bolt<sup>TM</sup> 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  $37^{\circ}\text{C}$ . Peptides were extracted from the gel pieces, dried down, and samples were re-dissolved in 30 $\mu\text{L}$ , 2.5% acetonitrile, 0.1% formic acid. 5 $\mu\text{L}$  of each digest was run by nanoLC-MS/MS using a 2h gradient on a 0.075 mm  $\times$  250 mm CSH C18 column feeding into a Q-Exactive HF mass spectrometer.

All MS/MS 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 PPM. 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.

**Co-IP in HEK 293T cells**—One 80% confluent 10-cm plate of 293T cells was transfected with 10 µg HA-PDIK1L and 10 µg of either Empty vector, or FLAG-SCP4, or FLAG-235aa-SCP4, or FLAG-SCP4<sup>D293A</sup>. Total 20 µg DNA was mixed with 60 µL 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 µL of ice-cold Cell Lysis Buffer (CST, #9803). Incubated on ice for 30 minutes. Centrifuged at 4°C for 10 minutes at 14,000 × 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 2-mercaptoethanol, followed by Western blot analysis.

**RNA-Seq**—For knock-out 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 µL chloroform and incubated for 3 minutes at room temperature, followed by centrifugation at 10,000 × 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 × 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 µL 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). 4 independent biological replicates for control samples, samples

from cells with SCP4 knock-out, and cells with STK35-PDIK1L double knock-out were sequenced in the same flow cell.

**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 4 replicates for each sample (Love et al., 2014). Pseudo-alignment to the reference human transcriptome (gen-code 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 4 replicates for each sample (Love et al., 2014). Pearson's product-moment correlations between global  $\log_2$  fold changes observed upon different knockouts in MOLM-13 cells were calculated using *cor()* function in R with default parameters. Gene Ontology analysis was performed using Metascape (Szklarczyk et al., 2014; Tripathi et al., 2015).

**Protein expression and purification**—The SCP4 phosphatase domain (encoding residues 156–466) was subcloned 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 OD<sub>600</sub> 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 2-mercaptoethanol (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 M 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.

**Phosphatase activity assay**—Initial rates of phosphatase activity were determined by incubating SCP4 (1 $\mu$ M) with various concentrations of phosphopeptide (0–100  $\mu$ M) in SCP4 assay buffer (50 mM Tris-acetate pH 5.0, 10 mM MgCl<sub>2</sub>) in a 40 $\mu$ L reaction volume for 10 minutes at 37°C. 20 $\mu$ L of the reaction was then quenched with 40 $\mu$ 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<sub>620</sub> in a Tecan Plate reader 200. The A<sub>620</sub> values were corrected by subtracting the A<sub>620</sub> 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.

**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% H<sub>2</sub>O) 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 couple to a Q Exactive HF mass spectrometer (both Thermo Fisher Scientific). Sample extracts (8  $\mu$ L) were injected onto a Sequant ZIC-pHILIC column (150 mm  $\times$  2.1 mm, 5  $\mu$ m) and guard column (20 mm  $\times$  2.1 mm, 5  $\mu$ m) 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  $\mu$ L/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.

## QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed in GraphPad Prism (versions 5 and 9), with details provided in the corresponding figure legends, such as the number of replicates and statistical tests employed. Error bars represent the mean plus or minus standard error of the mean, unless otherwise stated. No methods were used to determine whether the data met assumptions of the statistical approach.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGMENTS

We thank the members of the Vakoc, Weiss, and Zhang laboratories, the CSHL Core facilities, and shared resources for making this work possible. We thank the Proteomics & Metabolomics Facility (RRID: SCR\_021314), Nebraska Center for Biotechnology at the University of Nebraska-Lincoln, for the mass spectrometry analysis, specifically Mike Naldrett and Sophie Alvarez. The facility and instrumentation are supported by the Nebraska Research Initiative. We thank the CSHL Mass Spectrometry Shared Resource, which is supported by the Cancer Center Support Grant 5P30CA045508. Special thanks to Osama El Demerdash for help at the final stages of the manuscript submission. This work was supported by Cold Spring Harbor Laboratory NCI Cancer Center Support grant 5P30CA045508. Additional funding was provided to C.R.V. by the Pershing Square Sohn Cancer Research Alliance, National Institutes of Health grants R01 CA174793 and P01 CA013106, a Leukemia & Lymphoma Society Scholar Award, and a grant from the CSHL-Northwell Alliance. Y.J.Z. received grants from the National Institutes of Health (R01 GM104896 and 125882).

## DECLARATION OF INTERESTS

C.R.V. has received consulting fees from Flare Therapeutics, Roivant Sciences, and C4 Therapeutics, has served on the scientific advisory board of KSQ Therapeutics, Syros Pharmaceuticals, and Treeline Biosciences, has received research funding from Boehringer-Ingelheim and Treeline Biosciences, and has a stock option from Treeline Biosciences. All other authors declare no competing interests.

## REFERENCES

- Ahmed TA, Adamopoulos C, Karoulia Z, Wu X, Sachidanandam R, Aaronson SA, and Poulikakos PI (2019). SHP2 drives adaptive resistance to ERK signaling inhibition in molecularly defined subsets of ERK-dependent tumors. *Cell Rep.* 26, 65–78.e65. 10.1016/j.celrep.2018.12.013. [PubMed: 30605687]
- Alvira D, Naughton R, Bhatt L, Tedesco S, Landry WD, and Cotter TG (2011). Inhibition of protein-tyrosine phosphatase 1B (PTP1B) mediates ubiquitination and degradation of Bcr-Abl protein. *J. Biol. Chem* 286, 32313–32323. 10.1074/jbc.M111.249060. [PubMed: 21795709]
- Anders S, Pyl PT, and Huber W (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. 10.1093/bioinformatics/btu638. [PubMed: 25260700]
- Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, and Sultan C (1985). Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. *Ann Intern Med* 103, 620–625. [PubMed: 3862359]
- Bjelosevic S, Gruber E, Newbold A, Shembrey C, Devlin JR, Hogg SJ, Kats L, Todorovski I, Fan Z, Abreht TC, et al. (2021). Serine biosynthesis is a metabolic vulnerability in FLT3-ITD-driven acute myeloid leukaemia. *Cancer Discov*, 1582–1599. 10.1158/2159-8290.CD-20-0738. [PubMed: 33436370]
- Bray NL, Pimentel H, Melsted P, and Pachter L (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol* 34, 525–527. 10.1038/nbt.3519. [PubMed: 27043002]
- Bullinger L, Dohner K, and Dohner H (2017). Genomics of acute myeloid leukemia diagnosis and pathways. *J. Clin. Oncol* 35, 934–946. 10.1200/JCO.2016.71.2208. [PubMed: 28297624]
- Cagatay T, and Chook YM (2018). Karyopherins in cancer. *Curr. Opin. Cell Biol* 52, 30–42. 10.1016/j.ceb.2018.01.006. [PubMed: 29414591]
- Cao J, Yu Y, Zhang Z, Chen X, Hu Z, Tong Q, Chang J, Feng XH, and Lin X (2018). SCP4 promotes gluconeogenesis through FoxO1/3a dephosphorylation. *Diabetes* 67, 46–57. 10.2337/db17-0546. [PubMed: 28851713]
- Chang KT, Guo J, di Ronza A, and Sardiello M (2018). Aminode: identification of evolutionary constraints in the human Proteome. *Sci. Rep* 8, 1357. 10.1038/s41598-018-19744-w. [PubMed: 29358731]
- Cheng A, Kaldis P, and Solomon MJ (2000). Dephosphorylation of human cyclin-dependent kinases by protein phosphatase type 2C alpha and beta 2 isoforms. *J. Biol. Chem* 275, 34744–34749. 10.1074/jbc.M006210200. [PubMed: 10934208]
- Cherry M, and Williams DH (2004). Recent kinase and kinase inhibitor X-ray structures: mechanisms of inhibition and selectivity insights. *Curr. Med. Chem* 11, 663–673. 10.2174/0929867043455792. [PubMed: 15032722]

- Daver N, Schlenk RF, Russell NH, and Levis MJ (2019). Targeting FLT3 mutations in AML: review of current knowledge and evidence. *Leukemia* 33, 299–312. 10.1038/s41375-018-0357-9. [PubMed: 30651634]
- Dempster JM, Rossen J, Kazachkova M, Pan J, Kugener G, Root DE, and Tsherniak A (2019). Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines. *bioRxiv*, 720243. 10.1101/720243.
- Dhillon S (2019). Gilteritinib: first global approval. *Drugs* 79, 331–339. 10.1007/s40265-019-1062-3. [PubMed: 30721452]
- Djouder N, Tuerk RD, Suter M, Salvioni P, Thali RF, Scholz R, Vaahntomeri K, Auchli Y, Rechsteiner H, Brunisholz RA, et al. (2010). PKA phosphorylates and inactivates AMPKalpha to promote efficient lipolysis. *EMBO J.* 29, 469–481. 10.1038/emboj.2009.339. [PubMed: 19942859]
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. 10.1093/bioinformatics/bts635. [PubMed: 23104886]
- Elson A (2018). Stepping out of the shadows: oncogenic and tumor-promoting protein tyrosine phosphatases. *Int. J. Biochem. Cell Biol* 96, 135–147. 10.1016/j.biocel.2017.09.013. [PubMed: 28941747]
- Fahs S, Lujan P, and Kohn M (2016). Approaches to study phosphatases. *ACS Chem. Biol* 11, 2944–2961. 10.1021/acscchembio.6b00570.
- Goloudina AR, Kochetkova EY, Pospelova TV, and Demidov ON (2016). Wip1 phosphatase: between p53 and MAPK kinases pathways. *Oncotarget* 7, 31563–31571. 10.18632/oncotarget.7325. [PubMed: 26883196]
- Greenbaum H, Galper BL, Decter DH, and Eisenberg VH (2021). Endometriosis and autoimmunity: can autoantibodies be used as a non-invasive early diagnostic tool? *Autoimmun. Rev* 20, 102795. 10.1016/j.autrev.2021.102795. [PubMed: 33722753]
- Grove CS, and Vassiliou GS (2014). Acute myeloid leukaemia: a paradigm for the clonal evolution of cancer? *Dis. Model Mech* 7, 941–951. 10.1242/dmm.015974. [PubMed: 25056697]
- Hannon GJ, Casso D, and Beach D (1994). KAP: a dual specificity phosphatase that interacts with cyclin-dependent kinases. *Proc. Natl. Acad. Sci. U S A* 91, 1731–1735. 10.1073/pnas.91.5.1731. [PubMed: 8127873]
- Harding HP, Zhang Y, Zeng H, Novoa I, Lu PD, Calfon M, Sadri N, Yun C, Popko B, Paules R, et al. (2003). An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol. Cell* 11, 619–633. 10.1016/s1097-2765(03)00105-9. [PubMed: 12667446]
- Hayward D, Bancroft J, Mangat D, Alfonso-Perez T, Dugdale S, McCarthy J, Barr FA, and Gruneberg U (2019). Checkpoint signaling and error correction require regulation of the MPS1 T-loop by PP2A-B56. *J. Cell Biol* 218, 3188–3199. 10.1083/jcb.201905026. [PubMed: 31511308]
- 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. 10.1038/nbt.3290. [PubMed: 26121415]
- Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, and Skrzypek E (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* 43, D512–D520. 10.1093/nar/gku1267. [PubMed: 25514926]
- 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. 10.1038/nbt.2647. [PubMed: 23873081]
- 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* 34, 724–740.
- 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. [PubMed: 32822582]

- Kamps MP, and 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. 10.1128/mcb.6.3.751. [PubMed: 2430174]
- Kang N, Koo J, Wang S, Hur SJ, and Bahk YY (2016). A systematic study of nuclear interactome of C-terminal domain small phosphatase-like 2 using inducible expression system and shotgun proteomics. *BMB Rep.* 49, 319–324. 10.5483/bmbrep.2016.49.6.240. [PubMed: 26674342]
- Keller A, Nesvizhskii AI, Kolker E, and 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. 10.1021/ac025747h. [PubMed: 12403597]
- Kemmer D, Podowski RM, Arenillas D, Lim J, Hodges E, Roth P, Sonn-hammer EL, Hoog C, and Wasserman WW (2006). NovelFam3000–un-characterized human protein domains conserved across model organisms. *BMC Genomics* 7, 48. 10.1186/1471-2164-7-48. [PubMed: 16533400]
- Kim ES (2017). Midostaurin: first global approval. *Drugs* 77, 1251–1259. 10.1007/s40265-017-0779-0. [PubMed: 28612232]
- Kornev AP, Haste NM, Taylor SS, and 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. 10.1073/pnas.0607656103. [PubMed: 17095602]
- Kreitz J, Schonfeld C, Seibert M, Stolp V, Alshamleh I, Oellerich T, Steffen B, Schwalbe H, Schnutgen F, Kurrel N, and Serve H (2019). Metabolic plasticity of acute myeloid leukemia. *Cells* 8, 805. 10.3390/cells8080805.
- 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. 10.1074/jbc.M114.566729. [PubMed: 24755226]
- Liu S, Sun JP, Zhou B, and 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. 10.1073/pnas.0510506103. [PubMed: 16567630]
- Liu X, Yu R, Sun L, Garibotto G, Lin X, Wang Y, Thomas SS, Li R, and Hu Z (2017). The nuclear phosphatase SCP4 regulates FoxO transcription factors during muscle wasting in chronic kidney disease. *Kidney Int.* 92, 336–348. 10.1016/j.kint.2017.02.031. [PubMed: 28506762]
- Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]
- Lucet IS, Fantino E, Styles M, Bamert R, Patel O, Broughton SE, Walter M, Burns CJ, Treutlein H, Wilks AF, and Rossjohn J (2006). The structural basis of Janus kinase 2 inhibition by a potent and specific pan-Janus kinase inhibitor. *Blood* 107, 176–183. 10.1182/blood-2005-06-2413. [PubMed: 16174768]
- Mackay GM, Zheng L, van den Broek NJ, and Gottlieb E (2015). Analysis of cell metabolism using LC-MS and isotope tracers. *Methods Enzymol.* 561, 171–196. 10.1016/bs.mie.2015.05.016. [PubMed: 26358905]
- McCartney RR, Garnar-Wortzel L, Chandrashekarappa DG, and Schmidt MC (2016). Activation and inhibition of Snf1 kinase activity by phosphorylation within the activation loop. *Biochim. Biophys. Acta* 1864, 1518–1528. 10.1016/j.bbapap.2016.08.007. [PubMed: 27524664]
- Meszaros B, Erdos G, and 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. 10.1093/nar/gky384. [PubMed: 29860432]
- 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. 10.1182/bloodadvances.2019000820. [PubMed: 31698466]
- 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. 10.1038/ng.3984. [PubMed: 29083409]



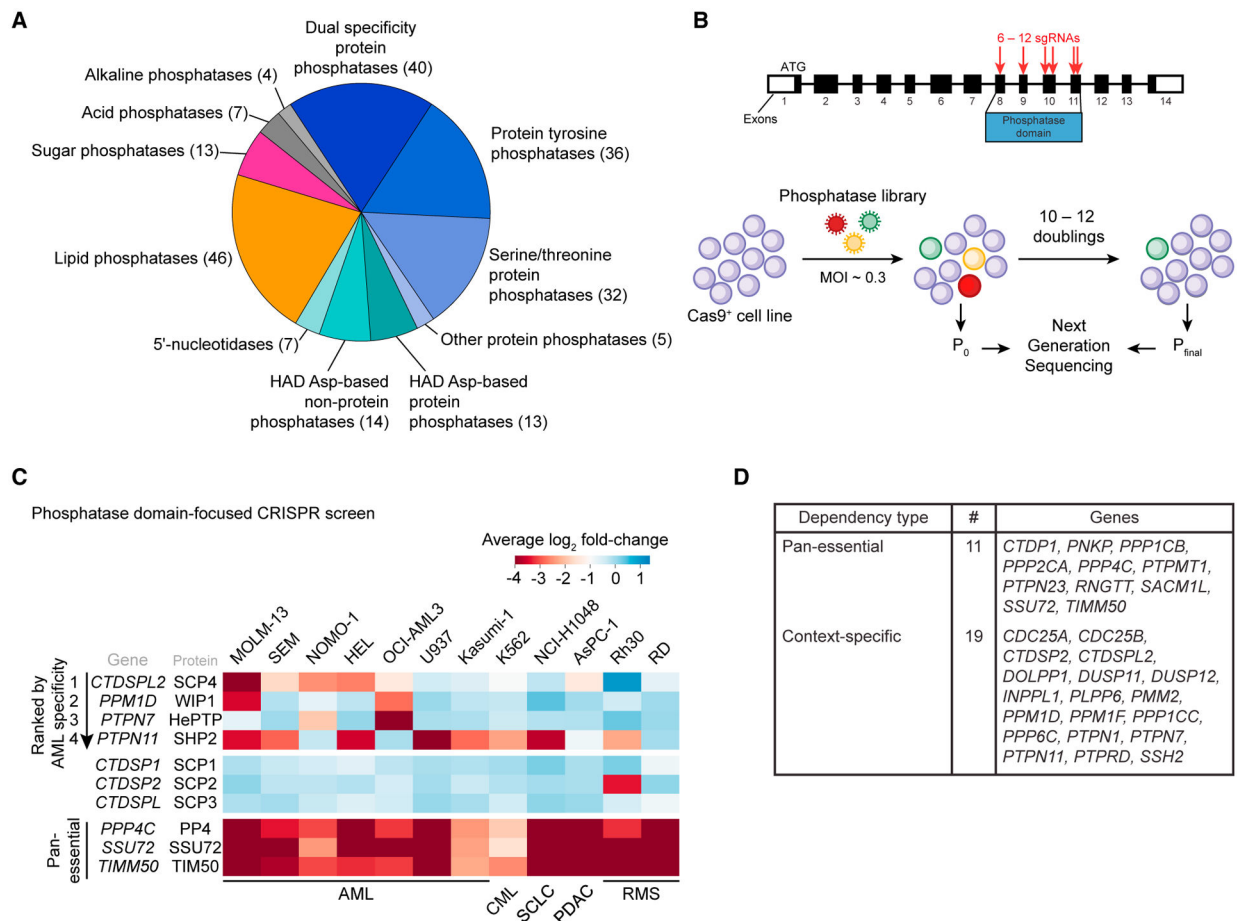
- Mulloy JC, Wunderlich M, Zheng Y, and Wei J (2008). Transforming human blood stem and progenitor cells: a new way forward in leukemia modeling. *Cell Cycle* 7, 3314–3319. 10.4161/cc.7.21.6951. [PubMed: 18948748]
- Nesvizhskii AI, Keller A, Kolker E, and Aebersold R (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem.* 75, 4646–4658. 10.1021/ac0341261. [PubMed: 14632076]
- 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. 10.1038/s41587-019-0344-3. [PubMed: 31819260]
- Patro R, Duggal G, Love MI, Irizarry RA, and Kingsford C (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. 10.1038/nmeth.4197. [PubMed: 28263959]
- Pereira O, Teixeira A, Sampaio-Marques B, Castro I, Girao H, and Ludovico P (2018). Signalling mechanisms that regulate metabolic profile and autophagy of acute myeloid leukaemia cells. *J Cell Mol Med* 22, 4807–4817. [PubMed: 30117681]
- Poon RY, and Hunter T (1995). Dephosphorylation of Cdk2 Thr160 by the cyclin-dependent kinase-interacting phosphatase KAP in the absence of cyclin. *Science* 270, 90–93. 10.1126/science.270.5233.90. [PubMed: 7569954]
- Raffel S, Klimmeck D, Falcone M, Demir A, Pouya A, Zeisberger P, Lutz C, Tinelli M, Bischel O, Bullinger L, et al. (2020). Quantitative proteomics reveals specific metabolic features of acute myeloid leukemia stem cells. *Blood* 136, 1507–1519. [PubMed: 32556243]
- Salmeen A, Andersen JN, Myers MP, Tonks NK, and Barford D (2000). Molecular basis for the dephosphorylation of the activation segment of the insulin receptor by protein tyrosine phosphatase 1B. *Mol. Cell* 6, 1401–1412. 10.1016/s1097-2765(00)00137-4. [PubMed: 11163213]
- Seifried A, Schultz J, and Gohla A (2013). Human HAD phosphatases: structure, mechanism, and roles in health and disease. *FEBS J.* 280, 549–571. 10.1111/j.1742-4658.2012.08633.x. [PubMed: 22607316]
- Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB, and Vakoc CR (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol* 33, 661–667. 10.1038/nbt.3235. [PubMed: 25961408]
- Short NJ, Kantarjian H, Ravandi F, and Daver N (2019). Emerging treatment paradigms with FLT3 inhibitors in acute myeloid leukemia. *Ther. Adv. Hematol* 10.1177/2040620719827310.
- Somerville TC, Matheny CJ, Spencer GJ, Iwasaki M, Rinn JL, Witten DM, Chang HY, Shurtleff SA, Downing JR, and Cleary ML (2009). Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells. *Cell Stem Cell* 4, 129–140. [PubMed: 19200802]
- Song H, Hanlon N, Brown NR, Noble ME, Johnson LN, and Barford D (2001). Phosphoprotein-protein interactions revealed by the crystal structure of kinase-associated phosphatase in complex with phosphoCDK2. *Mol. Cell* 7, 615–626. 10.1016/s1097-2765(01)00208-8. [PubMed: 11463386]
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, et al. (2014). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. 10.1093/nar/gku1003. [PubMed: 25352553]
- Takekawa M, Maeda T, and Saito H (1998). Protein phosphatase 2 $\alpha$  inhibits the human stress-responsive p38 and JNK MAPK pathways. *EMBO J.* 17, 4744–4752. 10.1093/emboj/17.16.4744. [PubMed: 9707433]
- Tarumoto Y, Lu B, Somerville TDD, Huang YH, Milazzo JP, Wu XS, Klingbeil O, El Demerdash O, Shi J, and Vakoc CR (2018). LKB1, salt-inducible kinases, and MEF2C are linked dependencies in acute myeloid leukemia. *Mol. Cell* 69, 1017–1027.e1016. 10.1016/j.molcel.2018.02.011. [PubMed: 29526696]
- Timm T, Balusamy K, Li X, Biernat J, Mandelkow E, and Mandelkow EM (2008). Glycogen synthase kinase (GSK) 3 $\beta$  directly phosphorylates Serine 212 in the regulatory loop and inhibits

- microtubule affinity-regulating kinase (MARK) 2. *J. Biol. Chem* 283, 18873–18882. 10.1074/jbc.M706596200. [PubMed: 18424437]
- Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM, DeJesus P, Che J, Mulder LC, et al. (2015). Meta- and orthogonal integration of influenza “OMICs” data defines a role for UBR4 in virus budding. *Cell Host Microbe* 18, 723–735. 10.1016/j.chom.2015.11.002. [PubMed: 26651948]
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, and Rozen SG (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 40, e115. 10.1093/nar/gks596. [PubMed: 22730293]
- Vainonen JP, Momeny M, and Westermarck J (2021). Druggable cancer phosphatases. *Sci. Transl Med* 13, eabe2967. 10.1126/sci-translmed.abe2967. [PubMed: 33827975]
- Van Galen P, Mbong N, Kreso A, Schoof EM, Wagenblast E, Ng SWK, Krivdova G, Jin L, Nakauchi H, and Dick JE (2018). Integrated Stress Response Activity Marks Stem Cells in Normal Hematopoiesis and Leukemia. *Cell Rep* 25, 1109–1117. [PubMed: 30380403]
- Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, and Sabatini DM (2015). Identification and characterization of essential genes in the human genome. *Science* 350, 1096–1101. 10.1126/science.aac7041. [PubMed: 26472758]
- Wani S, Sugita A, Ohkuma Y, and Hirose Y (2016). Human SCP4 is a chromatin-associated CTD phosphatase and exhibits the dynamic translocation during erythroid differentiation. *J. Biochem* 160, 111–120. 10.1093/jb/mvw018. [PubMed: 26920047]
- 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. 10.1016/j.ccr.2008.04.020. [PubMed: 18538732]
- Wei Y, Huang YH, Skopelitis DS, Iyer SV, Costa ASH, Yang Z, Kramer M, Adelman ER, Klingbeil O, Demerdash OE, et al. (2021). SLC5A3-dependent myo-inositol auxotrophy in acute myeloid leukemia. *Cancer Discov.* 10.1158/2159-8290.CD-20-1849.
- Weis K (2003). Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. *Cell* 112, 441–451. 10.1016/s0092-8674(03)00082-5. [PubMed: 12600309]
- Winans S, Flynn A, Malhotra S, Balagopal V, and Beemon KL (2017). Integration of ALV into CTDSPL and CTDSPL2 genes in B-cell lymphomas promotes cell immortalization, migration and survival. *Oncotarget* 8, 57302–57315. 10.18632/oncotarget.19328. [PubMed: 28915671]
- Wu Z, Liu J, Hu S, Zhu Y, and 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. 10.1159/000487172. [PubMed: 29414823]
- Wunderlich M, Mizukawa B, Chou FS, Sexton C, Shrestha M, Saunthararajah Y, and Mulloy JC (2013). AML cells are differentially sensitive to chemotherapy treatment in a human xenograft model. *Blood* 121, e90–e97. 10.1182/blood-2012-10-464677. [PubMed: 23349390]
- Xiao S, Li S, Yuan Z, and 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. 10.21037/atm-19-4402. [PubMed: 32953737]
- Xie T, Peng W, Yan C, Wu J, Gong X, and Shi Y (2013). Structural insights into RIP3-mediated necroptotic signaling. *Cell Rep.* 5, 70–78. 10.1016/j.celrep.2013.08.044. [PubMed: 24095729]
- Xie T, Saleh T, Rossi P, and Kalodimos CG (2020). Conformational states dynamically populated by a kinase determine its function. *Science* 370, eabc2754. 10.1126/science.abc2754. [PubMed: 33004676]
- Xu Y, Milazzo JP, Somerville TDD, Tarumoto Y, Huang YH, Ostrander EL, Wilkinson JE, Challen GA, and Vakoc CR (2018). A TFIID-SAGA perturbation that targets MYB and suppresses acute myeloid leukemia. *Cancer Cell* 33, 13–28.e18. 10.1016/j.ccell.2017.12.002. [PubMed: 29316427]
- Yang H, Zhu J, Wang G, Liu H, Zhou Y, and 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. 10.3389/fcell.2020.582695. [PubMed: 33117809]

- Yuan X, Bu H, Zhou J, Yang CY, and Zhang H (2020). Recent advances of SHP2 inhibitors in cancer therapy: current development and clinical application. *J. Med. Chem* 63, 11368–11396. 10.1021/acs.jmedchem.0c00249. [PubMed: 32460492]
- Zhang J, Zhou B, Zheng CF, and Zhang ZY (2003). A bipartite mechanism for ERK2 recognition by its cognate regulators and substrates. *J. Biol. Chem* 278, 29901–29912. 10.1074/jbc.M303909200. [PubMed: 12754209]
- Zhao Y, Liu J, Chen F, and Feng XH (2018). C-terminal domain small phosphatase-like 2 promotes epithelial-to-mesenchymal transition via snail dephosphorylation and stabilization. *Open Biol.* 8, 170274. 10.1098/rsob.170274. [PubMed: 29618518]
- Zhao Y, Xiao M, Sun B, Zhang Z, Shen T, Duan X, Yu PB, Feng XH, and 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. 10.1074/jbc.M114.568964. [PubMed: 25100727]
- Zhong X, Prinz A, Steger J, Garcia-Cuellar MP, Radsak M, Bentaher A, and 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. 10.1182/bloodadvances.2018025866. [PubMed: 30463913]
- Zuber J, McJunkin K, Fellmann C, Dow LE, Taylor MJ, Hannon GJ, and Lowe SW (2011). Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nat. Biotechnol* 29, 79–83. 10.1038/nbt.1720. [PubMed: 21131983]

**Highlights**

- Phosphatase-domain-focused CRISPR screening applied to cancer cell lines
- SCP4 phosphatase is an acquired catalytic dependency in AML
- STK35/PDIK1L kinases bind to the catalytically active SCP4
- SCP4 removes inhibitory phosphorylation from the STK35/PDIK1L activation loop



**Figure 1. Phosphatase-domain-focused CRISPR screening identifies context-specific dependencies in human cancer cell lines**

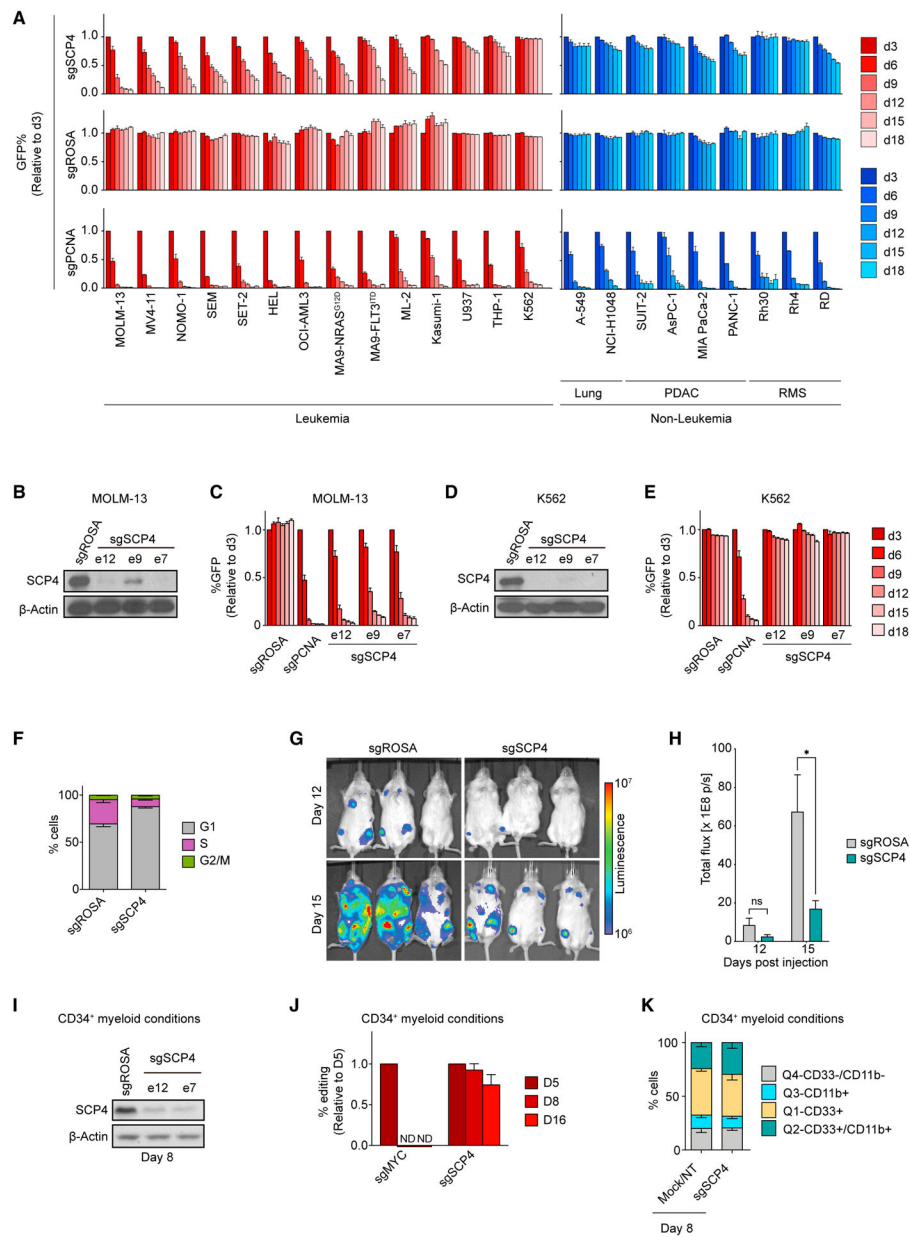
(A) Categorization of human phosphatase domains targeted by CRISPR-Cas9 genetic screens in this study.

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

(C) Extracted essentiality scores for phosphatases demonstrating AML-bias or pan-essentiality. Plotted is the  $\log_2$  fold change of sgRNA abundance during ~11 population doublings. The effects of individual sgRNAs targeting each domain were averaged. AML, acute myeloid leukemia; CML, chronic myeloid leukemia; SCLC, small-cell lung cancer; PDAC, pancreatic ductal adenocarcinoma; RMS, rhabdomyosarcoma.

(D) Summary of dependencies discovered in the screen.

See also Figure S1 and Table S1.



**Figure 2. SCP4 is an acquired dependency in human AML**

(A) Competition-based proliferation assays in 23 human cancer cell lines infected with the indicated sgRNAs. PDAC, pancreatic ductal adenocarcinoma; RMS, rhabdomyosarcoma.  $n = 3$ .

(B) Western blot of whole-cell lysates from MOLM-13 cells on day 5 post-infection with the indicated sgRNAs.

(C) Competition-based proliferation assay in MOLM-13 cells infected with the indicated sgRNAs.  $n = 3$ .

(D) Western blot of whole-cell lysates from K562 cells on day 5 post-infection with the indicated sgRNAs.

(E) Competition-based proliferation assay in K562 cells infected with the indicated sgRNAs. n = 3.

(F) Quantification of the different cell-cycle stages in MOLM-13 cells on day 5 post-infection with the indicated sgRNAs. n = 3.

(G) Bioluminescence imaging of NSG mice transplanted with luciferase<sup>+</sup>/Cas9<sup>+</sup> MOLM-13 cells infected with either sgROSA or sgSCP4.

(H) Quantification of bioluminescence intensity. n = 3. p value was calculated by unpaired Student's t-test. \*p < 0.05.

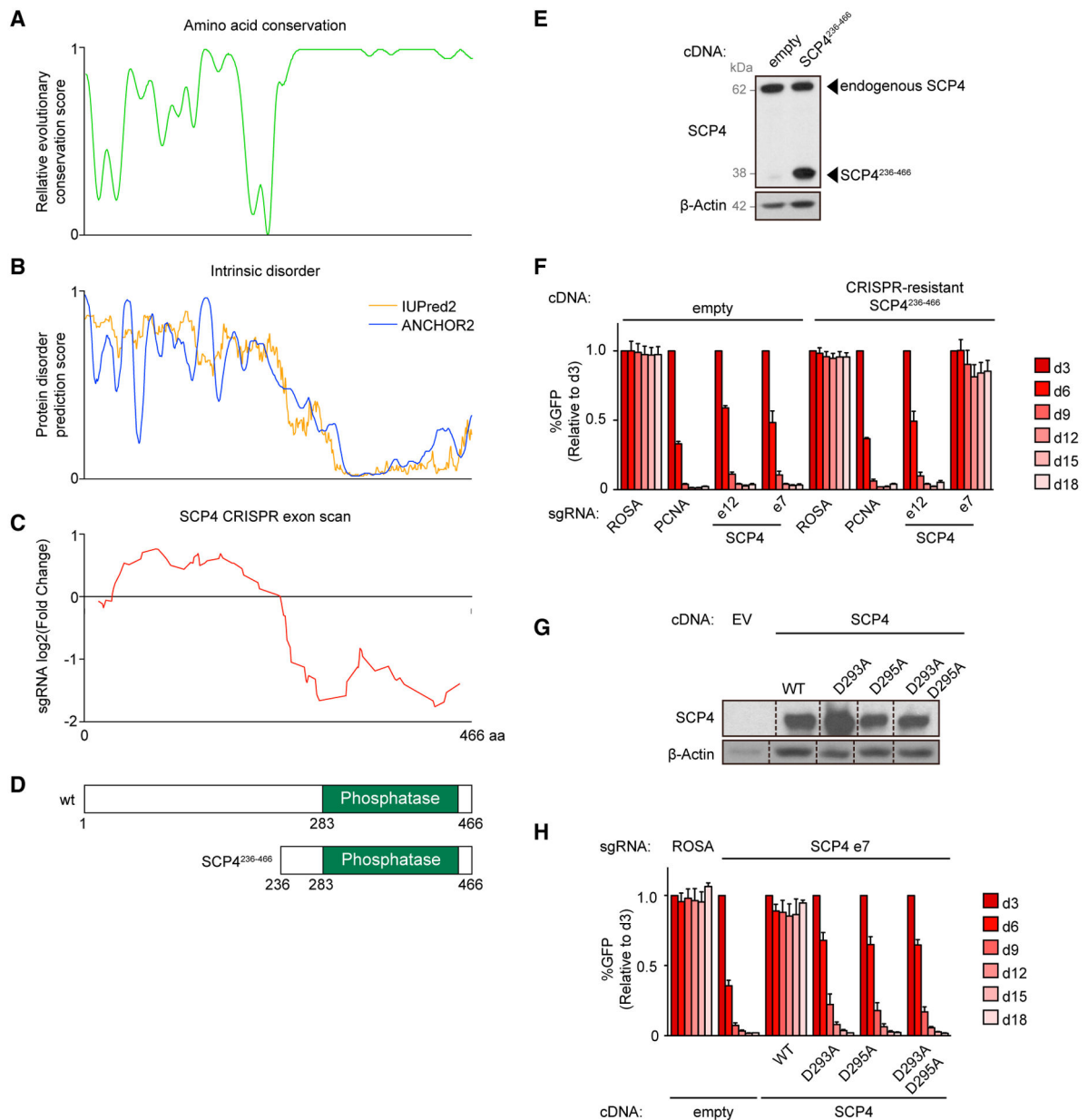
(I) Western blot of whole-cell lysates from CD34<sup>+</sup> cells electroporated with Cas9 loaded with the indicated sgRNAs on day 8 post-electroporation, with culturing in myeloid conditions.

(J) qRT-PCR analysis of indels presence in CD34<sup>+</sup> cells electroporated with Cas9 loaded with the indicated sgRNAs over the course of culturing in myeloid conditions. The effects of individual sgRNAs for SCP4 were averaged. n = 4.

(K) Quantification of the flow cytometry analysis of myeloid differentiation of CD34<sup>+</sup> cells electroporated with Cas9 loaded with the indicated sgRNAs on day 8 post-electroporation, with culturing in myeloid conditions. The effects of individual negative controls and sgRNAs for SCP4 were averaged. n = 4.

All bar graphs represent the mean ± 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, ROSA, Mock, and NT, negative controls; PCNA and MYC, positive controls.

See also Figure S2.



**Figure 3. The catalytic phosphatase function of SCP4 is essential in AML**

(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. Based on data from Meszaros et al. (2018).

(C) Running average of  $\log_2$  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) Domain architectures of human SCP4 and the truncated version of SCP4 used in this study.

(E) Western blot of whole-cell lysates from MOLM-13 cells stably expressing empty vector or CRISPR-resistant SCP4<sup>236-466</sup>.

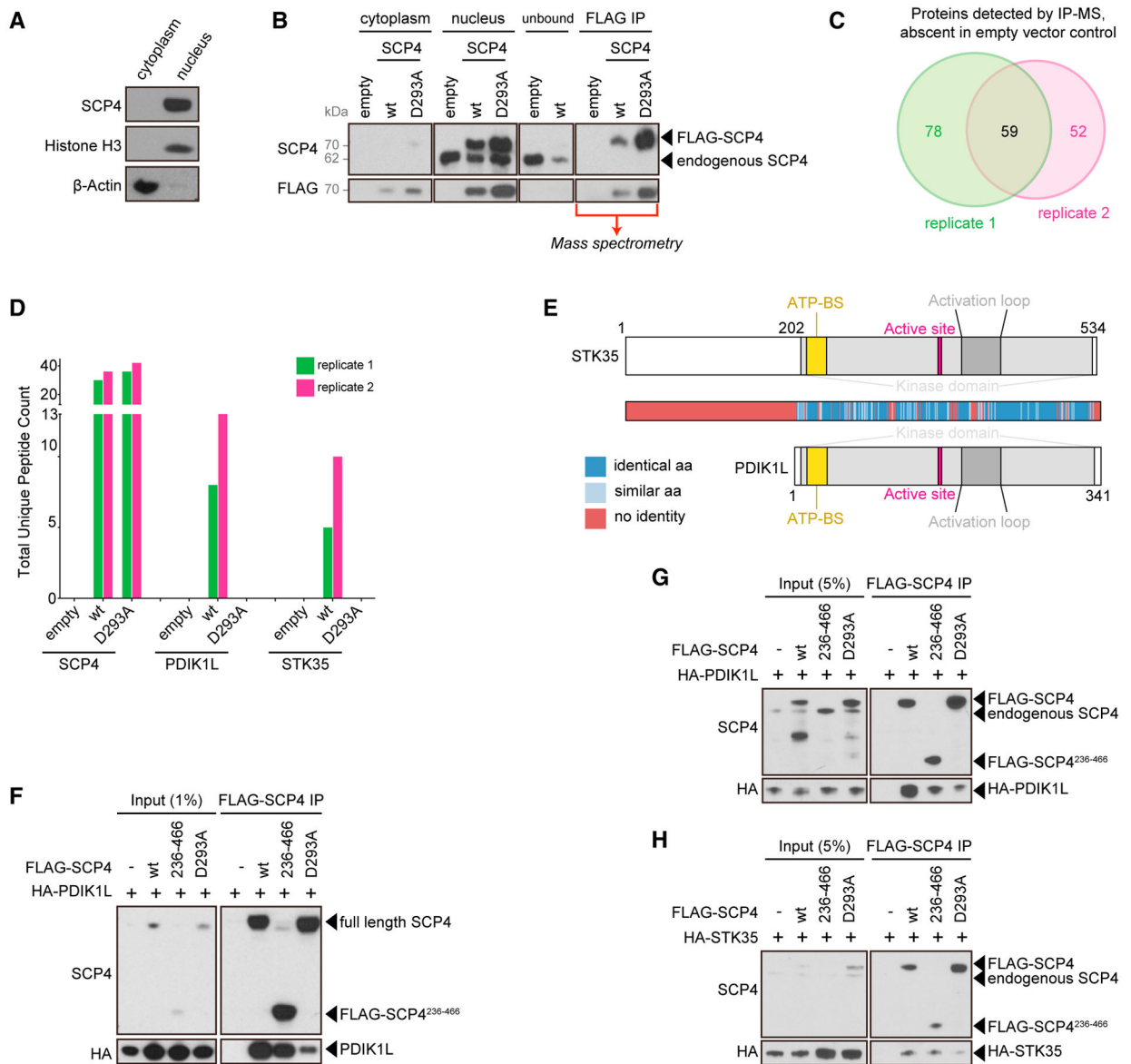


(F) Competition-based proliferation assay in MOLM-13 cells stably expressing empty vector or CRISPR-resistant SCP4<sup>236-466</sup> infected with the indicated sgRNAs. n = 3.

(G) Western blot of whole-cell lysates from MOLM-13 cells stably expressing empty vector (EV; unloaded) or CRISPR-resistant wild type (WT) and catalytic mutants of SCP4. Vertical black dashed lines indicate omitted lanes from the same gel and same exposure.

(H) Competition-based proliferation assay in MOLM-13 cells stably expressing empty vector or CRISPR-resistant WT and catalytic mutants of SCP4 infected with the indicated sgRNAs. n = 3.

All bar graphs represent the mean  $\pm$  SEM. ROSA, negative control; PCNA, positive control. See also Figure S3 and Table S2.



**Figure 4. STK35 and PDIK1L bind selectively to the catalytically active form of SCP4**

(A) Western blot of cytoplasm and nuclear fractions from MOLM-13 cells.

(B) Representative FLAG-SCP4 affinity purification western blot analysis for the subsequent mass spectrometry (MS) analysis. Cytoplasm and nucleus fractions from MOLM-13 cells stably expressing empty vector, FLAG-SCP4 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 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.

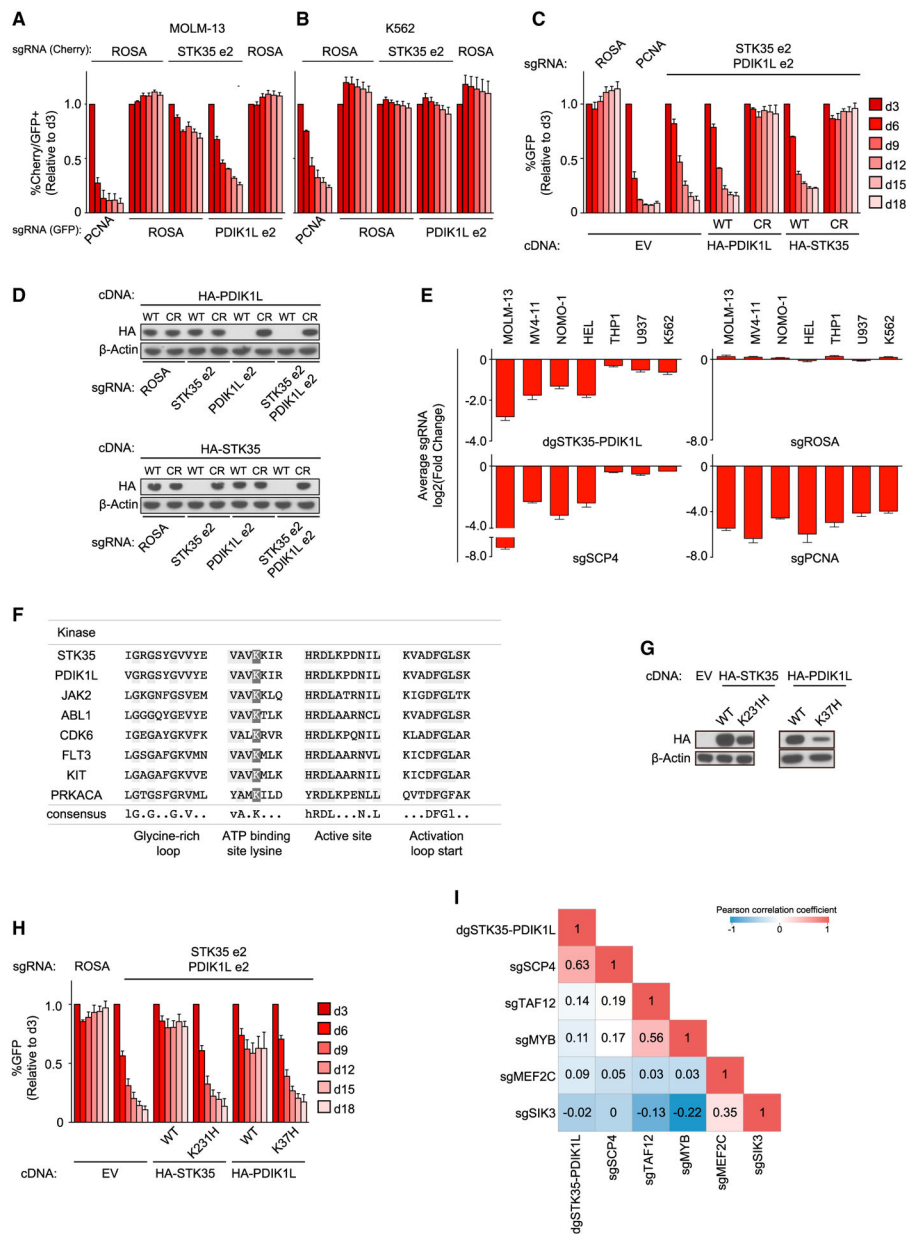
(C) Venn diagram depicting the overlap between proteins detected by MS and absent in empty vector controls in the 2 independent biological replicates.

(D) Total unique peptide counts for SCP4, PDIK1L, and STK35 detected by MS in the 2 independent biological replicates.

(E) Domain architectures and homology heat-map of human STK35 and PDIK1L. ATP-BS, ATP-binding site.

(F–H) Immunoprecipitation followed by western blotting performed with the indicated antibodies. The whole-cell lysate was prepared from HEK293T 24 h post-transfection with the indicated constructs (F). The nuclear lysates were prepared from the human MOLM-13 cells stably expressing the indicated constructs (G and H). –, empty vector; WT, wild-type FLAG-SCP4; 236–466, FLAG-SCP4<sup>236–466</sup>; D293A, catalytic mutant FLAG-SCP4<sup>D293A</sup>; IP, immunoprecipitation. Note: degradation bands appear in the WT and D293A input at ~50 kDa and in D293A at ~40 kDa.

See also Figure S4.



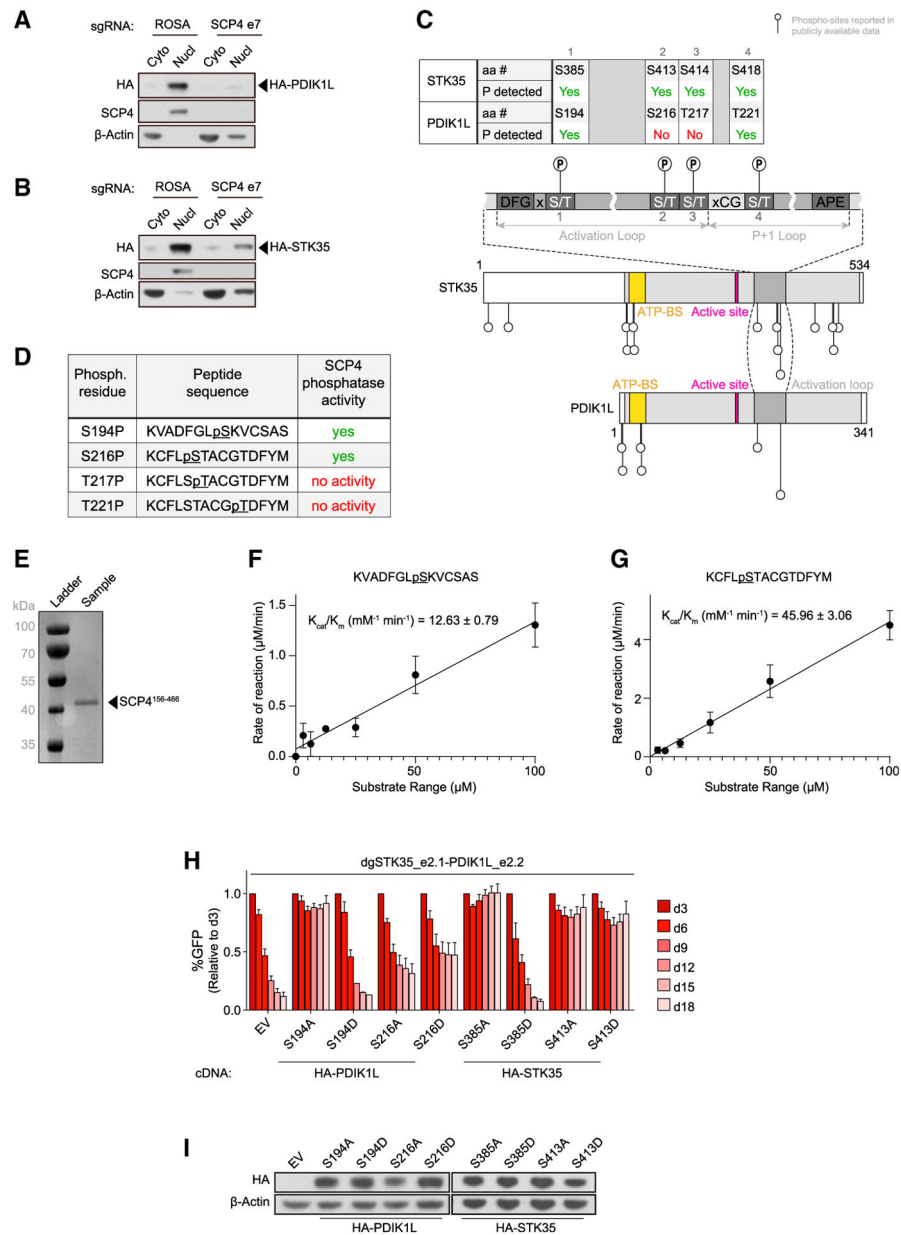
**Figure 5. STK35 and PDIK1L function redundantly in the same genetic pathway as SCP4**  
(A and B) Competition-based proliferation assays in MOLM-13 (A) and K562 (B) cells co-infected with GFP-linked sgRNA and mCherry-linked sgRNA. Double mCherry<sup>+</sup>/GFP<sup>+</sup> population depletion indicates loss of cell fitness due to genetic redundancy. n = 3.  
(C) Competition-based proliferation assay in MOLM-13 cells stably expressing EV, WT, or CRISPR-resistant (CR) HA-PDIK1L and HA-STK35 infected with the indicated sgRNAs. n = 3.  
(D) Western blot of whole-cell lysates from MOLM-13 cells stably expressing WT and CR HA-PDIK1L and HA-STK35 on day 5 post-infection with the indicated sgRNAs.  
(E) Summary of competition-based proliferation assays in the indicated cell lines. Plotted is the fold change (log<sub>2</sub>) of sgRNA<sup>+</sup>/GFP<sup>+</sup> cells after 18 days in culture (average of triplicates).

(F) Alignment of residues defining ATP binding site for human protein kinases. Residues that are conserved among the kinase 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.

(G) Western blot of whole-cell lysates from MOLM-13 cells stably expressing EV or CR WT and catalytic mutants of STK35 and PDIK1L in MOLM-13.

(H) Competition-based proliferation assay in MOLM-13 cells stably expressing EV or CR WT and catalytic mutant versions of HA-STK35 and HA-PDIK1L infected with the indicated sgRNAs.  $n = 3$ .

(I) Correlation matrix for the global  $\log_2$  fold changes of gene expression relative to negative controls in each independent experiment across 15,095 genes. All bar graphs represent the mean  $\pm$  SEM. “dg” refers to the bi-cistronic vector. ROSA, negative control; PCNA, positive control. See also Figure S5.



### Figure 6. SCP4 functions upstream of STK35/PDK11L to support AML proliferation

(A and B) Western blot of cytoplasm (Cyto) and nucleus (Nucl) fractions of MOLM-13 cells stably expressing HA-PDK11L (A) or HA-STK35 (B) on day 5 post-infection with the indicated sgRNAs. Shown is a representative experiment of 3 independent biological replicates.

(C) Schematics of phosphorylation sites reported in the publicly available phosphoproteomics datasets on STK35 and PDK11L relative to their domain architectures (Hornbeck et al., 2015; Ochoa et al., 2020). aa #, amino acid number; P, phosphorylation; ATP-BS, ATP-binding site.

(D) Phosphorylated peptides assayed for SCP4 phosphatase activity *in vitro*.

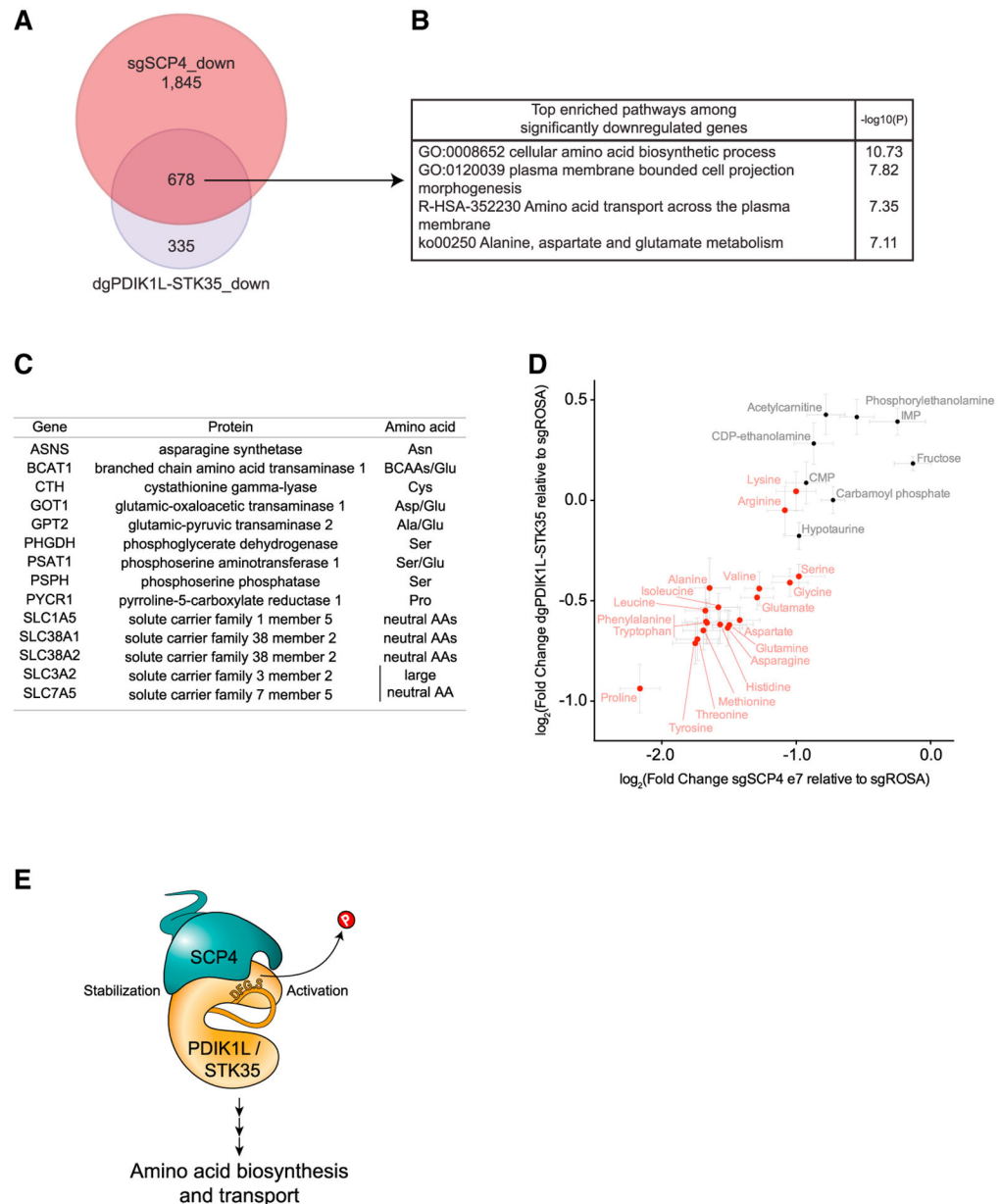
(E) Recombinant SCP4 protein purity as assessed by SDS-PAGE and Coomassie blue staining. His-SUMO-(TEV)-SCP4 was expressed in BL21 (DE3) cells and purified by affinity, anion exchange, and gel filtration chromatography. Molecular weight markers are shown for reference.

(F and G) Phosphatase activity of SCP4 against the indicated peptides plotted for kinetic fitting. Each measurement was conducted in triplicate with standard deviations shown as error bars.

(H) Competition-based proliferation assay in MOLM-13 cells stably expressing EV or the CR HA-PDIK1L or HA-STK35 constructs harboring the indicated amino acid substitutions infected with the indicated sgRNAs.  $n = 3$ .

(I) Western blot of whole-cell lysates from MOLM-13 cells stably expressing EV or CR HA-PDIK1L or HA-STK35 constructs harboring the indicated amino acid substitutions.

All bar graphs represent the mean  $\pm$  SEM. See also Figure S6.



**Figure 7. SCP4-STK35/PDIK1L complex is needed to sustain amino acid levels in AML**  
 (A) Venn diagram depicting the overlap between statistically significant downregulated genes in MOLM-13 cells upon SCP4 knockout and STK35/PDIK1L double knockout. DeSeq2 (n = 4).  
 (B) Ontology analysis of overlapping statistically significant 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 of.  
 (D) 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 ( $n = 6$ ). Amino acids are in red, and there are a few unchanged metabolites in black for reference. Shown is a representative experiment of 3 independent biological replicates.

(E) Model.

See also Figure S7.

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti- $\beta$ -Actin, HRP-linked	Sigma-Aldrich	Cat# A3854-200UL
Mouse monoclonal anti-FLAG	Sigma-Aldrich	Cat# F1804
Rabbit polyclonal anti-H3K4me3	Sigma-Aldrich	Cat# 07-473
Rat monoclonal anti-HA (clone 3F10), HRP-linked	Roche	Cat# 12013819001
Rabbit monoclonal anti-SCP4	Cell Signaling Technology	Cat# 6932
Rabbit polyclonal anti-mouse IgG, HRP-linked	Agilent/Dako	Cat# P026002-2
Donkey anti-rabbit IgG, HRP-linked	Cytiva/Amersham	Cat# NA934
Bacterial and virus strains		
MegaX DH10B T1R Electrocomp Cells	Invitrogen	Cat# C640003
Chemicals, peptides, and recombinant proteins		
2-Mercaptoethanol	Sigma-Aldrich	Cat# M6250
2x Laemmli Sample Buffer	BIO-RAD	Cat# 1610737
Agarose, Standard, Low Electroendosmosis (EEO)	Avantor	Cat# A426-07
Anti-FLAG <sup>®</sup> M2 Affinity Gel	Sigma-Aldrich	Cat# A2220-10ML
Blasticidin S HCl	Thermo Fisher Scientific	Cat# A1113903
Blasticidin S HCl	Thermo Fisher Scientific	Cat# A1113903
BsmBI	New England BioLabs	Cat# R0739
Cell Lysis Buffer	Cell Signaling Technology	Cat# 9803
Chloroform	Sigma-Aldrich	Cat# 288306
D-Luciferin	Goldbio	Cat# 115144-35-9
Dimethyl Sulfoxide	Sigma-Aldrich	Cat# D2650
DMEM with 4.5 g/L glucose, L-glutamine & sodium pyruvate	Cellgro	Cat# 10-013-CV
DNA Polymerase I, Large (Klenow) Fragment	New England Biolabs	Cat# M0210L
EZview <sup>™</sup> Red Anti-HA Affinity Gel	Sigma-Aldrich	Cat# E6779-1ML
Geneticin Selective Antibiotic (G-418 Sulfate)	Thermo Fisher Scientific	Cat# 10131035
Geneticin Selective Antibiotic (G418 Sulfate)	Thermo Fisher Scientific	Cat# 10131035
HEPES	Life Technologies	Cat# 15630-130
ImmPACT DAB peroxidase (HRP) substrate	Vector	Cat# SK-4105
Isopropanol	Sigma-Aldrich	Cat# 190674
Klenow Fragment (3'-5' exo-)	New England Biolabs	Cat# M0212L
Opti-MEM <sup>®</sup>	Thermo Fisher Scientific	Cat# 31985062
Penicillin/Streptomycin	Thermo Fisher Scientific	Cat# 15140122
Phusion Flash High-Fidelity PCR Master Mix	Thermo Fisher Scientific	Cat# F548S
Phusion <sup>®</sup> Hot Start Flex DNA Polymerase	New England Biolabs	Cat# M0535
Pierce <sup>™</sup> Anti-HA Magnetic Beads	Thermo Fisher Scientific	Cat# 88837
Polybrene	EMD Millipore	Cat# TR-1003-G
Polyethylenimine, Linear, MW 25,000 (PEI 25000)	Polysciences	Cat# 23966-1
Puromycin dihydrochloride	Sigma-Aldrich	Cat# P8833

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RPMI 1640 w/ L-glutamine	Fisher Scientific	Cat# MT10040CV
T4 DNA polymerase	New England Biolabs	Cat# M0203L
T4 polynucleotide kinase	New England Biolabs	Cat# M0201L
TRIzol Reagent	Thermo Fisher Scientific	Cat# 15596018
Critical commercial assays		
Agilent High Sensitivity DNA Kit	Agilent Genomics	Cat# 5067-4626
FITC Annexin V Apoptosis Detection Kit	BD Biosciences	Cat# 556547
FITC BrdU Flow Kit	BD Biosciences	Cat# 559619
Gibson Assembly Master Mix	New England Biolabs	Cat# E2611
In-Fusion HD Cloning Kit	Takara Bio	Cat# 638909
MiniElute PCR Purification Kit	QIAGEN	Cat# 28004
MiSeq Reagent Kit V3 150-cycle	Illumina	Cat# MS-102-3001
Pierce™ BCA Protein Assay Kit	Thermo Fisher Scientific	Cat# 23225
Power SYBR Green Master Mix	Thermo Fisher Scientific	Cat# 4368577
QIAamp DNA Mini Kit	QIAGEN	Cat# 51304
QIAquick Gel Extraction Kit	QIAGEN	Cat# 28704
Quick Ligation Kit	New England Biolabs	Cat# M2200L
TruSeq RNA Sample Prep Kit v2	Illumina	Cat# RS-122-2001
Deposited data		
RNA-Seq	This study	GEO: GSE174169
RNA-Seq	Tarumoto et al. (2018)	GEO: GSE109491
RNA-Seq	Xu et al. (2018)	GEO: GSE104308
Human reference genome GRCh37/hg19	Genome Reference Consortium	<a href="https://www.ncbi.nlm.nih.gov/grc/human">https://www.ncbi.nlm.nih.gov/grc/human</a>
Human reference genome GRCh37/hg38	Genome Reference Consortium	<a href="https://www.ncbi.nlm.nih.gov/grc/human">https://www.ncbi.nlm.nih.gov/grc/human</a>
Raw Mass Spectrometry Data Files	This study	ProteomeXchange Consortium with the dataset identifier PXD029769. Additional identifiers <a href="https://doi.org/10.25345/C5786Q">https://doi.org/10.25345/C5786Q</a> , MassIVE (MSV000088383)
Raw Metabolomics Data Files	This study	MetaboLights database with identifier MTBLS3707
Experimental models: Cell lines		
Human: A549	ATCC	CCL-185
Human: AsPC-1	ATCC	CRL-1682
Human: HEK293T	ATCC	CRL-3216
Human: HEL	ATCC	TIB-180
Human: K562	ATCC	CCL-243
Human: Kasumi-1/Cas9	Robert Roeder	N/A
Human: MA9-ITD	Wunderlich et al. (2013)	N/A
Human: MA9-RAS	Wunderlich et al. (2013)	N/A
Human: MIAPaCa-2	ATCC	CRL-142
Human: ML-2	DSMZ	ACC 15
Human: MOLM-13	DSMZ	ACC 554
Human: MV4-11	ATCC	CRL-9591

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human: NCI-H1048 [H1048]	ATCC	CRL-5853
Human: NOMO-1	DSMZ	ACC 542
Human: OCI-AML3	DSMZ	ACC 582
Human: PANC-1	ATCC	CRL-1469
Human: RD	ATCC	CCL-136
Human: RH30	DSMZ	ACC 489
Human: RH4	NA	RRID: CVCL_5916
Human: SEM	DSMZ	ACC 546.
Human: SET-2	DSMZ	ACC 608
Human: SUIT-2	JCRB	JCRB1094
Human: THP-1	ATCC	TIB-202
Human: U937	ATCC	CRL-1593.2
Mouse: NIH3T3	ATCC	CRL-1658
Mouse: RN2	Zuber et al. (2011)	N/A
Experimental models: Organisms/strains		
Mouse: NSG	The Jackson Laboratory	Stock No: 005557
Oligonucleotides		
sgRNA sequences, see Tables S3 and S4	This study	N/A
Primers, see Tables S3 and S4	This study	N/A
Recombinant DNA		
LRG (Lenti_sgRNA_EFS_GFP)	Shi et al. (2015)	Addgene plasmid # 65656
LRG2.1	Tarumoto et al. (2018)	Addgene plasmid # 108098
LRCherry2.1	Tarumoto et al. (2018)	N/A
LRG2.1-Blast	Klingbeil, unpublished	N/A
LentiV_Cas9_puro	Tarumoto et al. (2018)	Addgene plasmid # 108100
LentiV_Neo	Tarumoto et al. (2018)	N/A
Lenti-luciferase-P2A-Neo	Wei et al. (2021)	Addgene plasmid # 105621
psPAX2	N/A	Addgene plasmid # 12260
<i>CTDSPL2</i> cDNA	Horizon Discovery	Clone ID: 5744745
partial <i>STK35</i> cDNA	Horizon Discovery	Clone ID: 9021751
<i>PDIK1L</i> cDNA	Horizon Discovery	Clone ID: 4828997
Software and algorithms		
Aminode	Chang et al. (2018)	<a href="http://www.aminode.org/">http://www.aminode.org/</a>
CSHL sgRNA Design Tool	El Demerdash, unpublished	<a href="https://mojica.cshl.edu">https://mojica.cshl.edu</a>
DESeq2	Love et al. (2014)	<a href="https://bioconductor.org/packages/release/bioc/html/DESeq2.html">https://bioconductor.org/packages/release/bioc/html/DESeq2.html</a>
FlowJo	TreeStar Inc.	<a href="https://www.flowjo.com">https://www.flowjo.com</a>
GraphPad Prism v.5 and v.9	GraphPad Software	<a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a>
HTSeq-Count v.0.6.1p1	Anders et al. (2015)	<a href="https://htseq.readthedocs.io/en/master/">https://htseq.readthedocs.io/en/master/</a>
HUGO Gene Nomenclature Committee (HGNC)	N/A	<a href="https://www.genenames.org">https://www.genenames.org</a>
IUPred2A	Meszaros et al. (2018)	<a href="https://iupred2a.elte.hu">https://iupred2a.elte.hu</a>

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Kallisto	Bray et al. (2016)	<a href="https://pachterlab.github.io/kallisto/">https://pachterlab.github.io/kallisto/</a>
Metascape	Tripathi et al. (2015)	<a href="http://metascape.org">http://metascape.org</a>
Salmon v.1.0	Patro et al. (2017)	<a href="https://salmon-tddft.jp/wiki/Download-v.1.0.0">https://salmon-tddft.jp/wiki/Download-v.1.0.0</a>
STAR v.2.5.2b-1	Dobin et al. (2013)	<a href="https://github.com/alexdobin/STAR">https://github.com/alexdobin/STAR</a>
UCSC Genome Browser	UCSC	<a href="http://genome.ucsc.edu">http://genome.ucsc.edu</a>

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript