Genome-wide CRISPR-Cas9 Screen Identifies Leukemia-Specific Dependence on a Pre-mRNA Metabolic Pathway Regulated by DCPS

Yamauchi, Takuji
Division of Hematology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School

Masuda, Takeshi
Department of Pharmaceutical Microbiology, Faculty of Life Sciences, Kumamoto University

Canver, Matthew C.
Division of Hematology/Oncology, Boston Children's Hospital

Seiler, Michael
H3 Biomedicine, Inc.

他

出版情報：Cancer Cell. 33 (3), pp.386-400, 2018-03-12. Elsevier
バージョン：

Genome-wide CRISPR-Cas9 screen identifies leukemia-specific dependence on a pre-mRNA metabolic pathway regulated by the DCPS decapping enzyme

Takuji Yamauchi[1,2,3], Takeshi Masuda[4], Matthew C. Canver[5], Michael Seiler[6], Yuichiro Semba[2], Mohammad Shboul[7], Mohammed Al-Raqad[7,8], Manami Maeda[1], Vivien A. C. Schoonenberg[5], Mitchel A. Cole[5], Claudio Macias Trevino[5], Yuichi Ishikawa[1], Qiuming Yao[9], Michitaka Nakano[2], Fumio Arai[3], Stuart H. Orkin[5], Bruno Reversade[7], Silvia Buonamici[6], Luca Pinello[9], Koichi Akashi[2,10], Daniel E. Bauer[5] and Takahiro Maeda[1,10]*

Affiliations: [1]Division of Hematology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Department of Medicine and Biosystemic Science, Kyushu University Graduate School of Medical Sciences, Fukuoka 812-8582, Japan. [3]Department of Stem Cell Biology and Medicine, Graduate School of Medical Sciences, Kyushu University. [4]Department of Pharmaceutical Microbiology, Faculty of Life Sciences, Kumamoto University, Kumamoto 862-0973, Japan. [5]Division of Hematology/Oncology, Boston Children's Hospital, Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Stem Cell Institute, Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA. [6]H3 Biomedicine, Inc., Cambridge, MA 02139, USA. [7]Institute of Medical Biology, A*STAR, 8A Biomedical Grove, Singapore 138648, Singapore. [8]Al-Balqa Applied University, Faculty of Science, Al-Salt, 19117, Jordan. [9]Department of Pathology & Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA.[10]Center for Cellular and Molecular Medicine, Kyushu University Hospital, Fukuoka 812-0054, Japan.

***Corresponding Author and Lead Contact:** t_maeda@cancer.med.kyushu-u.ac.jp

SUMMARY

To identify novel targets for acute myeloid leukemia (AML) therapy, we performed genome-wide CRISPR-Cas9 screening using AML cell lines, followed by a second screen in vivo. Here, we show that the mRNA decapping enzyme scavenger (*DCPS*) gene is essential for AML cell survival. The DCPS enzyme interacted with components of pre-mRNA metabolic pathways, including spliceosomes, as revealed by mass spectrometry. RG3039, a DCPS inhibitor originally developed to treat spinal muscular atrophy, exhibited anti-leukemic activity via inducing pre-mRNA mis-splicing. Humans harboring germline bi-allelic *DCPS* loss-of-function mutations do not exhibit aberrant hematologic phenotypes, indicating that *DCPS* is dispensable for human hematopoiesis. Our findings shed light on a pre-mRNA metabolic pathway and identify DCPS as a target for AML therapy.

SIGNIFICANCE

Genome-wide CRISPR-Cas9 knockout screening is a powerful tool for functional genomics. We performed genome-wide CRISPR-Cas9 dropout screens utilizing AML lines exhibiting a normal karyotype and harboring functionally-normal *Trp53*, which enabled straightforward interpretation of screening results. We then selected potentially actionable targets using databases relevant to chemical inhibitors and human genetics. We identified DCPS as a drug target for AML therapy. We report that RG3039, a DCPS inhibitor proven safe in a phase I clinical trial, exhibits anti-leukemia activity and identify a pre-mRNA metabolic pathway required for AML survival. Our work suggests a potential opportunity to ‘‘re-purpose’ DCPS inhibitors for AML therapy.

INTRODUCTION

Acute myeloid leukemia (AML) is a devastating disease with a long-term survival rate of less than 30% (Ferrara and Schiffer, 2013). Recent progress has been made to define its mechanisms, and sequencing studies now provide a near-complete picture of the AML genome (Welch et al., 2012). Nonetheless, to devise urgently needed therapies, functional studies are necessary to assess the significance of AML-associated mutations (Boehm and Hahn, 2011; Garraway and Lander, 2013; Lawrence et al., 2014).

Successful application of the *S. pyogenes*-derived type 2 clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 system for genome editing is transforming the landscape of genetic research in many organisms (Cho et al., 2013; Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013). Furthermore, given its high efficiency and flexibility, the system is ideal for use in genome-wide recessive genetic screens and proof-of-principle studies using cancer cell lines demonstrate the potential of this technology to identify genes essential for cancer cell survival (Koike-Yusa et al., 2014; Shalem et al., 2014; Shi et al., 2015; Wang et al., 2014). Since AML cells generally exhibit low mutational burdens (Alexandrov et al., 2013) and mutation of the *TP53* tumor suppressor gene has a significant impact on AML prognosis (Zhang et al., 2016), it is

critical to perform a screen employing AML lines whose genetic background, namely *TP53* status, are well-defined.

RESULTS

A genome-wide CRISPR-Cas9 screen identifies DCPS as an AML essential gene

To establish AML cell lines with a relatively "clean" genetic background, we first generated AML in mice by transducing either the *CALM/AF10* or *MLL/AF9* leukemia oncogene into mouse bone marrow hematopoietic stem cells (HSCs) and then transferred cells to sub-lethally irradiated recipients. Primary AML cells were harvested 3-6 months later, serially transplanted three times, and then cultured in the presence of cytokines. The resultant two independent lines were then transduced with the Cas9 nuclease (**Figure 1A**). Cells of both lines exhibited normal karyotypes (**Figure S1A**). To perform the genome-wide CRISPR-Cas9 screen, we used the mouse lentivirus-based GeCKO v2 library, which contains 130,209 single-guide RNAs (sgRNAs) targeting 20,611 protein-coding genes and 1,175 miRNAs (Sanjana et al., 2014; Shalem et al., 2014). Cas9-expressing AML cells of both lines were transduced with the library (day 0) and treated with puromycin on day 1. We passaged the cells 5-6 times over a 16-day incubation period, while maintaining at least 500 cells per sgRNA throughout (**Figure S1B**). Genomic DNA

was isolated from cells on days 3 and 18 and deep-sequenced to measure read counts of each sgRNA. Changes in abundance of each sgRNA were assessed using the MAGeCK program (Li et al., 2014; Shalem et al., 2014). We obtained over 400 million mapped reads per sample, suggesting that at least 600 cells were transduced with each sgRNA (**Figure S1C**). Strikingly, sgRNAs targeting *Trp53* were among the most enriched after a 16-day incubation of both AML lines, indicating intact TP53 activity in both lines (**Figures 1B** and **1C**). We identified nearly 1,700 dropout genes in each line at a false discovery rate (FDR) of 0.25, with significant overlap between the lines (**Figure 1D**, **Figure S1D** and **Table S1**). As expected, genes encoding components of basal cellular machineries were highly enriched in dropout genes (**Figure S1E**). Dropout genes were abundantly expressed in primary mouse CALM/AF10 or MLL/AF9 leukemia cells, an observation strongly suggesting those genes are functional in AML cells and that sgRNA off-target effects are negligible (**Figure S2A**). Overall, we identified 2256 dropout genes at a FDR of 0.25 using two AML lines (**Figure 1D**). sgRNAs targeting the genes with a well-defined function in leukemogenesis, among them, *Kras*, *Nras*, *Bcl2l1*, *Jak1*, *Jak2*, *Brd4* and *Brd9*, were significantly depleted, confirming quality and efficiency of our dropout screens (**Figure S2B**).

To identify potentially actionable targets, we selected genes meeting the following criteria: 1) they encoded a protein with an available inhibitor and/or 2) their germline mutation loss-of-function phenotype was relatively moderate based on the literature or the human exome-sequencing database (Lek et al., 2016; Lim et al., 2014; Narasimhan et al., 2016; Saleheen et al., 2017; Sulem et al., 2015). We excluded genes encoding components of basal cellular machineries (e.g. histones, ribosomal proteins, or polymerases), as targeting these factors would likely be deleterious to normal tissues (**Figure 1D**). Since AML lines used in initial screens were maintained in the presence of cytokines, it was possible that some dropout genes encoded essential effectors of cytokine signaling, but were dispensable for activity of primary AML cells. To address this issue, we performed a second in vivo validation screen targeting 470 genes (**Figure 1D** and **Table S2**). Cas9-expressing MLL/AF9 cells were transduced with a library containing 8 sgRNAs per gene plus 100 non-targeting sgRNAs and then transferred into sub-lethally irradiated recipients (**Figure S2C**). DNA samples from pre- and post-transplant were deep-sequenced with sgRNA abundance determined using MAGeCK software (**Figure S2D**). Nearly one fourth of the 470 genes tested in the second screen were dispensable in vivo and another one fourth exhibited only moderate effects (not shown). Non-targeting sgRNAs showed

variable effects (**Figure S2E**). Overall, we identified 130 genes necessary for AML cell survival both in vitro and in vivo for further evaluation (**Figure 1D**).

Among genes significantly depleted in our primary screen was the mRNA decapping enzyme scavenger (*Dcps*), which encodes a mRNA 5' cap binding enzyme implicated in mRNA decay (Milac et al., 2014) (**Figure 1E**). Read counts for each *Dcps*-targeted sgRNA significantly decreased in both AML lines over the 16-day incubation period (**Figure 1F**), which also demonstrated significant dropout in the second screen (**Figure 1G**). DCPS protein harbors an N-terminal domain that shares structural homology to yeast mRNA export factor Mex67 and a C-terminal domain that contains a histidine triad (HIT) sequence (His274, His276, and His278) in which His276 is critical for decapping activity (Gu et al., 2004; Han et al., 2005). To determine DCPS amino acid residues necessary for AML cell survival, we performed a negative selection CRISPR-Cas9 mutagenesis scan (Shi et al., 2015) of all *Dcps* coding exons using the Ctps (cytidine 5'-triphosphate synthase) gene as control (**Table S3**). sgRNAs targeting either the N- or C-terminal DCPS domains, namely those targeting aa 230-240 or the HIT sequence, respectively, were significantly decreased after the 16-day incubation, strongly suggesting a critical role for those residues in AML survival (**Figure 1H** and **1I**). In contrast, 238 sgRNAs targeting all Ctps coding exons were largely dispensable (**Figure S2F**).

The DCPS inhibitor RG3039 slows proliferation and induces AML cell differentiation

We next asked whether DCPS inactivation inhibits proliferation of human AML cell lines. To this end, we generated lentivirus-based, GFP-tagged shRNA targeting DCPS and evaluated the proportions of GFP-positive (shRNA+) and -negative (shRNA-) cells by FACS over time (**Figure 2A**). Efficient DCPS knockdown in the GFP-positive fraction was confirmed by Western blot (**Figure 2B**). Cells that expressed shRNA-DCPS, but not those transduced with scrambled shRNA, demonstrated a proliferative disadvantage as compared to non-transduced cells, indicating a toxic effect of shRNA-DCPS in AML cells (**Figure 2A**). DCPS knockdown induced cell cycle arrest (**Figure 2C**) and apoptosis (**Figure 2D**). Notably, DCPS-knockdown MOLM-13 cells underwent myelo/monocytic differentiation, as evidenced by FACS analysis and morphological examination (**Figure 2E** and **2F**).

The DCPS inhibitor RG3039, an orally-active quinazoline derivative, was originally developed to treat spinal muscular atrophy (SMA) (Jarecki et al., 2005). We examined RG3039 binding to DCPS protein in AML cells via a cellular thermal shift assay (CETSA) (Martinez Molina et al., 2013; Xu et al., 2016). To do so, we treated MOLM-13 cells with RG3039 or control DMSO vehicle and then raised resulting cell lysates to various temperatures in order to test whether DCPS protein is protected against

heat-induced denaturation in the presence of RG3039. Soluble fractions of protein lysates were collected and subjected to Western blot. While DCPS protein was not detected in lysates of DMSO control cells subjected to higher temperature, DCPS protein remained soluble in lysates of RG3039-treated cells dose-dependently and thus was detectable by Western blot (**Figures 2G** and **2H**). Time course CETSA analysis showed that engagement of ligand with DCPS occurred immediately after RG3039 administration and then gradually decreased over 24 hours (not shown).

We next assessed anti-leukemia effects of RG3039 by treating 4 human AML lines with RG3039 for 11 days and generating growth curves. That analysis showed that RG3039 had dose-dependent anti-proliferative effects, although compound sensitivity varied among lines (**Figure 2I**). We also assessed the effects of RG3039 on the cell cycle via a 5-ethynyl-2´-deoxyuridine (EdU) incorporation assay. While EdU was efficiently incorporated into DMSO-treated control cells, EdU incorporation was significantly decreased in RG3039-treated cells, and the proportion of cells in S phase was extremely low in treated versus untreated cells (**Figures 2J** and **2K**). RG3039-treated cells also underwent apoptosis after 72 hr of drug treatment, as revealed by Annexin V stain (**Figure 2L**). As observed in DCPS-knockdown AML cells, RG3039 treatment induced differentiation of human (**Figures S3A** and **S3B**) and mouse (**Figures S3C** and **S3D**) AML.

Importantly, observed anti-leukemic effects of RG3039 were caused via TP53-independent mechanisms, as CALM/AF10 and MLL/AF9 AML cells whose TP53 function was abrogated by CRISPR-Cas9-mediated *Trp53* knockout were also sensitive to the compound (**Figure S3E**). Preferential DCPS dependency of leukemia cells was also validated using the DEMETER, an siRNA-based gene dependency database (Tsherniak et al., 2017). Strikingly, leukemia cells (n=28) were significantly more dependent on DCPS than were cells from non-hematological tumors (n=454) (**Figure S3F**). Of note, the leukemia cell line least dependent on DCPS was THP-1, which is relatively insensitive to shRNA-mediated DCPS knockdown or RG3039 treatment (**Figures 2A** and **2I**).

DCPS binds to the nuclear RNA processing machinery

Since DCPS protein was predominantly nuclear in human primary AML cells (**Figure S4A**), we hypothesized that DCPS primarily functions in that compartment rather than in the cytoplasmic mRNA decay pathway. To search for nuclear proteins potentially interacting with DCPS, we undertook immunoprecipitation (IP) with an anti-DCPS antibody of lysates of MOLM-13 cells followed by mass spectrometry analysis. Analysis was done in triplicate using non-specific immunoglobulins as control. Western blot analysis of IP'd proteins indicated successful pull-down of endogenous DCPS protein in each

experiment (**Figure 3A**). Among highly significant interactors, we identified components of pre-mRNA processing machineries including spliceosomes, the transcription-export complex (TREX), and the nuclear pore complex (NUP), as well as complexes functioning in pre-rRNA processing (**Figures 3B**, **3C** and **Table S4**). DCPS also bound to NuRD (Nucleosome Remodeling Deacetylase) subunits, suggesting a potential function in transcriptional regulation. Importantly, sgRNAs targeting genes encoding DCPS-interacting proteins were generally depleted in our CRISPR-Cas9 screens, indicating that binding events are functionally significant (**Figures 3D**). The cap-binding complex (CBC), which consists of CBP20 (also known as NCBP2) and CBP80 (NCBP1), reportedly binds the 5' cap of an mRNA in the nucleus and functions in RNA export and pre-mRNA splicing (Izaurralde et al., 1995; Izaurralde et al., 1994; Visa et al., 1996). We therefore used publicly available datasets to ask whether DCPS interacts with proteins that also participate in the CBC complex (Andersen et al., 2013). We found that some DCPS-interacting proteins are common to the CBC complex, including proteins also found in the TREX and NUP complex (**Figure S4B**).

DCPS inactivation causes pre-mRNA mis-splicing and induces a type I interferon response in AML cells

Given that DCPS binds the splicing machinery, we asked whether DCPS inhibition would impair pre-mRNA splicing in AML cells. To assess this possibility, we treated AML cells or granulocyte macrophage progenitors (GMPs), the normal counterpart of AML cells, with RG3039 and performed RNA-seq to determine potential transcriptome-wide splicing changes (**Figure 4A**). Of note, CETSA analysis showed comparable binding affinity between RG3039 and DCPS protein in cells, regardless of cell type (**Figure S4C**), suggesting that RG3039 inhibits DCPS enzymatic activity comparably in GMPs and AML cells. RNA-seq data were analyzed using a bioinformatic pipeline enabling quantification of exon-exon junctions without predetermining alternative splicing models or annotating splice junctions as described (Darman et al., 2015). We observed 493 and 704 mis-splicing events in AML cells at 6 hr and 10 hr after RG3039 treatment, respectively, while there were significantly fewer in GMPs (294 and 416 events at 6 hr and 10 hr, respectively) (**Figure 4B** and **Table S5**). Alternative 5' splice site (ss) selection was most frequently observed in both AML cells and GMPs (**Figure 4B**), although mis-spliced genes differed between the two cell types (**Figure 4C**). Mis-spliced mRNAs containing premature termination codons undergo nonsense-mediated mRNA decay (NMD), preventing production of aberrant proteins (Lykke-Andersen and Jensen, 2015). Bioinformatic prediction analysis revealed that approximately 40% of genes aberrantly-spliced at either

alternative 3' or 5' splice sites were NMD-sensitive in AML cells, but this outcome was not evident in GMPs (**Figure 4D**). Transcripts predicted to undergo NMD in AML cells were, in fact, significantly down-regulated after 10 hours of RG3039 treatment (**Figure 4D**). Notably, most aberrant splicing events observed in AML cells occurred at the first exon (or the first intron, in cases of intron retention) (**Figure 4E**).

Since DCPS associates with several pre-mRNA processing machineries, its inhibition could significantly impact pre-mRNA metabolism in ways not limited to pre-mRNA splicing. In fact, a genome-wide CRISPR screen, in which RG3039 was added to the medium of Cas9-expressing MLL/AF9 cells from days 10 to 18, failed to identify sgRNAs uniquely enriched in the presence of RG3039 (Table S1), suggesting that RG3039 exerts anti-leukemic activity via multiple mechanisms rather than a single pathway. To assess overall effects of DCPS inhibition on the leukemia transcriptome, we performed Gene Set Enrichment Analysis (GSEA) and functional annotation clustering using the Metascape program (http://metascape.org/). Both analyses identified gene signatures representing a type-I interferon response in RG3039-treated AML cells but not in GMPs (**Figures 4F, S4D** and **Table S5**). This activity may account for anti-leukemic effects observed following DCPS inhibition. Of note, genes aberrantly-spliced in AML cells were not necessarily enriched for factors functioning in IFN signaling (**Figure S4E**).

DCPS is dispensable for steady-state hematopoiesis in mice and humans

To determine effects of RG3039 on normal mouse hematopoiesis in vivo, we treated wild-type C57BL/6J mice with RG3039 daily for 12 days and examined peripheral blood (PB) counts over a time course (**Figure S5A**). RG3039/DCPS engagement was confirmed via CETSA in bone marrow (BM) cells the day after the last RG3039 injection (**Figure S5A**). While RG3039 treatment was tolerable and did not cause gross side effects, it induced a mild leukocytopenia at higher dosage (**Figure S5B**). BM cellularity was also slightly reduced with no cell-type specificity in RG3039-treated mice (**Figures S5C** and **S5D**). When mice were treated with RG3039 for a longer period (28 days), we observed mild transient anemia, which was normalized within a week of drug withdrawal (**Figure S5E**). Nonetheless, these data indicate that mice largely tolerate pharmacological inhibition of DCPS enzymatic activity.

To evaluate effects of DCPS deficiency on human hematopoiesis, we employed a xenotransplant model, in which human cord blood-derived CD34 cells reconstitute human hematopoiesis in mice (Yamauchi et al., 2013). To do so, we initiated RG3039 treatment 6 weeks after transplant and administered drug daily for 14 days (**Figure 5A**). We observed comparable proportions of blood cells of human origin ($hCD45^+$) in peripheral blood in

DMSO (vehicle)- and RG3039-treated mice (**Figure 5B**). Mice were euthanized after 14 injections (days), at which time we examined proportions of $hCD45^+$ cells and cellularity in BM. A small portion (2×10^6 cells) of BM mononuclear cells (BMMNCs) was transferred to irradiated recipients (6 mice per experimental group) for the 2nd transplant. While RG3039 treatment did not alter reconstitution of human hematopoiesis in recipients' BM upon 1st transplant (**Figure 5C**), subsets of BMMNCs were slightly skewed toward the myeloid lineage (**Figure 5D**). RG3039 treatment also did not alter proportions of hematopoietic stem/progenitor cells (HSPCs) (**Figure 5E**).

We next assessed effects of DCPS inhibition on human HSPC function via 2nd transplantation assay (**Figure 5A**) and observed comparable engraftment efficiency, regardless of BM source: engraftment of human blood cells was observed in 4/6, 3/6 and 4/6 of recipient mice that were transplanted with BM cells from DMSO-, RG3039 (10 mg/kg) - and RG3039 (20 mg/kg) treated donors in consistent with our previous study (Yamauchi et al., 2013) (**Figure S5F**). Furthermore, proportions of $hCD45^+$ cells in BM were comparable among experimental groups, indicating that DCPS inhibition does not grossly perturb human HSPC function (**Figure S5F**).

To explore effects of DCPS deficiency on normal hematopoiesis in humans, we examined PB counts of three children harboring germline homozygous loss-of-function

mutations as well as heterozygous relatives in a Jordanian family reported previously (Ng et al., 2015). The variant (c.201+2T>C) at the first splice donor site of intron 1 (**Figure 5F**) led to complete loss of the most abundant DCPS isoform as well as its decapping activity, leaving minimal levels of a novel DCPS isoform barely detectable at the protein level (Ng et al., 2015). Three individuals homozygous for the loss-of-function variant (III-1, -2 and -6) and the two carriers (II-8 and -9) exhibited normal blood counts (**Figures 5G** and **5H**), indicating that DCPS is dispensable for steady-state hematopoiesis in humans.

RG3039 exhibits anti-leukemia effects in human AML xenograft models

Next, we explored potential anti-leukemia effects of RG3039 in vivo in a mouse model of human AML. Specifically, we tested RG3039 efficacy using patient-derived xenograft (PDX) AML models established from three human AML lines (Townsend et al., 2016) (**Figure S6A**). After xenotransplantation, proportions of AML cells per total mononuclear cells in PB were examined over time by FACS, and RG3039 treatment was initiated once AML cells constituted 0.1-0.5% (**Figure 6A**). In the first cohort of mice, we monitored proportions of AML cells in PB until the last day of a two-week drug treatment (**Figure 6B**) and then euthanized mice to assess proportions of AML cells in BM (**Figure 6C**). RG3039 exhibited anti-leukemic activity, as evidenced by the lower leukemia burden in PB

and BM, although treatment response varied by lines tested (**Figures 6B** and **6C**). RG3039 binding to DCPS protein in AML cells in vivo was confirmed by CETSA using AML cells in BM (**Figure S6B**). In the second cohort, RG3039 was administered intraperitoneally for two weeks once AML cells were detected in PB as described above, and survival curves were generated. RG3039-treated mice survived significantly longer than vehicle-treated mice, indicative of therapeutic efficacy of RG3039 monotherapy against AML in vivo (**Figure 6D**). Of note, FHIT (fragile histidine triad), a newly-identified scavenger decapping enzyme (Taverniti and Séraphin, 2015), is differentially expressed in AML cells and normal hematopoietic cells (**Figure S6C**): Western blot analysis showed that human AML cells (cell lines, PDX lines and primary AML samples) express FHIT protein at significantly lower levels than do normal hematopoietic cells (cord blood and BMMNCs) (**Figure S6C**). Thus, lack of FHIT-mediated decapping activity may render AML cells more vulnerable to *DCPS* deficiency, accounting in part for leukemia-specific sensitivity to DCPS depletion or DCPS inhibition.

DISCUSSION

Investigators have discovered genes essential for cancer using genome-wide CRISPR-Cas9 screening (Hart et al., 2015; Tzelepis et al., 2016; Wang et al., 2017). We established two

mouse AML lines whose genetic backgrounds are well-defined: both exhibit a normal karyotype and harbor functionally-normal *Trp53*. This experimental system is particularly relevant as it mimics human primary AML cells where *TP53* is less frequently mutated than in solid tumors (Welch et al., 2012). Furthermore, an in vivo validation screen controlled for potential artifacts emerging from in vitro culture conditions. Another unique feature of our screen was that we selected potentially-actionable targets using databases for chemical inhibitors (Google search and DGIdb (Wagner et al., 2016)) and human genomic databases [ExAC (Lek et al., 2016), gnomAD (http://gnomad.broadinstitute.org), OMIM, ClinVar]. This approach was inspired by recent discovery of low-frequency coding-sequence variants that alter the risk of coronary artery disease and the successful translation of that approach to drug development (Abifadel et al., 2003; Myocardial Infarction and Investigators, 2016; Sabatine et al., 2015).

DCPS was identified as an actionable target for AML after rigorous discovery processes, including unbiased genome-wide screens in vitro, "manual" selection of genes implicated in a human genetics database, and then an in vivo validation screen We focused on DCPS for the following reasons: 1) sgRNAs targeting DCPS were markedly depleted both in vitro and in vivo, with high statistical significance; 2) RG3039, a DCPS inhibitor, was available and tested safe in a phase I clinical trial (Van Meerbeke et al., 2013),

suggesting it could be "re-purposed' as AML therapy; and 3) individuals homozygous for a *DCPS* loss-of-function mutation survive but exhibit intellectual disability and craniofacial and neuromuscular abnormalities, suggesting DCPS is dispensable in adulthood. Domain-mapping experiments further confirmed that DCPS is essential for AML cell survival and identified amino acid residues required for that activity.

RG3039 was originally developed as a drug to treat SMA (Jarecki et al., 2005). It inhibits DCPS catalytic activity via binding to a site within the HIT sequence, interfering with DCPS binding to the 7-methylguanosine (m^7G) cap of mRNA (Gu et al., 2004). RG3039 and its derivatives significantly improve symptoms and survival in SMA mouse models at similar doses that used in our in vivo studies (Gogliotti et al., 2013; Gopalsamy et al., 2017; Van Meerbeke et al., 2013). Importantly, RG3039 was granted Orphan Drug and Fast Track designations from the FDA and has been judged safe in a phase I trial in healthy volunteers (Gogliotti et al., 2013; Van Meerbeke et al., 2013), although details of the study (e.g. treatment schedule, dosage, side effects) were not currently available in publicly-available databases. Our data also indicate that RG3039-mediated DCPS inhibition does not grossly alter normal hematopoiesis. Furthermore, BM cells from RG3039-treated mice reconstituted human hematopoiesis similarly to controls in second transplant experiments, suggesting that DCPS inhibition does not perturb HSC/progenitor function.

Importantly, individuals with Al-Raqad syndrome (MIM 616459) who harbor biallelic loss-of-function germline *DCPS* mutations exhibit normal peripheral blood counts (Ahmed et al., 2015; Ng et al., 2015). While the loss-of-function allele (c.201+2T>C) reportedly expresses an aberrant transcript, its mRNA levels are very low and the endogenous protein is barely detected in patient-derived primary fibroblasts, which also lack detectable DCPS enzymatic activity (Ng et al., 2015). Since DCPS inhibitors with more potency and specificity are being developed (Gopalsamy et al., 2017), it would be interesting to test their anti-leukemia activity.

CETSA analysis confirmed binding of RG3039 and DCPS in AML cells as previously reported (Martinez Molina et al., 2013; Xu et al., 2016). RG3039 suppressed proliferation of human AML cells in vitro and in vivo. Since amino acid residues outside the HIT sequence that bind RG3039 were necessary for AML survival in a domain-mapping experiment, we expect that pharmacological inhibition of DCPS function beyond its enzymatic activity could exert more potent anti-leukemia effects. In fact, DCPS knockout (via CRISPR/Cas9) or knockdown (via shRNA) promotes more potent anti-leukemic activity than inhibition of its enzymatic activity by RG3039 treatment. For example, following DCPS knockdown in AML cells, expression of differentiation markers was more robust than was seen following RG3039 treatment. In this regard, a

Cereblon-based target degradation system (Winter et al., 2015) might be an interesting option to completely inactivate DCPS function.

It is unclear why AML cells are more vulnerable to DCPS deficiency than are cells functioning in normal human hematopoiesis. It also remains unknown exactly how DCPS inhibition causes cell cycle arrest, cellular apoptosis and differentiation in AML cells. Considering that DCPS interacts with multiple pre-mRNA processing machineries, either directly or indirectly (e.g. through RNA or chromatin), DCPS inhibition could broadly alter pre-mRNA metabolism in the nucleus, including pre-mRNA splicing.

RNA-seq experiments revealed that RG3039-induced mis-splicing is much less pronounced in GMPs, the normal counterpart of AML cells . Furthermore, genes mis-spliced to create alternative 3'ss and 5'ss vary between AML cells and GMPs and exhibit distinct NMD-prediction patterns. Importantly, an IFN signature was not evident in RG3039-treated GMPs, indicating differential effects of DCPS inhibition on pre-mRNA splicing and the overall transcriptome between cell types. RG3039 induced aberrant splicing, namely alternative splicing at 5' and 3' splice sites of the first exon, in marked contrast to aberrant splicing observed in the *SF3B1*-mutant chronic lymphocytic leukemia (CLL) cells (Darman et al., 2015), *Sf3b1*-mutant mouse myeloid progenitors (Obeng et al., 2016) or *Srsf2*-mutant mouse AML cells (Lee et al., 2016). We do not yet know how DCPS

depletion leads to pre-mRNA mis-splicing. However, two explanations are plausible. The first involves the nuclear cap binding complex (CBC), which consists of a 5'cap binding protein NCBP20 (a.k.a. CBP20) and NCBP1 (a.k.a. CBP80) and is implicated in pre-mRNA splicing (Gonatopoulos-Pournatzis and Cowling, 2014). CBC reportedly interacts with components of small nuclear ribonucleoproteins (snRNPs), which are necessary for efficient co-transcriptional spliceosomal assembly (Pabis et al., 2013). Multiple lines of evidence suggest that interaction between the 5' end m7G cap and CBC is necessary to splice the first intron (Inoue et al., 1989; Izaurralde et al., 1995; Izaurralde et al., 1994; Jiao et al., 2013). Like CBC, DCPS, which binds to the m7G cap in the nucleus, may function as a platform to recruit spliceosomes co-transcriptionally. An alternative is that in the nucleus CBC and DCPS may compete for m7G cap-containing pre-mRNAs, and DCPS depletion would favor engagement of CBC with the m7G cap, perturbing pre-mRNA splicing, as described (Shen et al., 2008).

RG3039 treatment affects splicing of more than 300 genes in AML cells, and DCPS complexes with multiple pre-mRNA processing machineries. Given that a type I interferon response was the most enriched gene expression signature following RG3039 treatment in AML cells, we postulate that DCPS inhibition broadly alters pre-mRNA metabolism, including splicing, and produces aberrant transcripts that induce a cell-intrinsic

response, such as ribotoxic stress-induced apoptosis, an outcome reminiscent of that observed in RNAi-treated cells (Bridge et al., 2003; Wurtmann and Wolin, 2009).

We report that RG3039 monotherapy had anti-leukemic effects and prolonged survival of human AML PDX models. In this model, instead of treating mice immediately after AML cell transfer, we initiated RG3039 treatment once they constituted 0.1-0.5% of peripheral blood mononuclear cells, at which point AML cells constituted 70-90% of BM cells. Considering that leukemia burden at the time of treatment was very high, and that these AML cells harbor complex genomic backgrounds and are refractory to multiple chemotherapies, the observation of anti-leukemic activity in vivo by RG3039 as a monotherapy is significant. Since somatic mutations of spliceosomal genes are common in myelodysplastic syndromes (MDS) and AML (Lee et al., 2016) and mRNA splicing is reportedly impeded in those cells (Darman et al., 2015; Lee et al., 2016; Obeng et al., 2016), it would be interesting to test whether AML/MDS cells harboring a spliceosome mutation are particularly sensitive to DCPS inhibition. Along the same line, combinatorial treatment using RG3039 plus a compound targeting spliceosomal function, such as an SF3B1 inhibitor (Lee and Abdel-Wahab, 2016), might be a potential therapeutic approach worthy of investigation.

In summary, we identified DCPS as a target for AML therapy. While DCPS is implicated in the mRNA decay pathway, it is primarily localized to the nucleus and functions to maintain pre-mRNA metabolism; moreover, its loss renders leukemia cells vulnerable to cell cycle arrest and apoptosis. Since DCPS is dispensable for steady-state-hematopoiesis in humans, DCPS inhibitors, such as RG3039, warrant attention as potential therapeutic approaches for AML.

ACKNOWLEDGMENTS

We thank Hai Yi, Catherine Zhu and members in the Division of Hematology at BWH and Department of Medicine and Biosystemic Science at Kyushu University for assistance, advice and helpful discussion, Zach Herbert for help and advice on sequencing and Elise Lamar for critical reading of the manuscript. This work is supported in part by a Grant-in-Aid for JSPS Fellows (15J10130) (to T.Y.), an NIDDK Predoctoral National Research Service Award for an M.D./Ph.D. Fellowship (F30DK103359-01A1) (to M.C.C.), a Strategic Positioning Fund for Genetic Orphan Diseases from the Agency for Science, Technology, and Research in Singapore (to B.R), NIH R00HG008399 (to L.P.), NIDDK (K08DK093705, R03DK109232), NHLBI (DP2OD022716), Burroughs Wellcome Fund, and American Society of Hematology (to D.E.B.), NIH P01HL032262 and P30DK049216 grants (Center of Excellence in Molecular Hematology) (to S.H.O.), a Grant-in-Aid for Scientific Research (S)(16H06391)(to K.A.) and a Grant-in-Aid for Scientific Research (A) (17H01567), American Cancer Society (RSG-13-379-01-LIB) and an American Society of Hematology Bridge Grant (to T.M.).

AUTHOR CONTRIBUTIONS

T.Y., M.C.C., L.P., D.E.B. and T.M. designed CRISPR-Cas9 screen experiments. T.Y., Y.S. and T.M reviewed CRISPR screen data. T.M., M.M., Y.I., Y.S. and M.N. executed CRISPR-CAS9 experiments, cell biology experiments, RNA-Seq, western blot analysis, immunohistochemistry and in vivo mouse studies, supervised by F.A., L.P., S.H.O., K.A., D.E.B. and T.M. T.M. performed mass-spectrometry and analyzed proteomic data. M.S., S.B. Y.S., Q.Y. and L.P. analyzed RNA-Seq data. M.S., M. A. R. and B.R. provided patient data. M.C.C., V.A.C.S., M.A.C. and C.M.T. analyzed CRISPR saturation mutagenesis data, supervised by L.P. and D.E.B. T.Y and T.M. wrote the manuscript with help from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Abifadel, M., Varret, M., Rabès, J.-P., Allard, D., Ouguerram, K., Devillers, M., Cruaud, C., Benjannet, S., Wickham, L., Erlich, D.*, et al.* (2003). Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet *34*, 154-156.

Ahmed, I., Buchert, R., Zhou, M., Jiao, X., Mittal, K., Sheikh, T. I., Scheller, U., Vasli, N., Rafiq, M. A., Brohi, M. Q.*, et al.* (2015). Mutations in DCPS and EDC3 in autosomal recessive intellectual disability indicate a crucial role for mRNA decapping in neurodevelopment. Hum Mol Genet *24*, 3172-3180.

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A.-L.*, et al.* (2013). Signatures of mutational processes in human cancer. Nature *500*, 415-421.

Andersen, P. R., Domanski, M., Kristiansen, M. S., Storvall, H., Ntini, E., Verheggen, C., Schein, A., Bunkenborg, J., Poser, I., Hallais, M.*, et al.* (2013). The human cap-binding complex is functionally connected to the nuclear RNA exosome. Nat Struct Mol Biol *20*, 1367-1376.

Boehm, J. S., and Hahn, W. C. (2011). Towards systematic functional characterization of cancer genomes. Nat Rev Genet *12*, 487-498.

Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol *34*, 525-527.

Bridge, A. J., Pebernard, S., Ducraux, A., Nicoulaz, A.-L., and Iggo, R. (2003). Induction of an interferon response by RNAi vectors in mammalian cells. Nat Genet *34*, 263-264.

Canver, M. C., Lessard, S., Pinello, L., Wu, Y., Ilboudo, Y., Stern, E. N., Needleman, A. J., Galactéros, F., Brugnara, C., Kutlar, A.*, et al.* (2017). Variant-aware saturating mutagenesis using multiple Cas9 nucleases identifies regulatory elements at trait-associated loci. Nat Genet.

Canver, M. C., Smith, E. C., Sher, F., Pinello, L., Sanjana, N. E., Shalem, O., Chen, D. D., Schupp, P. G., Vinjamur, D. S., Garcia, S. P.*, et al.* (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature *527*, 192-197.

Cho, S. W., Kim, S., Kim, J. M., and Kim, J. S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol *31*, 230-232.

Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., and de Hoon, M. J. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422-1423.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas

systems. Science *339*, 819-823.
Darman, R. B., Seiler, M., Agrawal, A. A., Lim, K. H., Peng, S., Aird, D., Bailey, S. L., Bhavsar, E. B., Chan, B., Colla, S.*, et al.* (2015). Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. Cell Rep *13*, 1033-1045.
Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.
Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R.*, et al.* (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol *34*, 184-191.
Ferrara, F., and Schiffer, C. A. (2013). Acute myeloid leukaemia in adults. Lancet *381*, 484-495.
Garraway, L. A., and Lander, E. S. (2013). Lessons from the cancer genome. Cell *153*, 17-37.
Gogliotti, R. G., Cardona, H., Singh, J., Bail, S., Emery, C., Kuntz, N., Jorgensen, M., Durens, M., Xia, B., Barlow, C.*, et al.* (2013). The DcpS inhibitor RG3039 improves survival, function and motor unit pathologies in two SMA mouse models. Hum Mol Genet *22*, 4084-4101.
Gonatopoulos-Pournatzis, T., and Cowling, V. H. (2014). Cap-binding complex (CBC). Biochem J *457*, 231-242.
Gopalsamy, A., Narayanan, A., Liu, S., Parikh, M. D., Kyne, R. E., Fadeyi, O., Tones, M. A., Cherry, J. J., Nabhan, J. F., LaRosa, G.*, et al.* (2017). Design of Potent mRNA Decapping Scavenger Enzyme (DcpS) Inhibitors with Improved Physicochemical Properties To Investigate the Mechanism of Therapeutic Benefit in Spinal Muscular Atrophy (SMA). J Med Chem *60*, 3094-3108.
Gu, M., Fabrega, C., Liu, S.-W., Liu, H., Kiledjian, M., and Lima, C. D. (2004). Insights into the structure, mechanism, and regulation of scavenger mRNA decapping activity. Mol Cell *14*, 67-80.
Han, G. W., Schwarzenbacher, R., McMullan, D., Abdubek, P., Ambing, E., Axelrod, H., Biorac, T., Canaves, J. M., Chiu, H.-J., Dai, X.*, et al.* (2005). Crystal structure of an Apo mRNA decapping enzyme (DcpS) from Mouse at 1.83 A resolution. Proteins *60*, 797-802.
Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K. R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S.*, et al.* (2015). High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell.
Inoue, K., Ohno, M., Sakamoto, H., and Shimura, Y. (1989). Effect of the cap structure on pre-mRNA splicing in Xenopus oocyte nuclei. Genes Dev *3*, 1472-1479.
Izaurralde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C., and Mattaj, I. W. (1995). A cap-binding protein complex mediating U snRNA export. Nature *376*, 709-712.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I. W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell *78*, 657-668.
Jarecki, J., Chen, X., Bernardino, A., Coovert, D. D., Whitney, M., Burghes, A., Stack, J., and Pollok, B. A. (2005). Diverse small-molecule modulators of SMN expression found by high-throughput compound screening: early leads towards a therapeutic for spinal muscular atrophy. Hum Mol Genet *14*, 2003-2018.
Jiao, X., Chang, J. H., Kilic, T., Tong, L., and Kiledjian, M. (2013). A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. Mol Cell *50*, 104-115.
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816-821.
Koike-Yusa, H., Li, Y., Tan, E. P., Velasco-Herrera, M. D. C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol *32*, 267-273.
Lawrence, M. S., Stojanov, P., Mermel, C. H., Robinson, J. T., Garraway, L. A., Golub, T. R., Meyerson, M., Gabriel, S. B., Lander, E. S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. Nature *505*, 495-501.
Lee, S. C.-W., and Abdel-Wahab, O. (2016). Therapeutic targeting of splicing in cancer. Nat Med *22*, 976-986.
Lee, S. C.-W., Dvinge, H., Kim, E., Cho, H., Micol, J.-B., Chung, Y. R., Durham, B. H., Yoshimi, A., Kim, Y. J., Thomas, M.*, et al.* (2016). Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins. Nat Med *22*, 672-678.
Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., O'Donnell-Luria, A. H., Ware, J. S., Hill, A. J., Cummings, B. B.*, et al.* (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285-291.
Li, W., Xu, H., Xiao, T., Cong, L., Love, M. I., Zhang, F., Irizarry, R. A., Liu, J. S., Brown, M., and Liu, X. S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol *15*, 554.
Lim, E. T., Würtz, P., Havulinna, A. S., Palta, P., Tukiainen, T., Rehnström, K., Esko, T., Mägi, R., Inouye, M., Lappalainen, T.*, et al.* (2014). Distribution and medical impact of loss-of-function variants in the Finnish founder population. PLoS Genet *10*, e1004494.
Lykke-Andersen, S., and Jensen, T. H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat Rev Mol Cell Biol *16*, 665-677.

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., and Church, G. M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.
Martinez Molina, D., Jafari, R., Ignatushchenko, M., Seki, T., Larsson, E. A., Dan, C., Sreekumar, L., Cao, Y., and Nordlund, P. (2013). Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. Science *341*, 84-87.
Milac, A. L., Bojarska, E., and Wypijewska del Nogal, A. (2014). Decapping Scavenger (DcpS) enzyme: advances in its structure, activity and roles in the cap-dependent mRNA metabolism. Biochim Biophys Acta *1839*, 452-462.
Myocardial Infarction, G., and Investigators, C. A. E. C. (2016). Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. N Engl J Med.
Narasimhan, V. M., Hunt, K. A., Mason, D., Baker, C. L., Karczewski, K. J., Barnes, M. R., Barnett, A. H., Bates, C., Bellary, S., Bockett, N. A.*, et al.* (2016). Health and population effects of rare gene knockouts in adult humans with related parents. Science.
Ng, C. K. L., Shboul, M., Taverniti, V., Bonnard, C., Lee, H., Eskin, A., Nelson, S. F., Al-Raqad, M., Altawalbeh, S., Séraphin, B., and Reversade, B. (2015). Loss of the scavenger mRNA decapping enzyme DCPS causes syndromic intellectual disability with neuromuscular defects. Hum Mol Genet *24*, 3163-3171.
Obeng, E. A., Chappell, R. J., Seiler, M., Chen, M. C., Campagna, D. R., Schmidt, P. J., Schneider, R. K., Lord, A. M., Wang, L., Gambe, R. G.*, et al.* (2016). Physiologic Expression of Sf3b1(K700E) Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. Cancer Cell *30*, 404-417.
Pabis, M., Neufeld, N., Steiner, M. C., Bojic, T., Shav-Tal, Y., and Neugebauer, K. M. (2013). The nuclear cap-binding complex interacts with the U4/U6·U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. RNA *19*, 1054-1063.
Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods *14*, 417-419.
Rivas, M. A., Pirinen, M., Conrad, D. F., Lek, M., Tsang, E. K., Karczewski, K. J., Maller, J. B., Kukurba, K. R., DeLuca, D. S., Fromer, M.*, et al.* (2015). Effect of predicted protein-truncating genetic variants on the human transcriptome. Science *348*, 666-669.
Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.
Sabatine, M. S., Giugliano, R. P., Wiviott, S. D., Raal, F. J., Blom, D. J., Robinson, J., Ballantyne, C. M., Somaratne, R., Legg, J., Wasserman, S. M.*, et al.* (2015). Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. N Engl J Med *372*, 1500-1509.
Saleheen, D., Natarajan, P., Armean, I. M., Zhao, W., Rasheed, A., Khetarpal, S. A., Won, H.-H.,

Karczewski, K. J., O'Donnell-Luria, A. H., Samocha, K. E., *et al.* (2017). Human knockouts and phenotypic analysis in a cohort with a high rate of consanguinity. Nature *544*, 235-239.
Sanjana, N. E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nature methods *11*, 783-784.
Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G., and Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science *343*, 84-87.
Shen, V., Liu, H., Liu, S.-W., Jiao, X., and Kiledjian, M. (2008). DcpS scavenger decapping enzyme can modulate pre-mRNA splicing. RNA *14*, 1132-1142.
Shi, J., Wang, E., Milazzo, J. P., Wang, Z., Kinney, J. B., and Vakoc, C. R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nat Biotechnol *33*, 661-667.
Sulem, P., Helgason, H., Oddson, A., Stefansson, H., Gudjonsson, S. A., Zink, F., Hjartarson, E., Sigurdsson, G. T., Jonasdottir, A., Jonasdottir, A., *et al.* (2015). Identification of a large set of rare complete human knockouts. Nat Genet *47*, 448-452.
Taverniti, V., and Séraphin, B. (2015). Elimination of cap structures generated by mRNA decay involves the new scavenger mRNA decapping enzyme Aph1/FHIT together with DcpS. Nucleic Acids Res *43*, 482-492.
Townsend, E. C., Murakami, M. A., Christodoulou, A., Christie, A. L., Köster, J., DeSouza, T. A., Morgan, E. A., Kallgren, S. P., Liu, H., Wu, S.-C., *et al.* (2016). The Public Repository of Xenografts Enables Discovery and Randomized Phase II-like Trials in Mice. Cancer Cell *29*, 574-586.
Tsherniak, A., Vazquez, F., Montgomery, P. G., Weir, B. A., Kryukov, G., Cowley, G. S., Gill, S., Harrington, W. F., Pantel, S., Krill-Burger, J. M., *et al.* (2017). Defining a Cancer Dependency Map. Cell *170*, 564-576.e516.
Tzelepis, K., Koike-Yusa, H., De Braekeleer, E., Li, Y., Metzakopian, E., Dovey, O. M., Mupo, A., Grinkevich, V., Li, M., Mazan, M., *et al.* (2016). A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. Cell Rep *17*, 1193-1205.
Van Meerbeke, J. P., Gibbs, R. M., Plasterer, H. L., Miao, W., Feng, Z., Lin, M.-Y., Rucki, A. A., Wee, C. D., Xia, B., Sharma, S., *et al.* (2013). The DcpS inhibitor RG3039 improves motor function in SMA mice. Hum Mol Genet *22*, 4074-4083.
Visa, N., Izaurralde, E., Ferreira, J., Daneholt, B., and Mattaj, I. W. (1996). A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. J Cell Biol *133*, 5-14.
Wagner, A. H., Coffman, A. C., Ainscough, B. J., Spies, N. C., Skidmore, Z. L., Campbell, K. M., Krysiak, K., Pan, D., McMichael, J. F., Eldred, J. M., *et al.* (2016). DGIdb 2.0: mining clinically

relevant drug-gene interactions. Nucleic Acids Res *44*, D1036-1044.

Wang, T., Wei, J. J., Sabatini, D. M., and Lander, E. S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science *343*, 80-84.

Wang, T., Yu, H., Hughes, N. W., Liu, B., Kendirli, A., Klein, K., Chen, W. W., Lander, E. S., and Sabatini, D. M. (2017). Gene Essentiality Profiling Reveals Gene Networks and Synthetic Lethal Interactions with Oncogenic Ras. Cell.

Welch, J. S., Ley, T. J., Link, D. C., Miller, C. A., Larson, D. E., Koboldt, D. C., Wartman, L. D., Lamprecht, T. L., Liu, F., Xia, J.*, et al.* (2012). The origin and evolution of mutations in acute myeloid leukemia. Cell *150*, 264-278.

Winter, G. E., Buckley, D. L., Paulk, J., Roberts, J. M., Souza, A., Dhe-Paganon, S., and Bradner, J. E. (2015). Phthalimide conjugation as a strategy for in vivo target protein degradation. Science *348*, 1376-1381.

Wurtmann, E. J., and Wolin, S. L. (2009). RNA under attack: cellular handling of RNA damage. Crit Rev Biochem Mol Biol *44*, 34-49.

Xu, H., Gopalsamy, A., Hett, E. C., Salter, S., Aulabaugh, A., Kyne, R. E., Pierce, B., and Jones, L. H. (2016). Cellular thermal shift and clickable chemical probe assays for the determination of drug-target engagement in live cells. Org Biomol Chem.

Yamauchi, T., Takenaka, K., Urata, S., Shima, T., Kikushige, Y., Tokuyama, T., Iwamoto, C., Nishihara, M., Iwasaki, H., Miyamoto, T.*, et al.* (2013). Polymorphic Sirpa is the genetic determinant for NOD-based mouse lines to achieve efficient human cell engraftment. Blood *121*, 1316-1325.

Zhang, L., McGraw, K. L., Sallman, D. A., and List, A. F. (2016). The role of p53 in myelodysplastic syndromes and acute myeloid leukemia: molecular aspects and clinical implications. Leuk Lymphoma, 1-14.

FIGURE TITLE AND LEGENDS

Figure 1. Genome-wide CRISPR-Cas9 screens identify *Dcps* as an AML essential gene.

(A) Generation of Cas9-expressing mouse AML cell lines. Two mouse lines expressing Cas9 endonuclease (CALM/AF10-Cas9 and MLL/AF9-Cas9) were used for screens. **(B)** Genes significantly enriched or dropped-out after a 16-day incubation were identified using the MAGeCK program (Li et al., 2014; Shalem et al., 2014). Representative results (GeCKO library B screen in MLL/AF9 cells) of the enrichment screen are shown. A modified robust ranking aggregation (RRA) algorithm was used to rank sgRNAs based on p values (Li et al., 2014; Shalem et al., 2014). **(C)** Graphs show read counts of individual sgRNAs targeting *Trp53* before and after a 16-day incubation. P values were calculated using a Wilcoxon matched-pairs signed rank test. **(D)** Experimental schema of in vitro and in vivo CRISPR-Cas9 screens. Overall, we identified 130 AML essential genes for further evaluation. **(E)** A representative RRA score plot showing top dropout genes (GeCKO library B screen in CALM/AF10 cells) . **(F)** Read counts of sgRNAs targeting *Dcps* significantly decreased after a 16-day incubation in both AML lines. P values were calculated using a Wilcoxon matched-pairs signed rank test. **(G)** Read counts of single sgRNAs targeting *Dcps* before and 3 weeks after leukemia transfer are shown. Read counts of 7 sgRNAs are shown, as one out of 8 sgRNA (#14) was not detected on day 0. **(H)**

Domain-saturating DCPS mutagenesis. All NGG-restricted sgRNAs (n=154) were identified within *Dcps* coding exons. The pool was transduced into CALM/AF10-Cas9 cells and a dropout screen was performed. Read counts from final and initial time points were normalized to non-targeting guides and log2 fold-change in guide abundance was calculated. Guides were then mapped to protein by mapping double-stranded break sites to a codon. **(I)** Dropout scores (log2 fold-change) for each amino acid were mapped onto a structure publicly available in the Protein Data Bank (PDB ID: 1vlr). Since DCPS forms a homodimer, one monomer is depicted in pink for visual clarity, and scores are divided into bins of 1 log2 fold-change. Red asterisk indicates HIT sequence, where the 7-methylhuaninosine (m^7G) cap of mRNA binds. See also Figures S1, S2 and Tables S1, S2 and S3.

Figure 2. Treatment of AML cells with the DCPS inhibitor RG3039 induces cell-cycle arrest, apoptosis and differentiation. (A) Human AML cell lines were transduced with a lentivirus vector encoding an shRNA and GFP cassette, and the fraction of GFP-positive cells was measured at indicated times by FACS. At each time point that proportion was normalized to the number of GFP-positive cells present at day 3 (two days after transduction). Scrambled-shRNA served as control. The data is from one experiment. **(B)**

shRNA knockdown efficiencies were assessed by Western blot. Cont: control sample without transduction. Scr: scrambled shRNA. **(C)** EdU (5-ethynyl-2´-deoxyuridine) incorporation assay. MOLM-13 cells transduced with either scrambled or DCPS shRNA clone#69 were pulsed with EdU and proportions of EdU-positive cells were assessed by FACS on indicated days after transduction. **(D)** Proportions of apoptotic cells were assessed by Annexin V staining. **(E)** Expression levels of surface CD15 and CD11b in MOLM-13 cells were assessed by FACS following shRNA-mediated DCPS knockdown (day 5). **(F)** May-Giemsa stain of cytospin preparations showing monocytic differentiation of MOLM-13 cells upon DCPS knockdown. **(G)** Representative result of a cellular thermal shift assay (CETSA). **(H)** DCPS protein is resistant to heat denaturation in the presence of RG3039 dose-dependently. **(I)** Growth curves were generated for AML lines treated with RG3039. Data are represented as means ± SD. **(J)** MOLM-13 cells were pulsed with EdU and treated with either DMSO or RG3039. Proportions of EdU-positive cells and their DNA contents were analyzed by FACS 24 hr and 72 hr after RG3039 treatment. **(K)** Bar graphs show proportions of cells at each stage of the cell cycle after DMSO or RG3039 treatment. **(L)** Proportions of apoptotic cells were assessed by Annexin V stain following control DMSO or RG3039 treatment. Data (n=3) are represented as means ± SD. See also Figure S3.

Figure 3. DCPS protein interacts with components of pre-mRNA metabolic pathways.

(A) Immunoprecipitation (IP) of endogenous DCPS protein was performed in triplicate. DCPS was detected (red arrowhead) by Western blot when IP was performed with anti-DCPS antibody but not with control immunoglobulins (Ig), indicative of specificity (left). Asterisks denote non-specific immunoglobulin light and heavy chain signals. Protein lysates used for mass spectrometry were analyzed by SDS-PAGE and silver stained (right). **(B)** Volcano plot displaying results of triplicate IP/mass spectrometry experiments. Y-axis shows negative log10 p values, which represent reproducibility of events among three independent experiments; x-axis indicates log2 ratio of normalized protein abundance between anti-DCPS antibody and control Ig IPs. **(C)** Proposed model for nuclear DCPS function. DCPS complexes with components of pre-mRNA processing pathways including, spliceosomes, the transcription-export complex (TREX) and the nuclear pore complex (NUP). **(D)** Genes encoding DCPS-interacting proteins were found essential for AML survival in CRISPR-Cas9 screens. The Y-axis shows the lowest FDR q-value of each gene among 4 dropout screens (GeCKO library A and B screens in CALM/AF10 or MLL/AF9 cells); the X-axis indicates log2 ratio of normalized protein abundance between anti-DCPS antibody and control Ig IPs. See also Figure S4 and Table S4.

Figure 4. DCPS inhibition impedes pre-mRNA processing pathways in AML cells.

(A) RNA-Seq analysis of effects of DCPS inhibition on pre-mRNA splicing and transcriptome activity. RNA samples were prepared prior to and 6 hr and 10 hr after RG3039 treatment of CALM/AF10 mouse leukemia cells or GMPs. **(B)** Event counts and ratios relative to control of indicated splicing patterns (10 hr after RG3039 treatment) are shown. Data are represented as means ± SD. **(C)** Venn diagrams show overlap of mis-spliced genes between CALM/AF10 AML cells and GMPs. **(D)** Waterfall plots indicate NMD-sensitivities and expression changes of genes aberrantly-spliced following RG3039 treatment. **(E)** Locations of Mis-splicing events. **(F)** GSEA analysis of RG3039-treated AML cells. See also Table S5.

Figure 5. DCPS is dispensable for steady-state hematopoiesis in humans.

(A) Schematic representations of xenotransplant experiments. BM cells of mice treated with DMSO (vehicle) or RG3039 were transferred to secondary recipients to assess capacity to reconstitute human hematopoiesis. BRGS mice (Yamauchi et al., 2013) were used as recipients. **(B)** Proportions of human $CD45^+$ cells were assessed by FACS on indicated days after the first transplant. 6 mice per condition were analyzed at each time

point. **(C)** Dot graphs show proportions and numbers of $hCD45^+$ cells in BM after first transplant. n.s.: not significant. Data are represented as means ± SD (n=6 per condition). **(D)** Bar graphs show proportions of B cells ($hCD19^+$), T cells ($hCD3^+$) and myeloid cells ($hCD33^+$) in BM (n=6 per condition). Myeloid compartments were further defined using the myelo/monocytic markers CD11b, CD14 and CD15. Data are represented as means ± SD. P values were calculated using an unpaired t-test with Welch's correction. **(E)** Bar graphs show proportions of hematopoietic stem cells (HSCs: $CD34^+CD38^-$) and progenitors ($CD34^+CD38^+$) in BM. Data are represented as means ± SD (n=6 per condition). **(F)** Schematic representations of the germline loss-of-function mutation observed in a Jordanian family (Ng et al., 2015). Thymine (red) near the splice donor site downstream of exon 1 was mutated to cytosine, creating an alternative cryptic splice site (blue) and resulting in an in-frame premature termination 40bp downstream exon 1. **(G)** Family tree of the affected pedigree. Asterisks denote individuals whose peripheral blood counts were assessed in this study. **(H)** Peripheral blood counts of affected individuals and unaffected relatives are shown. See also Figure S5.

Figure 6. Treatment with a DCPS inhibitor has anti-leukemia effects in AML PDX models. (A) Workflow of PDX experiments. **(B)** Proportions of $hCD45^+$ cells in PB were

examined by FACS. Duration of RG3039 treatment (20 mg/kg for 14 days) is depicted in grey. Data are represented as means ± SD (n=4-5). **(C)** Mice were euthanized after RG3039 treatment, and BM leukemia burden was assessed by FACS. Dot graphs show proportions of $hCD45^+$ cells in BM. Data are represented as means ± SD (n=3-5). P values were calculated using an unpaired t-test with Welch's correction. **(D)** Survival curves (n=5 per group). Duration of RG3039 treatment (20 mg/kg for 14 days) is depicted in grey. P values were calculated using a Mantel-Cox log-rank test. See also Figure S6.

STAR★METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to the Lead Contact, Takahiro Maeda (t_maeda@cancer.med.kyushu-u.ac.jp).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

12-16 weeks-old C57BL/6 male mice were purchased from Jackson Laboratory. C57BL/6.*Rag2*^{null}*Il2rg*^{null}NOD-*Sirpa* (BRGS) mice were described previously (Yamauchi et al., 2013). 6-10 weeks-old BRGS mice were used for experiments. Mice were bred and maintained in individual ventilated cages and fed with autoclaved food and water at Kyushu University Animal Facility. All animal experiments were approved by the Institutional Animal Care and Use Committees, according to national and institutional guidelines.

Generation of Cas9-expressing mouse AML cell lines

Mouse bone marrow HSCs were harvested from male C57BL/6 mice and transduced with a retrovirus encoding the CALM/AF10 or MLL/AF9 oncogene. Transduced cells were transplanted into lethally-irradiated recipient mice. Primary AML cells were collected 3-6 months later, serially transplanted 3 times to obtain clonally homogenous population, and harvested to establish cell lines (Figure 1A). Cells were cultured in the presence of cytokines as described above and infected with lentivirus encoding both the *S. pyogenes* Cas9 protein and the blasticidin resistance gene (lentiCas9-Blast, Addgene). Blasticidin

(Sigma) selection was initiated 24 h later. *Trp53* mutant CALM/AF10 or MLL/AF9 AML cells were established by lentivirally transducing sgRNA against *Trp53* into parental AML lines, followed by Nutlin-3a, a Mdm2 inhibitor, treatment.

Cell culture

Mouse leukemia cells (CALM/AF10 and MLL/AF9) were cultured in Iscove's Modified Dulbecco's Media (IMDM) (Life Technologies) supplemented with 20% Fetal Bovine Serum (FBS) (Omega Scientific), 1% penicillin streptomycin (Life Technologies), and mouse SCF (20 ng/ml), mouse IL-3 (10 ng/ml) and mouse IL-6 (10 ng/ml) (all from PeproTech). Human leukemia cell lines (MOLM-13, OCI-AML3, MV4-11, and THP-1) were cultured in Roswell Park Memorial Institute (RPMI) medium (Life Technologies) with 10% FBS and 1% penicillin streptomycin.

Lentiviral transduction

Lentivirus transduction was performed as described (Shalem et al., 2014). In brief, HEK293T cells were cultured with Dulbecco’s Modified Eagle’s Medium (DMEM) (Life Technology) supplemented with 10% FBS and 1% penicillin stereptomycin. Cells were transfected with 6.7 μg psPAX2 (Addgene), 4.1 μg VSV-G (Addgene), and 10 μg lentiviral

vectors using 60 μg of linear polyethylenimine (Polysciences). Lentiviral supernatants were harvested at 48 and 72 hours post-transfection and concentrated by ultracentrifugation (24,000 rpm for 2 hours at 4℃ with a Beckman Coulter SW 32 Ti rotor). AML cells were plated into 12-well plates (0.5-3 x 10^6 cells per well) with medium supplemented with 8 μg/ml polybrene (Sigma) and spin-infected at 2,000 rpm for 2 hours at 37℃.

Xenotransplantation

Lineage (Lin)-depleted cord blood (CB) cells were obtained using magnetic beads (Lineage cell depletion kit; Miltenyi Biotec, Bergisch Gladbach, Germany). To reconstitute human hematopoiesis in mice, Lin^- CB cells (2×10^4 per mouse) were injected into irradiated (550 cGy) BRGS mice via the tail vein (Yamauchi et al., 2013). Mice were euthanized 6 weeks later, and their BM cells (2×10^6 per mouse) were transferred to irradiated secondary recipients to assess HSPC reconstitution capacity. Recipients of the second transplant were euthanized 6 weeks later, and engraftment of human blood cells in BM was assessed by FACS.

To establish patient-derived xenograft (PDX) models, three PDX lines were obtained from Leukemia/Lymphoma Xenograft Core (LLX) of the Dana-Farber Cancer Institute (DFCI). Somatic mutations in cells were assessed via the Rapid Heme Panel

(DFCI). 1-2 x 10^6 AML cells were injected into irradiated (550 cGy) BRGS mice (Yamauchi et al., 2013) via the tail vein. After transplantation, mice were given sterile water containing prophylactic enrofloxacin (Baytril; Bayer HealthCare).

Peripheral blood analysis in patients with biallelic *DCPS* loss-of-function

A Jordanian family consisting of three children (III-1, III-2 and III-6) exhibiting intellectual disability, craniofacial and neuromuscular abnormalities and two unaffected siblings (III-7 and III-8) were described previously (Ng et al., 2015). Peripheral blood counts of all were examined with standard methodology using an automated analyzer, with consent of patient or parent.

METHOD DETAILS

Immunohistochemistry

Immunohistochemistry was performed on paraffin sections with the conventional avidin–biotin–peroxidase method. Antigen retrieval was performed using citrate buffer. Endogenous peroxidases were quenched by incubating sections in 3% H_2O_2 in PBS for 5 min. Sections were blocked in PBS with 0.1% Tween-20 (PBST) for 1h. DCPS antibody (Santa Cruz,

A-12) was added overnight at 4℃. Slides were incubated with biotin-conjugated secondary antibodies (1:400; Vector Laboratories) using a VECTA-STAIN Elite ABC kit (Vector Laboratories) and developed with 3,3'-diaminobenzidine (DAB). Slides were mounted with Permount medium (Thermo Fisher Scientific).

Genome-wide CRISPR screen

The GeCKO v2 mouse library (Shalem et al., 2014) was purchased from Addgene. It consists of two half-libraries (A and B), containing altogether 130,209 sgRNAs (67,405 in Library A and 62,804 in Library B) targeting 20,611 protein-coding genes and 1,175 miRNAs plus 1,000 non-targeting control sgRNAs. The plasmid pool was prepared as described (https://www.addgene.org/pooled-library/zhang-human-gecko-v2/). The library representation was validated via Miseq (Illumina). Genome-wide screens were performed as described (Shalem et al., 2014). Briefly, $1.2x10^8$ AML cells (per half library) were transduced at a transduction efficiency of 30% with the viral pool to achieve an average coverage of more than x500. Cells were treated with puromycin (Sigma) 24 h after transduction (on day 1) and $3x10^6$ cells were harvested two days later (day 3) to obtain input DNA. Remaining cells were passaged 5-6 times over a 16-day incubation period. At least $3x10^7$ cells were maintained at any given time to ensure sgRNA representation (Figure

S1B). On day18, cells were harvested and genomic DNA extracted using a Blood & Cell Culture DNA Midi Kit (Qiagen).

sgRNA library design for the second screen

470 genes were chosen for a secondary screen in vivo as described (Figure 1D). For each gene, eight sgRNAs were chosen from the previously published mouse genome-wide sgRNA libraries "Brie" and "Asiago" (Doench et al., 2016). If fewer than 8 sgRNAs per gene were available within these libraries, sgRNA sequences were taken from the mouse "GeCKOv2" genome-wide library until 8 sgRNAs were obtained (Sanjana et al., 2014). The final library included 3,760 gene-targeting sgRNAs and 100 non-targeting sgRNAs as negative controls. Each sgRNA oligo was synthesized as described (Canver et al., 2015) and cloned using a Gibson Assembly master mix (New England Biolabs) into lentiGuide-Puro plasmid (Addgene plasmid no. 52963), which had been BsmBI-digested, PCR-purified, and dephosphorylated. Gibson Assembly products were transformed to electrocompetent cells (E. cloni, Lucigen). Sufficient colonies were isolated to ensure ~50x coverage of the library. The plasmid library was deep-sequenced using Miseq to confirm representation.

In vivo second screen in a mouse AML model

$2x10^7$ MLL-AF9-Cas9 cells were transduced with the viral pool at transduction efficiencies of 20% to achieve over x1,000 coverage (Figures S2C and S2D). Cells were treated with puromycin 24h after transduction and cultured 2 days, at which time $2x10^6$ puromycin-resistant cells were harvested to obtain input DNA. Remaining cells were transferred to sub-lethally irradiated recipient mice (Figure S2C). Mice were euthanized 3 weeks later, and AML cells were harvested from bone marrow to obtain output DNA samples.

Domain saturating mutagenesis

All NGG-restricted sgRNAs were identified within coding exons of *Dcps* (n=154) and *Ctps* (n=238) (Canver et al., 2017). 100 non-targeting sgRNAs were included as negative controls. Library synthesis was performed as per the second screening library, as noted above. Read counts from final and initial time points were normalized to control-nontargeting guides via the MAGeCK count function (Li et al., 2014), which was used to calculate log2 fold-changes in guide abundance. Guides were then mapped to the protein by mapping the double-stranded break site to the corresponding codon. Lastly, scores for amino acids with no assigned guide were interpolated via LOESS regression,

using known guide scores and location as input. Scores for each amino acid were then mapped onto the DCPS structure publicly available in the Protein Data Bank (PDB ID: 1vlr (Han et al., 2005)). Structure sequence was aligned to the sequence of the protein isoform to which guides were originally mapped using Biopythons pairwise2 module (Cock et al., 2009) (local alignment with Blosum62 matrix, opening gap cost -10, extension -0.5). Scores from guides mapping to the same amino acid were averaged. The protein structure was recolored in PyMOL (https://www.pymol.org) based on aligned scores, by placing them in the B-factor field. For visual clarity, scores were divided into bins of 1 log2 fold-change.

Library preparation for next generation sequencing

Library preparation was performed as described (Shalem et al., 2014). sgRNA inserts were PCR-amplified from 130 μg genomic DNA using Herculase 2 fusion DNA polymerase (Agilent). Resulting PCR products were purified and sequenced on a NextSeq500 sequencer (Illumina) to assess change in abundance of each sgRNA between initial and final cell populations.

shRNA-mediated DCPS knockdown

Lentivirus vectors containing shRNA against human DCPS were purchased from Sigma Aldrich (TRCN0000335923, TRCN0000335924 and TRCN0000005569). Vector fragments containing shRNA sequence were isolated by MfeI/SpeI digestion of the pLKO vector, gel-purified and sub-cloned into PPIG (pLKO Puro-IRES-GFP) lentivirus vectors. Lentivirus transduction was performed as described above.

Cellular thermal shift assay (CETSA)

CETSA was performed to assess binding of RG3039 to DCPS protein in cells as described (Martinez Molina et al., 2013; Xu et al., 2016).

Cell cycle and apoptosis assay

For cell cycle analysis, a 5-ethynyl-2'-deoxyuridine (EdU) incorporation assay was performed using the Click-iT Plus EdU Alexa Fluor 647 Flow Cytometry Assay Kit following the manufacturer's specifications (Life Technologies). FxCycle™ Violet (Life Technologies) was used for DNA content analysis. FITC-Annexin V (BioLegend) was used to identify apoptotic cells.

Immunoprecipitation/mass spectrometry

MOLM-13 cells were cultured in RPMI containing 10% FBS. Proteins were extracted with RIPA Buffer. DCPS complexes were pulled down using anti-DCPS antibody (Bethyl) followed by separation on SDS-PAGE. Proteins in gels were reduced using 10 mM dithiothreitol for 30 min, and alkylated using 50 mM iodoacetamide for 30 min. After overnight trypsin digestion, peptides were purified with SDB-XC StageTip. NanoLC-MS/MS were conducted using a Triple TOF 6600 (SCIEX) with an ekspert nanoLC 415 system (Eksigent Technologies). The mobile phase consisted of (A) 0.1% FA in water and (B) 0.1% FA in acetonitrile. A linear gradient of 5% B to 32% B in 60 min, 32% B to 80% B in 5 min and 80% B for 5 min was employed. Samples were injected onto a ChromXP C18CL (3 μm, 200 μm x 0.5 mm) column serving as a trap column and then separated on a ChromXP C18CL (3 μm, 75 μm x 15 mm) column. For Information Dependent Acquisition, precursor ions were scanned from 400 to 1250 with an accumulation time of 50 msec. For sequential window acquisition of all theoretical mass spectra (SWATH), precursor ions ware taken by a Variable Q1 Window Acquisition from 100 to 2000. MS data were subjected to a search against the Uniprot Human database with Protein Pilot V.5.0 (AB Sciex). The false discovery rate was set below 1%. The spectral library for SWATH analysis was generated from IDA data, and peak area of peptides was calculated by PeakView (AB Sciex).

RNA-Seq

Two independent primary AML lines per oncogene (CALM/AF10 or MLL/AF9) were obtained from the bone marrow of leukemic mice (CALM/AF10 #1 and #2 and MLL/AF9 #1 and #2; Figure S2A) and RNA was extracted using an RNeasy Mini Kit (Qiagen). Libraries were generated using the Ovation Mouse RNA-Seq System following the manufacturer's specifications (NuGEN). All libraries underwent 75 bp single-end sequencing on an Illumina NextSeq500 sequencer. For RNA-seq of RG3039-treated AML cells (Figure 4), CALM/AF10 cells or GMPs ($Lin^{-}Sca\text{-}1^{-}c\text{-}Kit^{+}Fc\gamma R^{+}CD34^{+}$) were treated with RG3039 (1μM) in the presence of cytokines, harvested over a 0, 6 and 10h time course, and RNA samples were prepared. Libraries were constructed and underwent 150 bp paired-end sequencing on an Illumina NextSeq500 sequencer.

QUANTIFICATION AND STATISTICAL ANALYSIS

Comparing dropout p values with gene expression

To investigate the relationship between dropout probability and gene expression, we used a scatter plot, as proposed by Tzelepis et al (Tzelepis et al., 2016). Dropout probability was obtained using MAGeCK software (Li et al., 2014). We used Salmon (Patro et al., 2017)

with default parameters for RNA-seq quantification, and averaged TPM values across three replicates per condition.

Quantification of RNA-Seq data and identification of differential splicing events

Raw sequence reads were aligned to mouse reference mm10 using the STAR aligner (Dobin et al., 2013). We assessed individual junction counts for each cohort by summing these counts over all samples in that cohort. Estimation of intron retention counts were performed by defining a 6 nt window over each exon-exon junction (3 nt in exon and 3 nt in intron), and all reads which completely overlapped this window were considered intronic (Darman et al., 2015). Individual junction usage in each cohort was then converted to a percent spliced in (PSI) measurement by dividing the count of that junction by the total count of all junctions sharing a splice site with it in that cohort. We assessed differential junction PSI between cohorts similar to (Darman et al., 2015), however a binomial z-test of logit-transformed PSI values was used to assess the likelihood of observation, and these p values were corrected using FDR. A q-value < 0.05 was considered significant.

Gene expression analysis

Isoform quantification was performed using Kallisto (Bray et al., 2016) against GENCODE M2. Total gene counts were compiled by summing isoform counts for all isoforms belonging of that gene. Differential analysis was performed using EdgeR (Robinson et al., 2010) to assess significance of gene expression changes between cohorts. Gene expression changes corresponding to a q-value < 0.05 were considered significant.

NMD prediction

Prediction of splice variant induction of nonsense mediated decay (NMD) was performed as previously described (Darman et al., 2015). Briefly, all RefSeq annotated transcripts for each gene containing a given significant splicing event were modified to contain the novel junction (i.e., intron retention events involve the removal of annotated junctions, alternative 3' splice sites cause the extension of downstream exons, etc). Each transcript was then translated in silico and it was determined whether a premature termination codon (PTC) had been inserted more than 50 nt from the last exon-exon junction, which is predicted to result in targeting by the NMD pathway (Rivas et al., 2015). If all transcripts associated with a given gene were predicted to result in degradation, that gene was predicted to be NMD-targeted. Similarly, if all transcripts were not predicted to contain a PTC that gene was predicted not to be targeted by the NMD pathway.

DATA AND SOFTWARE AVAILABILITY

RNA-seq data from CALM/AF10 AML or normal GMP cells were deposited in the Gene Expression Omnibus (GEO) database under accession number GSE107288.

SUPPLEMENTAL INFORMATION TITLES

Table S1. Summary of genome-wide screening results, Related to Figure1.

Table S2. Summary of in vivo second screening results, Related to Figure1.

Table S3. Summary of domain saturating mutagenesis results and a list of sgRNA sequences used, Related to Figure1.

Table S4. Summary of mass-spectrometry results, Related to Figure 3.

Table S5. Summary of RNA-seq results, Related to Figure 4.

KEY RESOURCES TABLE

The table highlights the genetically modified organisms and strains, cell lines, reagents, software, and source data **essential** to reproduce results presented in the manuscript. Depending on the nature of the study, this may include standard laboratory materials (i.e., food chow for metabolism studies), but the Table is **not** meant to be comprehensive list of all materials and resources used (e.g., essential chemicals such as SDS, sucrose, or standard culture media don't need to be listed in the Table). **Items in the Table must also be reported in the Method Details section within the context of their use.** The number of **primers and RNA sequences** that may be listed in the Table is restricted to no more than ten each. If there are more than ten primers or RNA sequences to report, please provide this information as a supplementary document and reference this file (e.g., See Table S1 for XX) in the Key Resources Table.

Please note that ALL references cited in the Key Resources Table must be included in the References list. Please report the information as follows:

- **REAGENT or RESOURCE:** Provide full descriptive name of the item so that it can be identified and linked with its description in the manuscript (e.g., provide version number for software, host source for antibody, strain name). In the Experimental Models section, please include all models used in the paper and describe each line/strain as: model organism: name used for strain/line in paper: genotype. (i.e., Mouse: OXTR$^{fl/fl}$: B6.129(SJL)-Oxtr$^{tm1.1Wsy/J}$). In the Biological Samples section, please list all samples obtained from commercial sources or biological repositories. Please note that software mentioned in the Methods Details or Data and Software Availability section needs to be also included in the table. See the sample Table at the end of this document for examples of how to report reagents.

- **SOURCE:** Report the company, manufacturer, or individual that provided the item or where the item can obtained (e.g., stock center or repository). For materials distributed by Addgene, please cite the article describing the plasmid and include "Addgene" as part of the identifier. If an item is from another lab, please include the name of the principal investigator and a citation if it has been previously published. If the material is being reported for the first time in the current paper, please indicate as "this paper." For software, please provide the company name if it is commercially available or cite the paper in which it has been initially described.

- **IDENTIFIER:** Include catalog numbers (entered in the column as "Cat#" followed by the number, e.g., Cat#3879S). Where available, please include unique entities such as RRIDs, Model Organism Database numbers, accession numbers, and PDB or CAS IDs. For antibodies, if applicable and available, please also include the lot number or clone identity. For software or data resources, please include the URL where the resource can be downloaded. Please ensure accuracy of the identifiers, as they are essential for generation of hyperlinks to external sources when available. Please see the Elsevier list of Data Repositories with automated bidirectional linking for details. When listing more than one identifier for the same item, use semicolons to separate them (e.g. Cat#3879S; RRID: AB_2255011). If an identifier is not available, please enter "N/A" in the column.
    - ***A NOTE ABOUT RRIDs:*** We highly recommend using RRIDs as the identifier (in particular for antibodies and organisms, but also for software tools and databases). For more details on how to obtain or generate an RRID for existing or newly generated resources, please visit the RII or search for RRIDs.

Please use the empty table that follows to organize the information in the sections defined by the subheading, skipping sections not relevant to your study. Please do not add subheadings. To add a row, place the cursor at the end of the row above where you would like to add the row, just outside the right border of the table. Then press the ENTER key to add the row. Please delete empty rows. Each entry must be on a separate row; do not list multiple items in a single table cell. Please see the sample table at the end of this document for examples of how reagents should be cited.

TABLE FOR AUTHOR TO COMPLETE

Please upload the completed table as a separate document. ***Please do not add subheadings to the Key Resources Table.*** *If you wish to make an entry that does not fall into one of the subheadings below, please contact your handling editor. (****NOTE:*** *For authors publishing in Current Biology, please note that references within the KRT should be in numbered style, rather than Harvard.)*

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
FITC-Annexin V	BioLegend	Cat#: 640905 RRID: AB_2561291
FITC anti-human CD15	BioLegend	Cat#: 301904 RRID: AB_314196
FITC anti-human CD33	BioLegend	Cat#: 303304 RRID: AB_314344
FITC anti-human CD34	BioLegend	Cat#: 343604 RRID: AB_1732005
PE anti-human CD11b	BioLegend	Cat#: 301306 RRID: AB_314158
PE anti-human CD38	BioLegend	Cat#: 303304 RRID: AB_314344
PE-Cy7 anti-human CD123	BioLegend	Cat#: 306010 RRID: AB_493576
APC anti-human CD45	BioLegend	Cat#: 304012 RRID: AB_314400
APC-Cy7 anti-human CD14	BioLegend	Cat#: 303506 RRID: AB_314358
APC-Cy7 anti-human CD19	BioLegend	Cat#: 302218 RRID: AB_314248
APC-Cy7 anti-human CD45RA	BioLegend	Cat#: 304128 RRID: AB_10708880
Pacific Blue anti-mouse CD45	BioLegend	Cat#: 103126 RRID: AB_493535
BV421 anti-mouse Gr-1	BioLegend	Cat#: 108433 RRID: AB_10900232
PE anti-mouse CD11b	BioLegend	Cat#: 101207 RRID: AB_312790
Rabbit polyclonal anti-DCPS	Bethyl	Cat#: A305-426A
Mouse monoclonal anti-DCPS	Santa Cruz	Cat#: sc-393226
Rabbit polyclonal anti-FHIT	ThermoFisher	Cat#: 71-9000 RRID: AB_2534004
Bacterial and Virus Strains		
shRNA Human DCPS	Sigma Aldrich	TRCN0000335923 TRCN0000335924 TRCN0000005569
Biological Samples		
Patient-derived xenografts (PDX)	Leukemia/Lymphoma Xenograft Core (LLX)	https://www.proxe.org
Bone marrow sections	Kyushu University Hospital	N/A

Umbilical Cord Blood	Kyushu Block Red Cross Blood Center, Japan Red Cross Society	N/A
Chemicals, Peptides, and Recombinant Proteins		
Mouse SCF	PeproTech	Cat#: 250-03
Mouse IL-3	PeproTech	Cat#: 213-13
Mouse IL-6	PeproTech	Cat#: 200-06
Nutlin-3a	Selleck	S8059
RG3039	MedKoo Biosciences	Cat#: 510322
Critical Commercial Assays		
Click-iT Plus EdU Alexa Fluor 647 Flow Cytometry Assay Kit	Life Technologies	Cat#: C10634
Ovation Mouse RNA-Seq System	NuGEN	Cat#: #0348-32
Deposited Data		
RNA-seq Raw Data	This paper	GSE107288
Western blot, cytospin, IHC Raw Data	This paper	https://data.mendeley.com/datasets/55pj8frskf/draft?a=0e3ae5c0-94e6-469a-94ee-df2bede78275
Experimental Models: Cell Lines		
Human: MOLM-13	DSMZ	Cat#: ACC-554 RRID: CVCL_2119
Human: OCI-AML3	DSMZ	Cat#: ACC-582 RRID: CVCL_1844
Human: MV4-11	ATCC	Cat#: CRL-9591 RRID: CVCL_0064
Human: THP-1	ATCC	Cat#: TIB-202 RRID: CVCL_0006
Human: HEK293T	ATCC	CRL-3216 RRID: CVCL_0063
Experimental Models: Organisms/Strains		
Mouse: C57BL/6	Jackson Laboratory	Cat#: 000664 RRID: IMSR_JAX:000664
Mouse: C57BL/6.*Rag2*^{null}*Il2rg*^{null}NOD-*Sirpa* (BRGS)	Yamauchi et al., 2013	N/A
Oligonucleotides		
sgRNAs sequence for 2nd screen, see Table S2	This paper	N/A
sgRNAs sequence for domain saturating mutagenesis, see Table S3	This paper	N/A
sgRNA targeting sequence: Trp53: AACAGATCGTCCATGCAGTG	This paper	N/A
Recombinant DNA		
lentiCas9-Blast	Sanjana et al., 2014	Addgene Plasmid #52962
psPAX2	N/A	Addgene Plasmid #12260
pMD2.G	N/A	Addgene Plasmid #12259
GeCKO v2 mouse library	Sanjana et al., 2014	Addgene Cat#: 1000000053

lentiGuide-Puro	Sanjana et al., 2014	Addgene Plasmid#: 52963
Software and Algorithms		
MAGeCK	Li et al., 2014	https://sourceforge.net/projects/mageck/
STAR aligner	Dobin et al., 2013	https://github.com/alexdobin/STAR
Kallisto	Bray et al., 2016	https://pachterlab.github.io/kallisto/
EdgeR	Robinson et al., 2010	http://bioconductor.org/packages/release/bioc/html/edgeR.html
Prism 5	GraphPad	https://www.graphpad.com
FlowJo V9.7.6	Tree Star	https://www.flowjo.com/
Other		

Figure 1

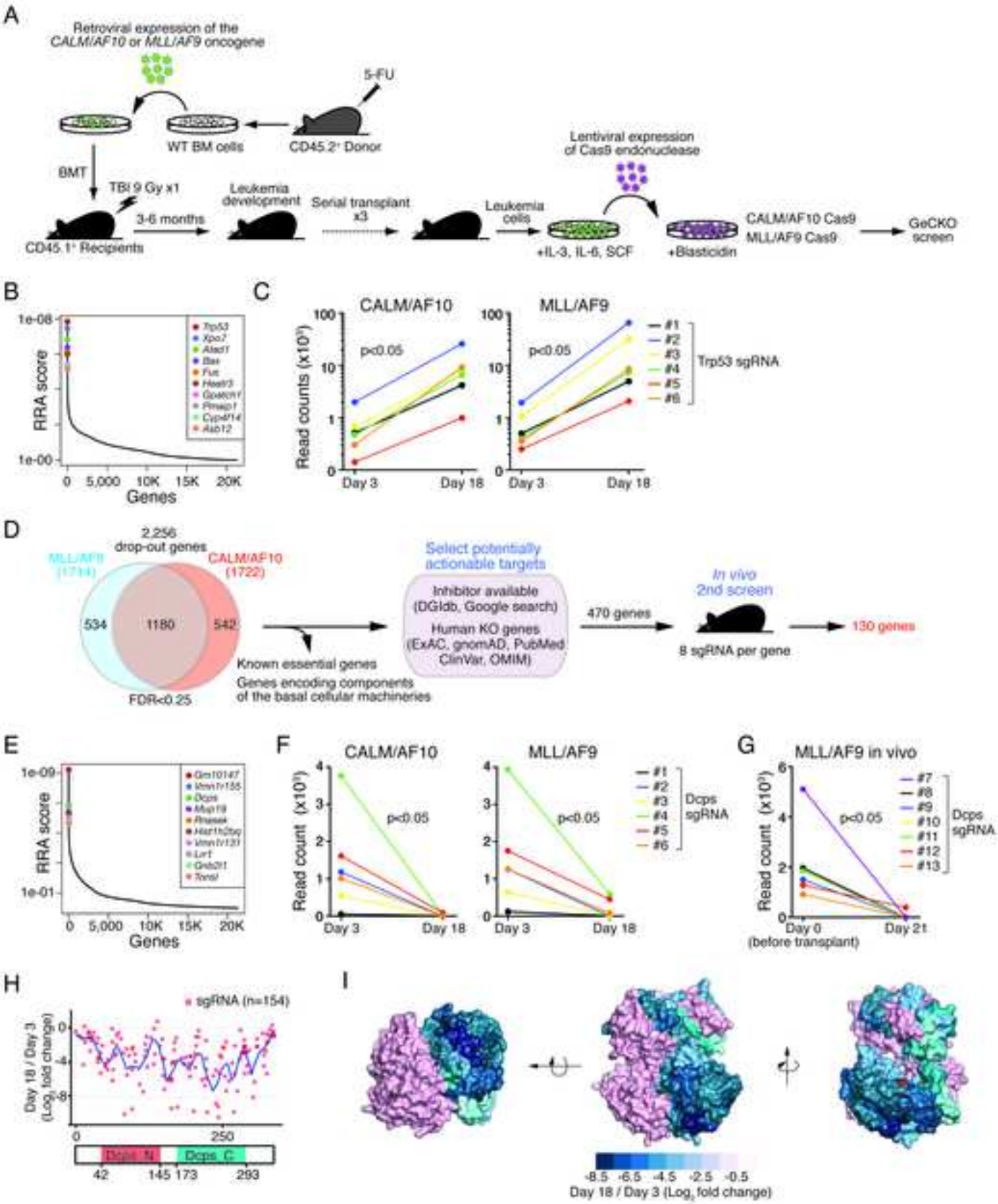


Figure 2

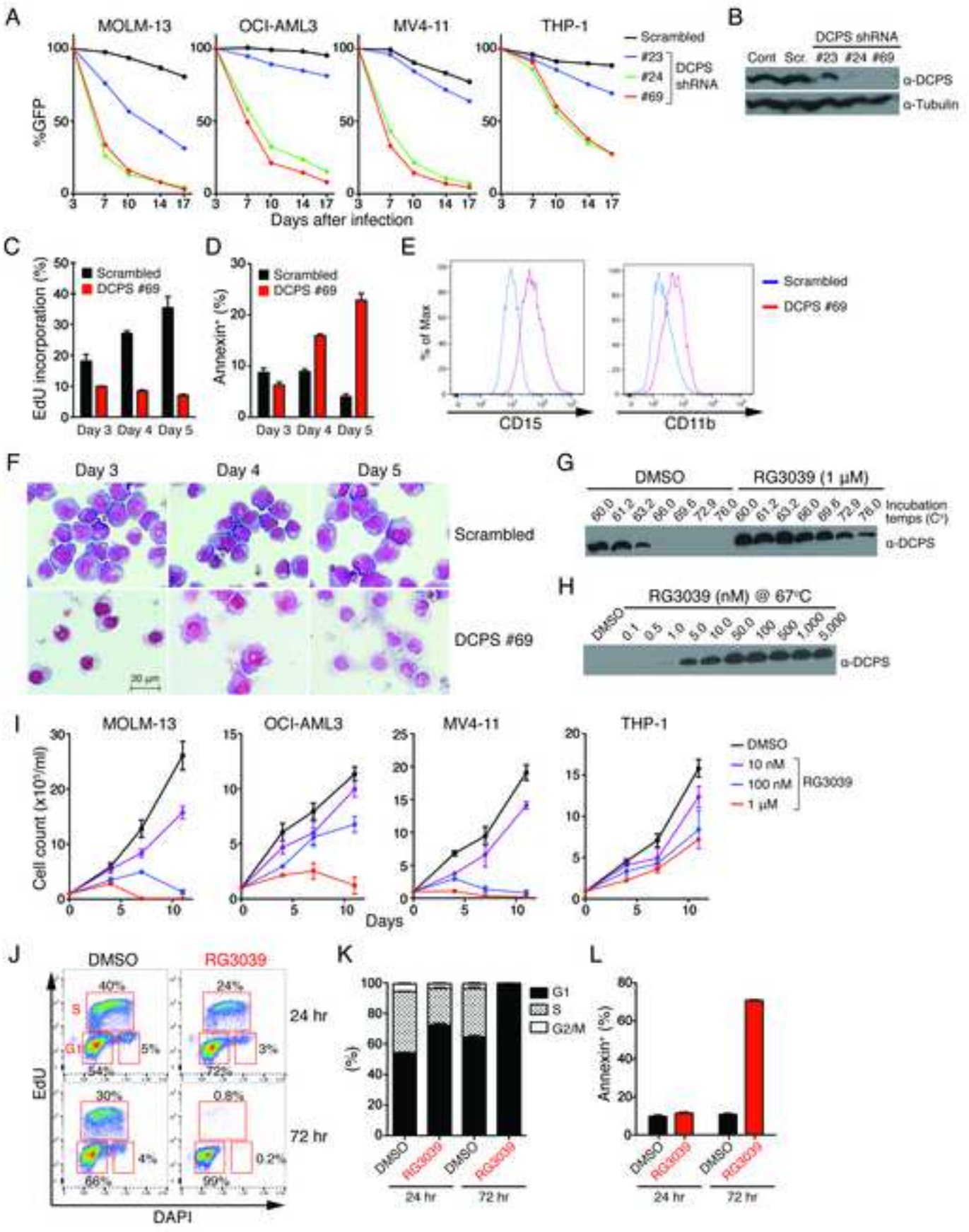


Figure 3

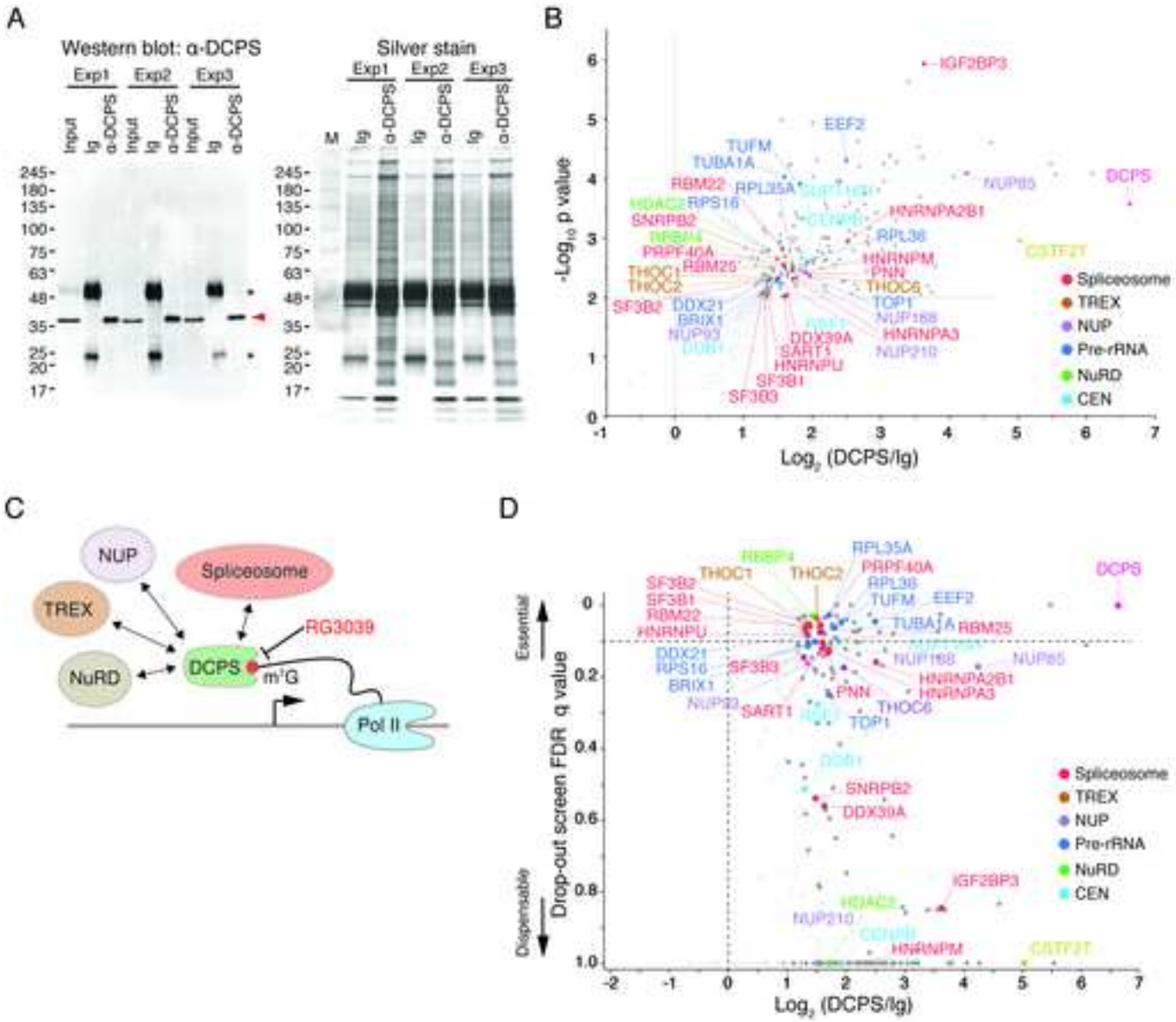

A
Western blot: α-DCPS
Exp1
Exp2
Exp3
Input
Ig
α-DCPS
Silver stain
M
245
180
135
100
75
63
48
35
25
20
17
B
-Log10 p value
Log2 (DCPS/Ig)
IGF2BP3
EEF2
TUFM
TUBA1A
RBM22
RPL35A
NUP85
DCPS
HDAC2
RPS16
HNRNPA2B1
SNRPB2
RPL36
CSTF2T
PRPF40A
HNRNPM
RBM25
PNN
THOC1
THOC6
THOC2
TOP1
SF3B2
DDX21
NUP188
BRIX1
HNRNPA3
DDX39A
SART1
NUP210
HNRNPU
SF3B1
SF3B3
Spliceosome
TREX
NUP
Pre-rRNA
NuRD
CEN
C
NUP
Spliceosome
TREX
RG3039
DCPS
NuRD
m7G
Pol II
D
Drop-out screen FDR q value
Essential
Dispensable
Log2 (DCPS/Ig)
HEXIM1
RPL35A
THOC1
THOC2
PRPF40A
DCPS
SF3B2
RPL36
SF3B1
TUFM
EEF2
RBM22
TUBA1A
RBM25
HNRNPU
DDX21
NUP188
NUP85
RPS16
SF3B3
HNRNPA2B1
BRIX1
HNRNPA3
PNN
NUP93
SART1
THOC6
TOP1
SNRPB2
DDX39A
IGF2BP3
HDAC2
NUP210
HNRNPM
CSTF2T
Spliceosome
TREX
NUP
Pre-rRNA
NuRD
CEN

Figure 4

Click here to download Figure Main Figures-04.tif

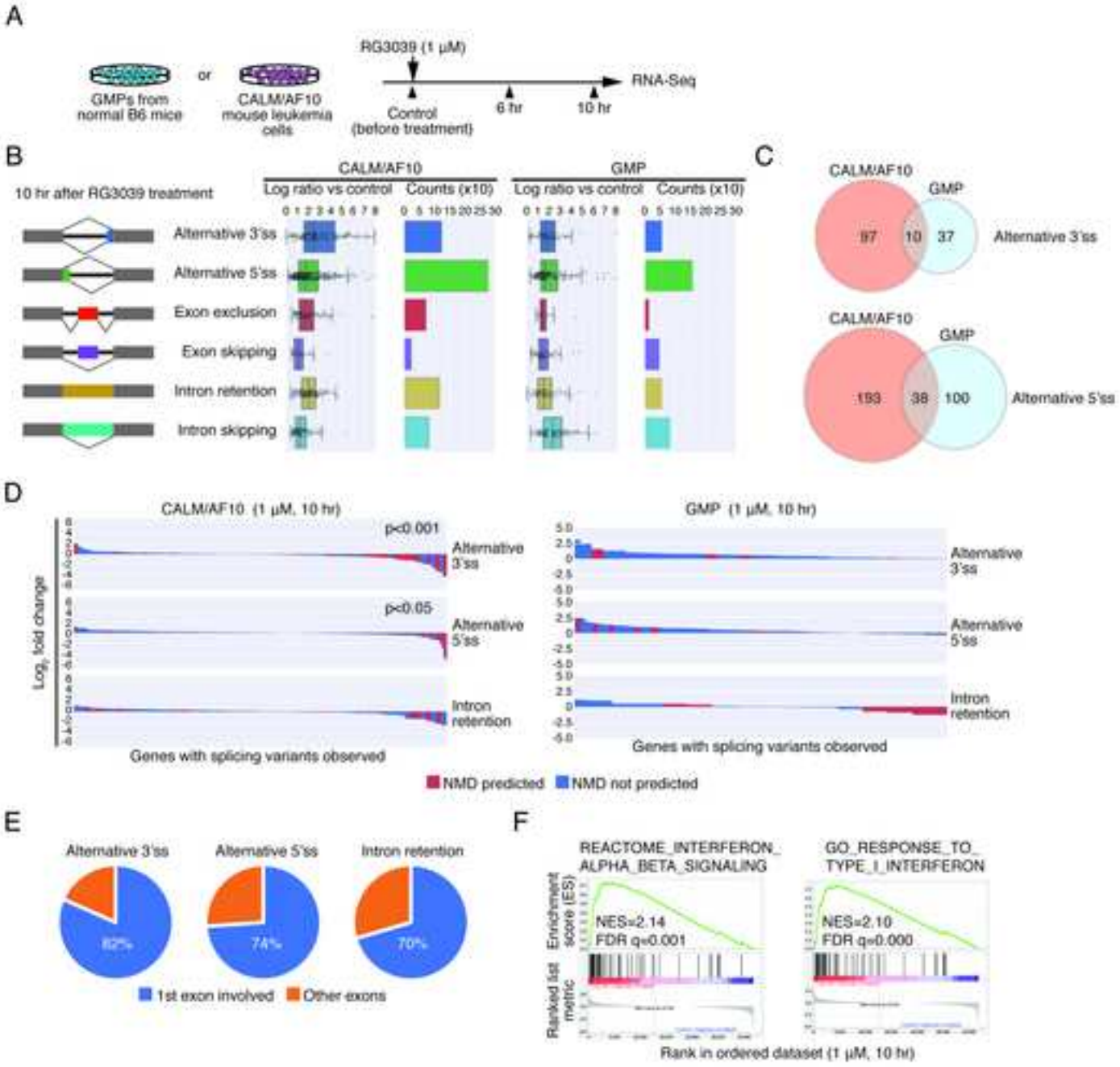
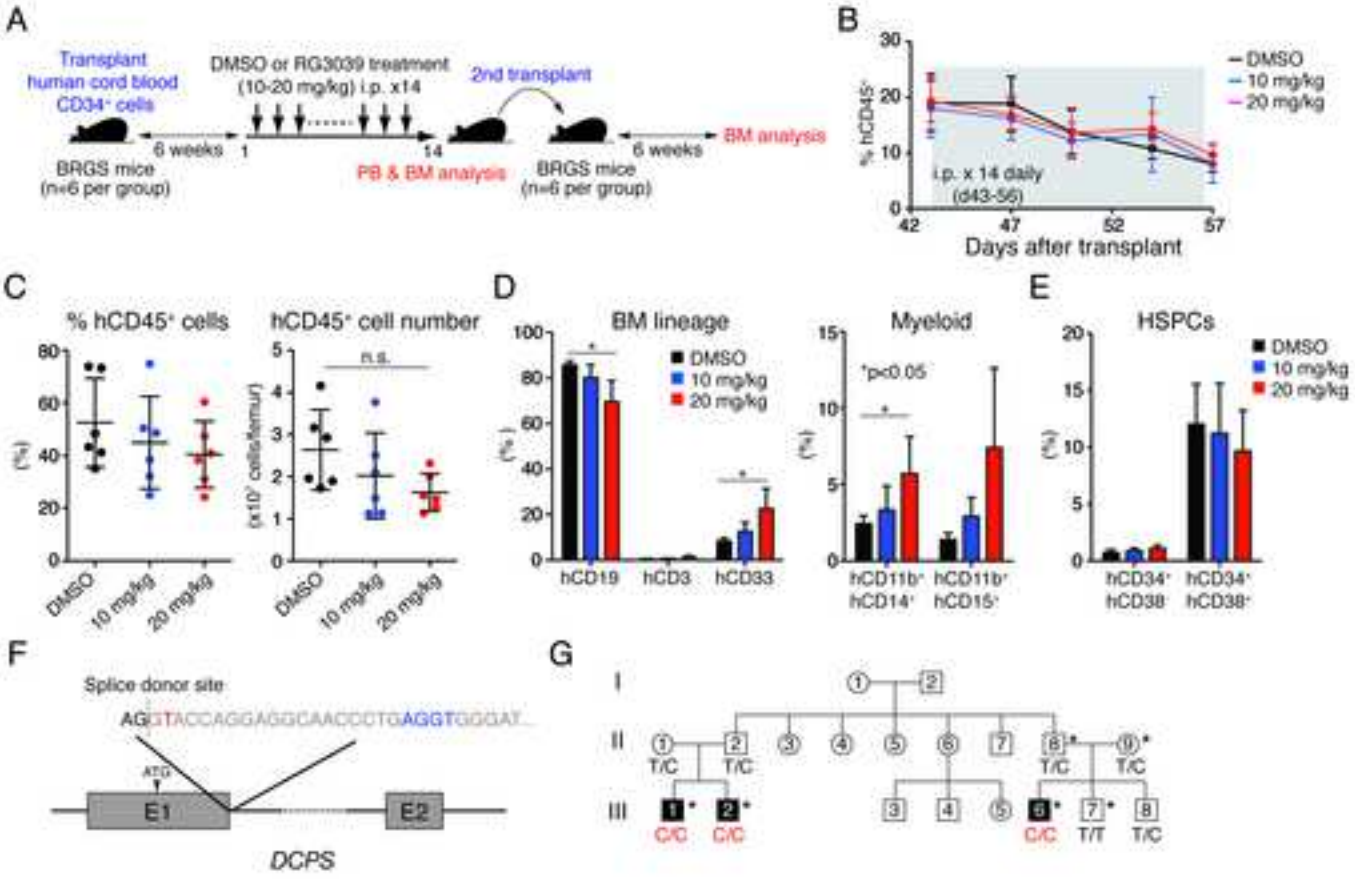


Figure 5

Click here to download Figure Main Figures_final-05.tif

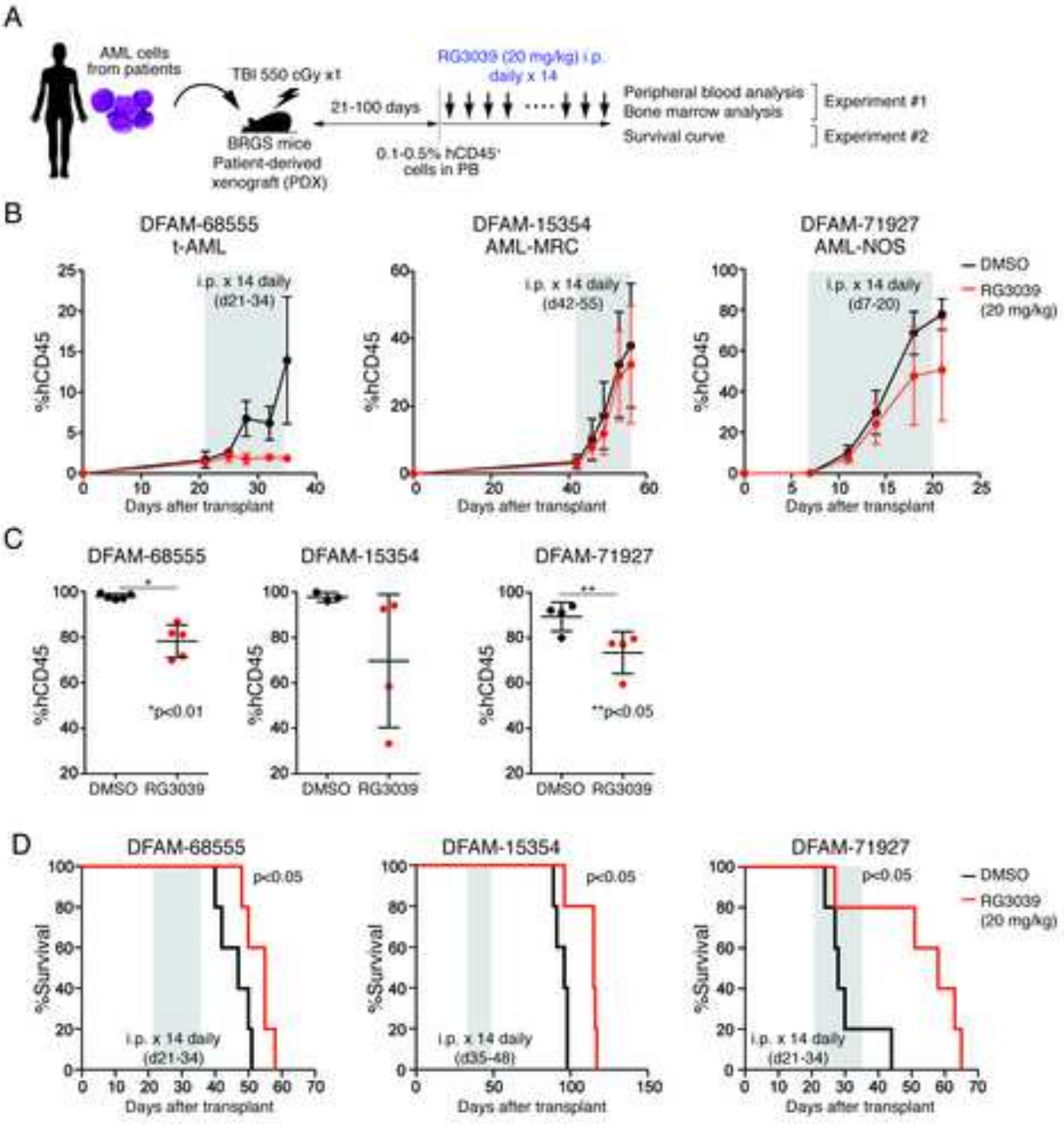
Figure 5



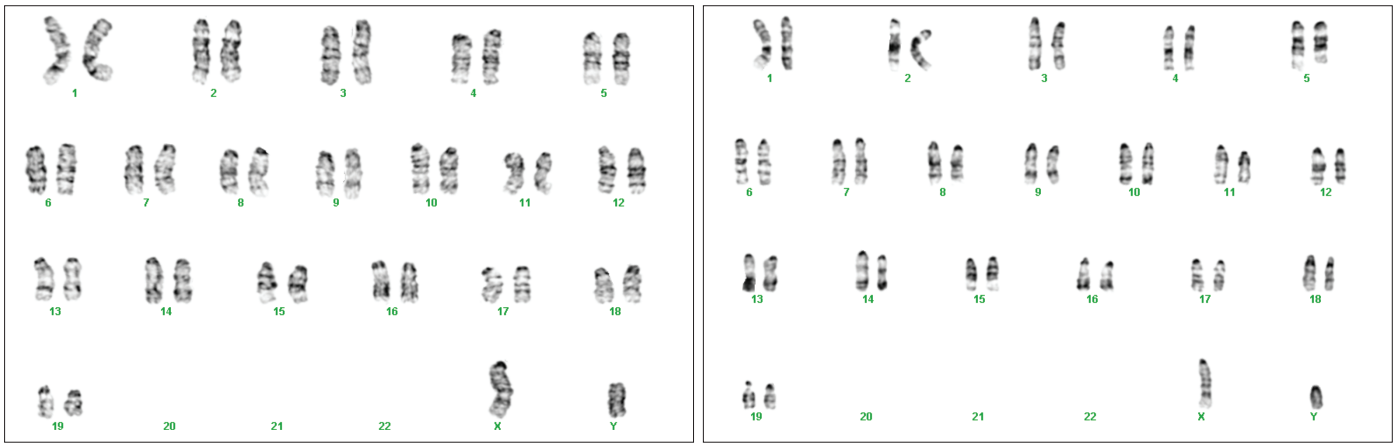
H

		III-1	III-2	III-6	III-7	II-8	II-9
Gender, Age		Male, 11 y.o.	Male, 9 y.o.	Male, 5 y.o.	Male, 9 y.o.	Male, 39 y.o.	Female, 36 y.o.
WBC	(10^3/μl)	10700	9200	6000	8750	8510	5800
Neut	(%)	68.5	61.5	50.6	48.8	53.8	54.3
Lymp	(%)	15.1	27.7	46.3	37.3	34.7	34.1
Mono	(%)	15.9	8.2	0.2	7.9	10.2	9.7
Eosino	(%)	0.2	2.4	N/A	5.7	1.1	1.6
Baso	(%)	0.3	0.2	N/A	0.3	0.2	0.3
RBC	(10^6/μl)	394	455	464	529	542	463
Hb	(g/dl)	11.6	12.0	11.8	14.5	14.9	11.9
Hct	(%)	33.2	35.7	35.1	40.5	43	36.2
MCV	(fL)	84.4	78.5	75.6	76.6	79.3	78.2
Platelet	(10^4/μl)	22.2	48.2	18.1	41.3	31.3	30.5

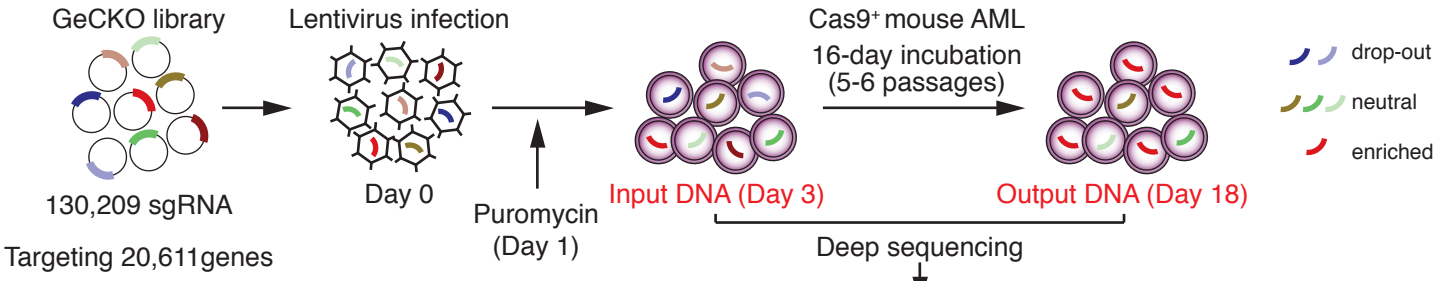
Figure 6



A



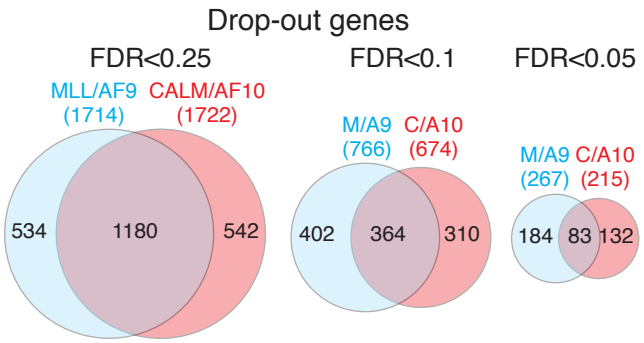
B



C

Sample	Mapped reads	Library sgRNA	Zero counts	Gini Index	Coverage
CALM-AF10_LibraryA_Day3	41,606,471	67,405	773	0.11	x617
CALM-AF10_LibraryA_Day18	40,699,845	67,405	6,006	0.2	x603
CALM-AF10_LibraryB_Day3	42,997,399	62,804	419	0.11	x685
CALM-AF10_LibraryB_Day18	44,087,417	62,804	6,641	0.21	x702
MLL-AF9_LibraryA_Day3	43,433,781	67,405	734	0.1	x644
MLL-AF9_LibraryA_Day18	50,203,365	67,405	6,711	0.21	x745
MLL-AF9_LibraryB_Day3	44,561,317	62,804	1,002	0.11	x710
MLL-AF9_LibraryB_Day18	44,572,208	62,804	5,614	0.2	x710

D



E

MLL/AF9

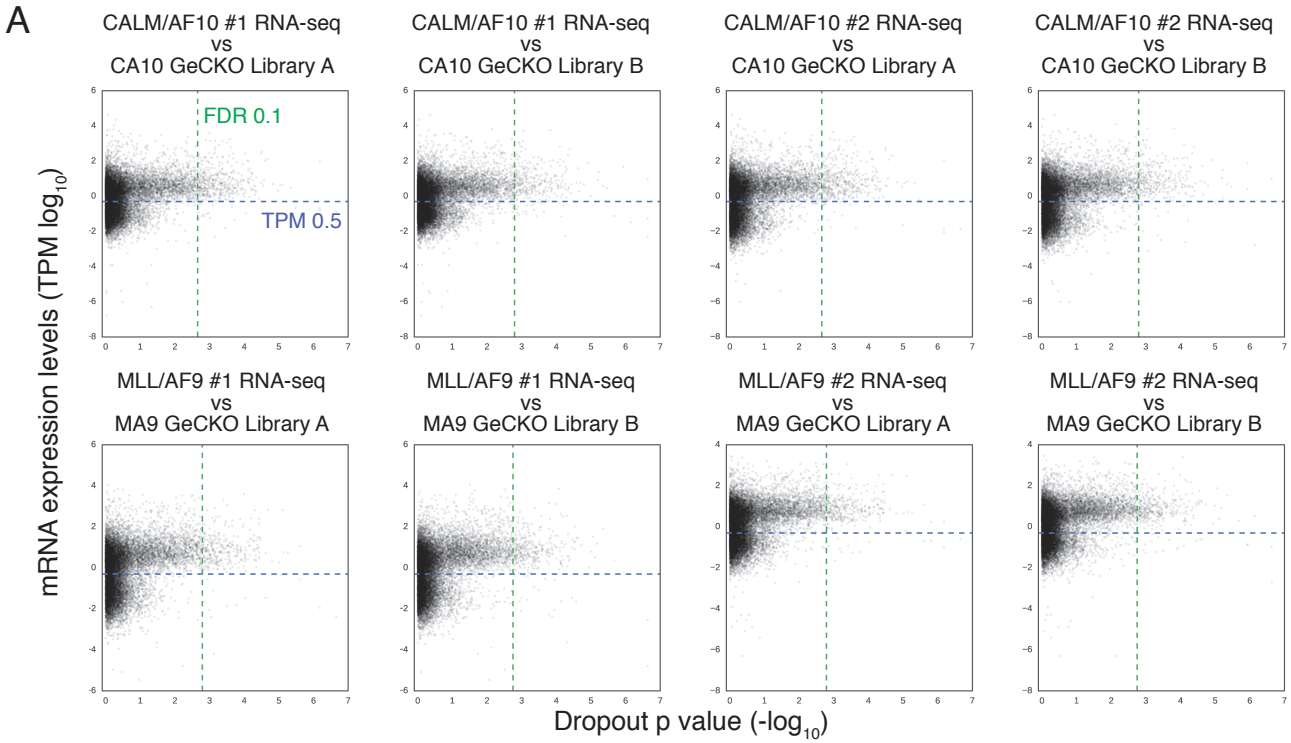
Node name (KEGG pathway)	p value
HSA03010_RIBOSOME	2.44E-17
HSA00970_AMINOACYL_TRNA_BIOSYNTHESIS	3.50E-15
HSA00240_PYRIMIDINE_METABOLISM	9.28E-11
HSA03020_RNA_POLYMERASE	2.78E-10
HSA03050_PROTEASOME	2.66E-08
HSA00190_OXIDATIVE_PHOSPHORYLATION	8.14E-08
HSA00230_PURINE_METABOLISM	1.08E-06

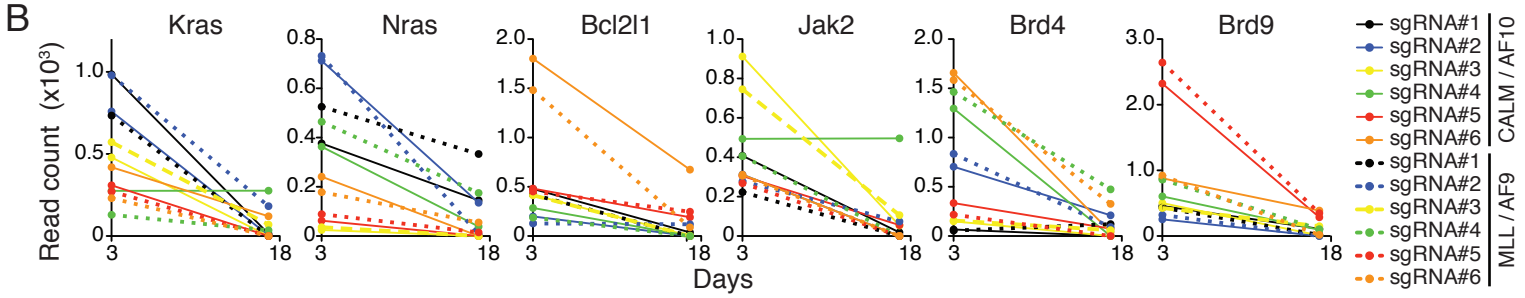
CALM/AF10

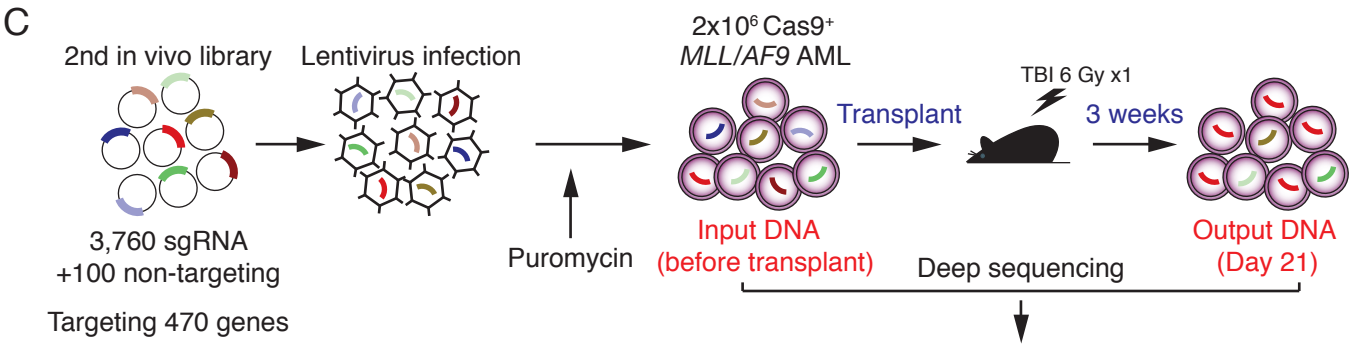
Node name (KEGG pathway)	p value
HSA03010_RIBOSOME	4.69E-21
HSA00190_OXIDATIVE_PHOSPHORYLATION	4.22E-13
HSA00970_AMINOACYL_TRNA_BIOSYNTHESIS	6.79E-12
HSA03050_PROTEASOME	9.19E-09
HSA03020_RNA_POLYMERASE	1.47E-08
HSA00240_PYRIMIDINE_METABOLISM	3.13E-08
HSA04110_CELL_CYCLE	3.39E-06
HSA04120_UBIQUITIN_MEDIATED_PROTEOLYSIS	6.31E-05

Figure S1. Related to Figure 1. CRISPR-Cas9 screens in mouse AML cell lines with normal karyotypes.

(A) Representative results of karyotype analysis. **(B)** Schematic representations of the GeCKO screen. **(C)** Sequencing read counts of each sample. Read mapping and gini index calculations were performed using the MAGeCK count program. **(D)** Venn diagrams show numbers of dropout genes at indicated FDR values. **(E)** Pathway analysis of dropout genes shows enrichment of signatures representing basal cellular machineries.

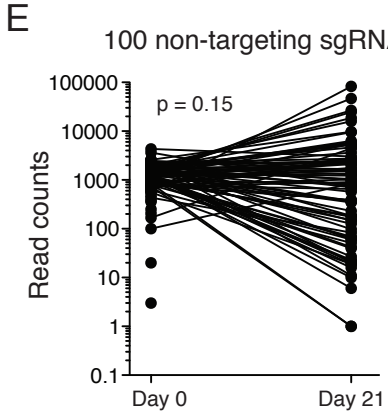






D

Sample	Mapped reads	Library sgRNA	Zero counts	Gini Index	Coverage
MLL-AF9 before transplant	4,966,871	3,860	11	0.05	x1287
MLL-AF9 3wks after transplant	3,235,660	3,860	2,359	0.76	x838



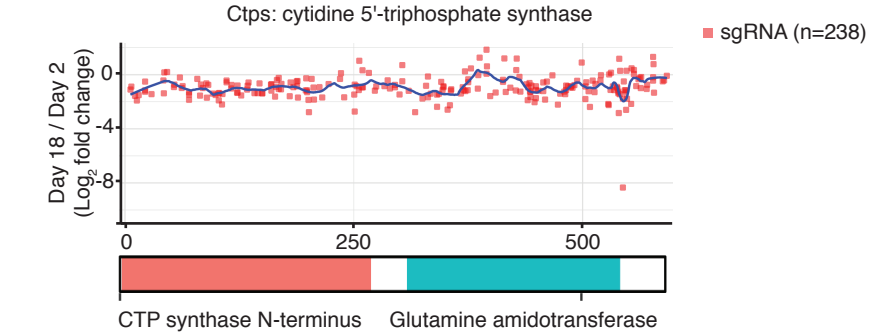


Figure S2. Dropout genes and validation screens: additional information. Related to Figure 1.

(A) RNA-seq experiments were performed using 4 primary mouse AML lines established independently in vivo (CALM/AF10 #1, CALM/AF10 #2, MLL/AF9 #1 and MLL/AF9 #2). mRNA expression levels (y-axis: TPM values) and the dropout p value (x axis: minus $\log_{10}$ values in GeCKO screens) of each gene were plotted. **(B)** Graphs show read counts of individual sgRNAs targeting well-known genes essential for AML, before and after a 16-day incubation. **(C)** Workflow of second screen in vivo. Cas9-expressing MLL/AF9 cells were transduced with a pool of sgRNAs targeting 470 selected genes (8 sgRNA per gene) and 100 non-targeting sgRNAs and then transferred to sub-lethally irradiated recipients. Change in abundance of each sgRNA between the initial and final cell population was analyzed using the MAGeCK program. **(D)** Sequencing results of DNA samples obtained before and after transplant. **(E)** Read counts of each non-targeting sgRNA before and after transplant were plotted. **(F)** Results of CTPS domain-mapping.

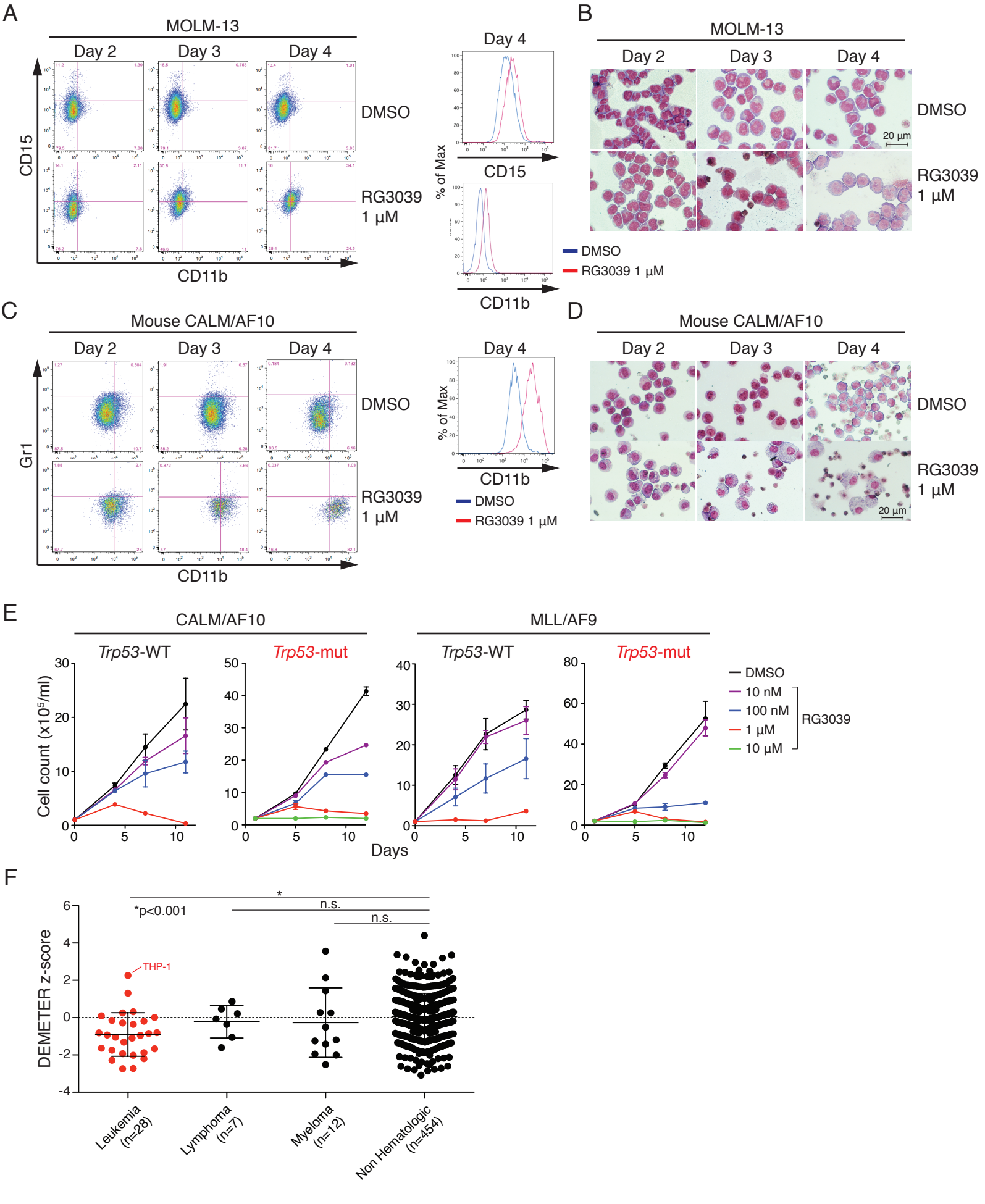
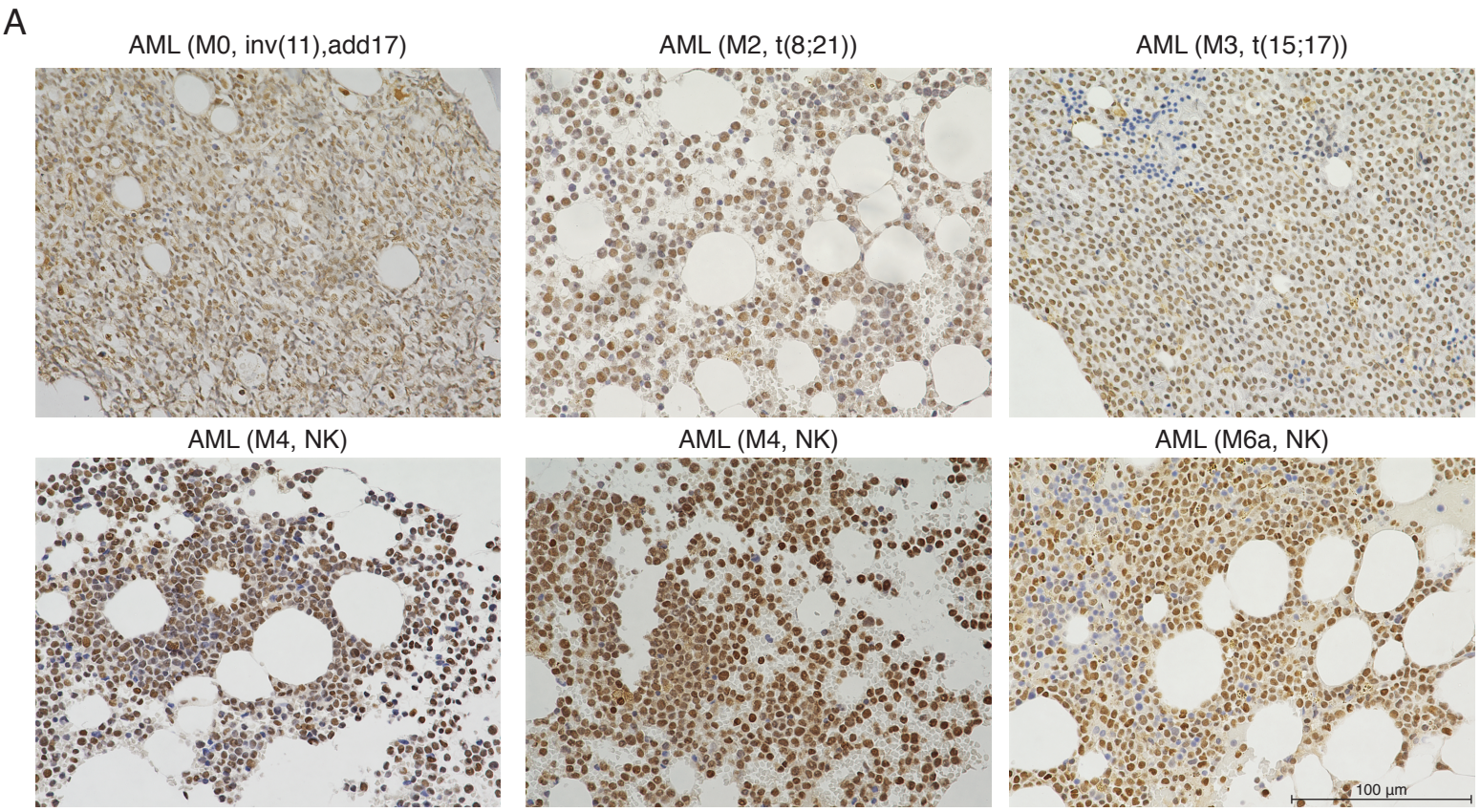
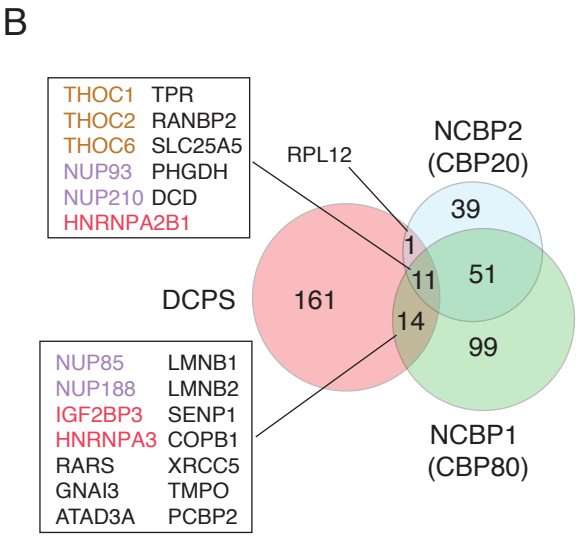
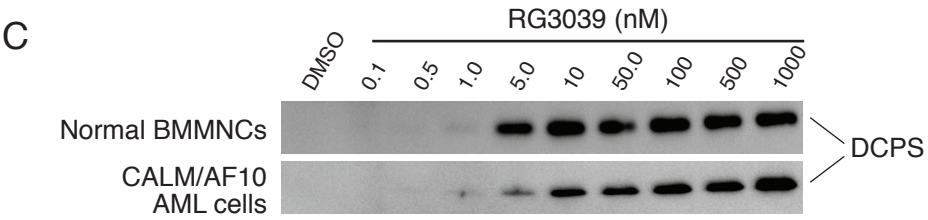
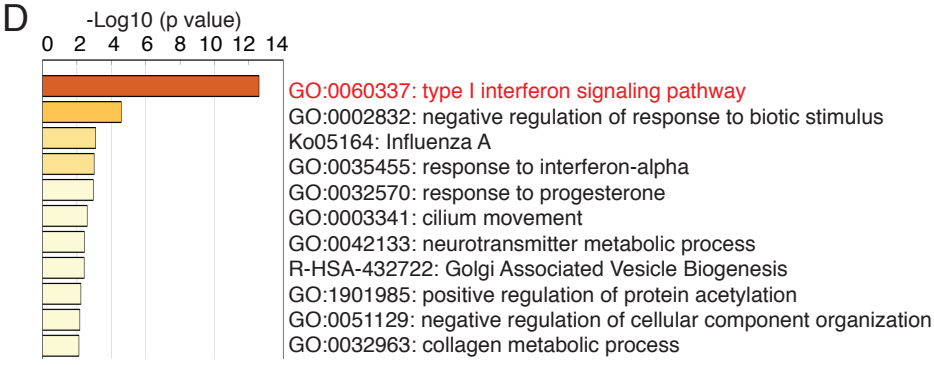

A
MOLM-13
Day 2
Day 3
Day 4
CD15
CD11b
DMSO
RG3039
1 µM
Day 4
% of Max
CD15
CD11b
DMSO
RG3039 1 µM
B
MOLM-13
Day 2
Day 3
Day 4
DMSO
RG3039
1 µM
20 µm
C
Mouse CALM/AF10
Day 2
Day 3
Day 4
Gr1
CD11b
DMSO
RG3039
1 µM
Day 4
% of Max
CD11b
DMSO
RG3039 1 µM
D
Mouse CALM/AF10
Day 2
Day 3
Day 4
DMSO
RG3039
1 µM
20 µm
E
CALM/AF10
MLL/AF9
Trp53-WT
Trp53-mut
Trp53-WT
Trp53-mut
Cell count (x10⁵/ml)
Days
DMSO
10 nM
100 nM
1 µM
10 µM
RG3039
F
*p<0.001
*
n.s.
n.s.
DEMETER z-score
THP-1
Leukemia (n=28)
Lymphoma (n=7)
Myeloma (n=12)
Non Hematologic (n=454)

Figure S3. RG3039 slows proliferation and induces AML cell differentiation via p53-independent mechanisms. Related to Figure 2. (A) Representative FACS plots of MOLM-13 cells treated with either DMSO (vehicle) or RG3039. **(B)** May-Giemsa stain of MOLM-13 cells treated with either DMSO or RG3039. **(C)** Representative FACS plots of mouse CALM/AF10 cells treated with either DMSO or RG3039. **(D)** May-Giemsa stain of CALM/AF10 cells. **(E)** *Trp53*-WT or mutant mouse AML cells were treated with RG3039, and growth curves generated as described in Figure 2I. Data are represented as means ± SD. **(F)** Dot graphs show DCPS dependency of 501 cancer cell lines representing leukemia, lymphoma, multiple myeloma and non-hematologic tumors. Data are shown as means ± SD. The Y-axis represents DEMETER z-score of shRNA targeting DCPS (https://depmap.org/rnai/genedeps?gene=DCPS).


A
AML (M0, inv(11),add17)
AML (M2, t(8;21))
AML (M3, t(15;17))
AML (M4, NK)
AML (M4, NK)
AML (M6a, NK)
100 µm


B
THOC1 TPR
THOC2 RANBP2
THOC6 SLC25A5
NUP93 PHGDH
NUP210 DCD
HNRNPA2B1
RPL12
NCBP2
(CBP20)
39
1
11
51
DCPS
161
14
99
NCBP1
(CBP80)
NUP85 LMNB1
NUP188 LMNB2
IGF2BP3 SENP1
HNRNPA3 COPB1
RARS XRCC5
GNAI3 TMPO
ATAD3A PCBP2


C
RG3039 (nM)
DMSO
0.1
0.5
1.0
5.0
10
50.0
100
500
1000
Normal BMMNCs
CALM/AF10
AML cells
DCPS


D
-Log10 (p value)
0 2 4 6 8 10 12 14
GO:0060337: type I interferon signaling pathway
GO:0002832: negative regulation of response to biotic stimulus
Ko05164: Influenza A
GO:0035455: response to interferon-alpha
GO:0032570: response to progesterone
GO:0003341: cilium movement
GO:0042133: neurotransmitter metabolic process
R-HSA-432722: Golgi Associated Vesicle Biogenesis
GO:1901985: positive regulation of protein acetylation
GO:0051129: negative regulation of cellular component organization
GO:0032963: collagen metabolic process

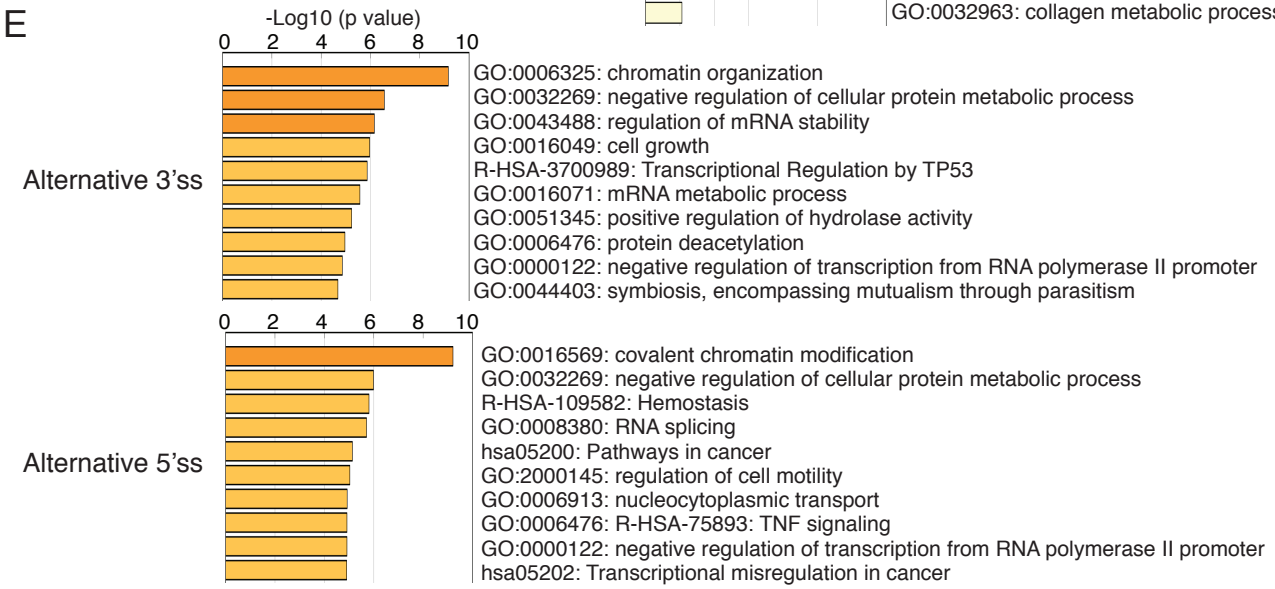

E
-Log10 (p value)
0 2 4 6 8 10
Alternative 3'ss
GO:0006325: chromatin organization
GO:0032269: negative regulation of cellular protein metabolic process
GO:0043488: regulation of mRNA stability
GO:0016049: cell growth
R-HSA-3700989: Transcriptional Regulation by TP53
GO:0016071: mRNA metabolic process
GO:0051345: positive regulation of hydrolase activity
GO:0006476: protein deacetylation
GO:0000122: negative regulation of transcription from RNA polymerase II promoter
GO:0044403: symbiosis, encompassing mutualism through parasitism
0 2 4 6 8 10
Alternative 5'ss
GO:0016569: covalent chromatin modification
GO:0032269: negative regulation of cellular protein metabolic process
R-HSA-109582: Hemostasis
GO:0008380: RNA splicing
hsa05200: Pathways in cancer
GO:2000145: regulation of cell motility
GO:0006913: nucleocytoplasmic transport
GO:0006476: R-HSA-75893: TNF signaling
GO:0000122: negative regulation of transcription from RNA polymerase II promoter
hsa05202: Transcriptional misregulation in cancer

Figure S4. DCPS protein localizes to the nucleus and associates with factors involved in pre-mRNA metabolic pathways. Related to Figures 3 and 4. (A) Immunohistochemistry (IHC) for DCPS protein in BM sections of six human AML cases. Shown are FAB subtype and karyotype in each case. Brown staining: DCPS protein. DCPS is predominantly nuclear in AML cells. NK: normal karyotype. **(B)** CBC, which consists of NCBP1 and NCBP2, and DCPS share some interacting proteins. NCBP1 and NCBP2 binding proteins were described previously (Andersen et al., 2013). **(C)** Affinity between RG3039 and DCPS was assessed by CETSA. **(D)** Results of gene expression signature analysis of RG3039-treated AML cells using the Metascape program (http://metascape.org/). **(E)** Metascape analysis of aberrantly-spliced genes following RG3039 treatment of CALM/AF10 AML cells.

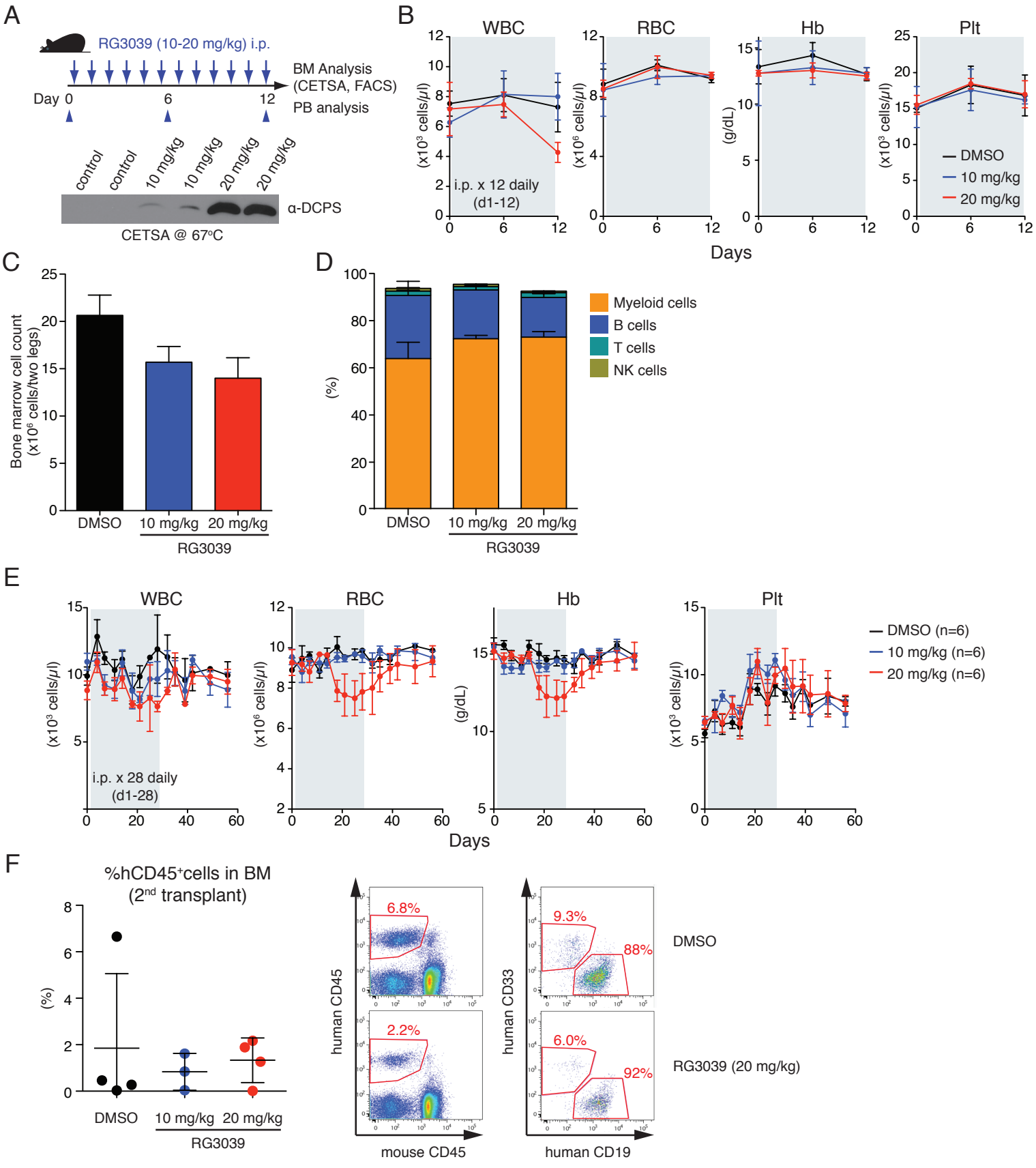
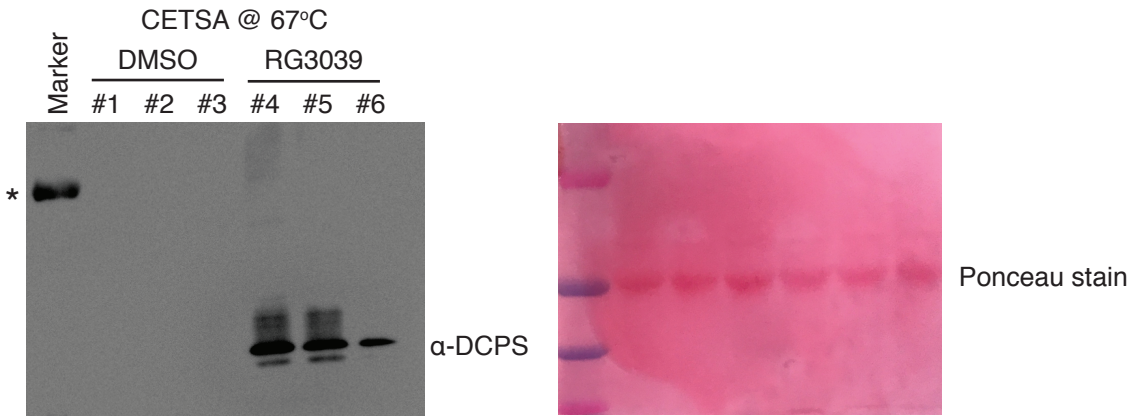

A
RG3039 (10-20 mg/kg) i.p.
BM Analysis
(CETSA, FACS)
PB analysis
Day 0
6
12
control
control
10 mg/kg
10 mg/kg
20 mg/kg
20 mg/kg
α-DCPS
CETSA @ 67°C
B
WBC
RBC
Hb
Plt
(x10³ cells/μl)
(x10⁶ cells/μl)
(g/dL)
i.p. x 12 daily
(d1-12)
DMSO
10 mg/kg
20 mg/kg
Days
C
Bone marrow cell count
(x10⁶ cells/two legs)
DMSO
10 mg/kg
20 mg/kg
RG3039
D
(%)
Myeloid cells
B cells
T cells
NK cells
DMSO
10 mg/kg
20 mg/kg
RG3039
E
WBC
RBC
Hb
Plt
(x10³ cells/μl)
(x10⁶ cells/μl)
(g/dL)
i.p. x 28 daily
(d1-28)
Days
DMSO (n=6)
10 mg/kg (n=6)
20 mg/kg (n=6)
F
%hCD45⁺cells in BM
(2nd transplant)
(%)
DMSO
10 mg/kg
20 mg/kg
RG3039
human CD45
mouse CD45
human CD33
human CD19
6.8%
2.2%
9.3%
88%
6.0%
92%
DMSO
RG3039 (20 mg/kg)

Figure S5. Effects of DCPS inhibition in vivo. Related to Figure 5. (A) DMSO or RG3039 (10 mg/kg or 20 mg/kg) was administered intraperitoneally daily for 12 days, and peripheral blood counts were analyzed before and after treatment (day 6 and day 12). CETSA was performed using BM cells the day after the last RG3039 (or DMSO) injection. **(B)** Peripheral blood counts were analyzed by a hematology analyzer before and after RG3039 treatment. Data are represented as means ± SD. WBC: white blood cell; RBC: red blood cell; Hb: hemoglobin; Plt: platelet. **(C)** BM cell counts were measured the day after the last RG3039 (or DMSO) injection. Data are represented as means ± SD. **(D)** Proportions of myeloid, B, T and NK cells in BM mononuclear cells were analyzed by FACS. Data are represented as means ± SD. **(E)** DMSO (vehicle) or RG3039 (10 mg/kg or 20 mg/kg) was administered intraperitoneally daily for 28 days, and peripheral blood counts were analyzed every 5-7 days. Treatment periods are depicted in gray. Data are represented as means ± SD. **(F)** Engraftment of human hematopoietic cells upon a second transplant. The Y-axis shows proportions of $hCD45^+$ cells in BM 6 weeks after the second transplant. Data are represented as means ± SD. Representative BM FACS plots are shown.

A

PDX	DFAM-68555-V1	DFAM-15354-V2	DFAM-71927-V3
WHO classification	Therapy-related AML	AML-MRC	AML-NOS
FAB classification	M5a	N/A	N/A
Karyotype	46,XX,t(9;11)(p22;q23)[3]/47,idem,+8[8]/50,idem,+6,+8,+9,der(9)t(9;11)(p22;q23),+13[cp5] t(9;11)(p22;q23) *MLL-AF9* Hyperdiploid, Trisomy 8	47, XY, +11 Trisomy 11	46,XX,dup(2)(q21q33),t(8;16)(p11;p13),psu,dic(22;1)(p11;p11)[10]/46,XX,dup(1)(q32q42),t(8;16),psu,dic(9;1)(q34;p11)[4]/46,XX,t(8;16),psu dic (19;1)(p13;p11) [4]/46,XX[2] t(8;16)(p11;p13) KAT6A-CREBBP
Molecular details	*FLT3* NM_004119 c.2027A>C p.N676T - in 79.6% of 314 reads Gain of *PIM1* (on 6p), gain of *RAD21* (on 8q), gain of *FLT3* and *PDS5B* (on 13q), gain of *CALR*, *NOTCH3*, *JAK3* and *MEF2B* (on 19p), gain of *CEBPA*, *CNOT3* and *U2AF2* (on 19q), gain of *RUNX1* and *U2AF1* (on 21q).	*DNMT3A* NM_175629 c.1868_1868insT p.Y623fs* - in 4.0% of 375 reads *FLT3* NM_004119 c.2027A>C p.N676T - in 88.7% of 309 reads (*FLT3*-ITD) Gain of *PIM1* (on 6p), gain of *RAD21* (on 8q), gain of *FLT3* and *PDS5B* (on 13q).	*NOTCH1* NM_017617 c.7230_7230insTG p.P2411fs* - in 52.1% of 121 reads Gain of *PIM1* (on 6p) *TP53* mutation (p.V173M)

B



C

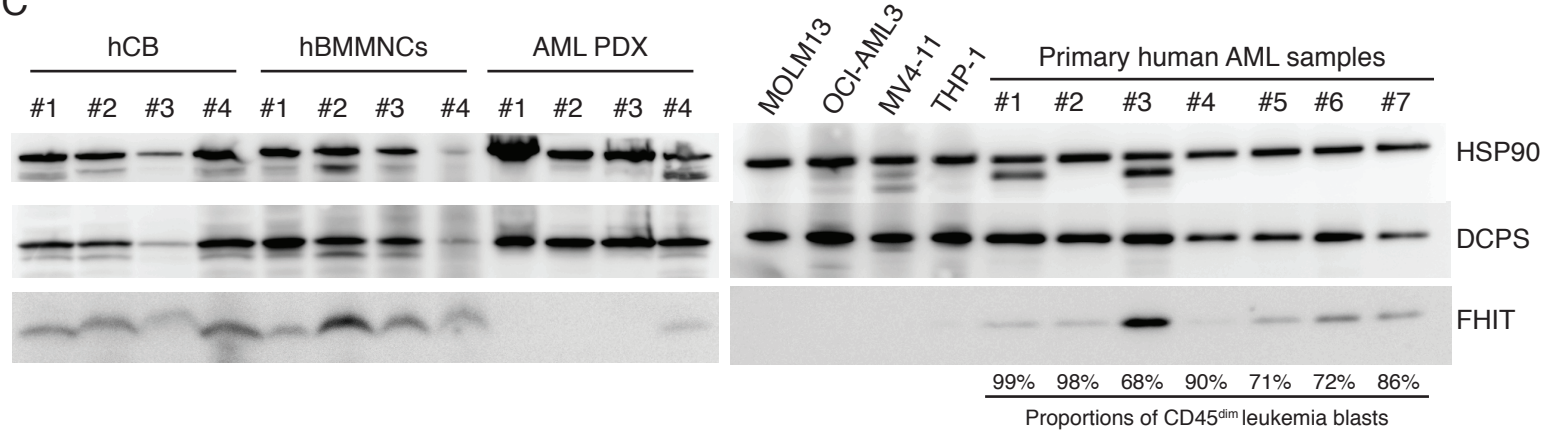


Figure S6. Characteristics of AML PDX lines and Western blot analysis for FHIT. Related to Figure 6. (A) Characteristics of three PDX lines. Information regarding WHO clarification, FAB classification and karyotype was obtained from the ProXe website. Mutational status as assessed by the Rapid Heme Panel (DFCI). **(B)** Representative results of CETSA using human AML cells harvested from BM of DMSO- or RG3039-treated PDX mice. For CETSA analysis AML cells in BM (DFAM-15354) were harvested the day after the last RG3039 (or DMSO) injection. Asterisk denotes non-specific signal. Equal protein loading was validated by Ponceau-S staining. **(C)** Western blot analysis of FHIT using anti-FHIT antibody.