

AWARD NUMBER: W81XWH-15-1-0425

TITLE:Identifying Neurofibromin-Specific Regulatory Nodes for Therapeutic Targeting in NF1

PRINCIPAL INVESTIGATOR: Frank P. McCormick, PhD

CONTRACTING ORGANIZATION: University of California, San Francisco  
San Francisco 94143

REPORT DATE: December 29, 2018

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

**REPORT DOCUMENTATION PAGE**Form Approved  
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE</b> 12/29/18		<b>2. REPORT TYPE</b> Final Report		<b>3. DATES COVERED</b> 30-SEP-2015 -29-SEP-2018	
<b>4. TITLE AND SUBTITLE</b> Identifying Neurofibromin-Specific Regulatory Nodes for Therapeutic Targeting in NF1				<b>5a. CONTRACT NUMBER</b> W81XWH-15-1-0425	
				<b>5b. GRANT NUMBER</b> 0010711237-0001	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b> Frank McCormick, Evan Markegard, Osamu Tetsu  E-Mail: frank.mccormick@ucsf.edu				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> University of California, Helen Diller Family Comprehensive Cancer Center, San Francisco, CA 94158-9001				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>  U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>  Approved for Public Release; Distribution Unlimited					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b> The overall goal of the project is to determine how neurofibromin is regulated, and to explore the hypothesis that loss of neurofibromin activity leads to up-regulation of specific receptors. We are building on our earlier discovery, that neurofibromin depends on the adapter protein SPRED1, to function, and we are utilizing the latest technical innovations including CRISPR technology to find genes that regulate neurofibromin SPRED function. To date we have demonstrated that oncogenic EGFR signaling disrupts Spred1-NF1 binding. Mass spectrometry was performed on cells overexpressing EGFR858R to identify potential phosphorylation sites on Spred1 and NF1 that could disrupt Spred1-NF1 binding by steric hindrance. A serine phosphorylation site on Spred1 was identified in which a phosphomimetic and phosphodeficient mutant decreased or increased Spred1-NF1 binding, respectively. Therefore, phosphorylation of Spred1 at this site by a serine kinase downstream of oncogenic EGFR may disrupt Spred1-NF1 binding. Our findings provide one potential mechanism by which oncogenic EGFR signaling disrupts negative feedback to allow for constitutive Ras signaling. Furthermore, this work may elucidate a novel kinase therapeutic target for restoring NF1 mediated inhibition of Ras.					
<b>15. SUBJECT TERMS</b> Neurofibromin, Spred1, Spred2, neurofibromatosis, therapeutic targeting					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>
<b>a. REPORT</b>	<b>b. ABSTRACT</b>	<b>c. THIS PAGE</b>			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	95	<b>19b. TELEPHONE NUMBER</b> (include area code)

## Table of Contents

	<u>Page</u>
1. Introduction .....	1
2. Keywords.....	1
3. Accomplishments... ..	1
4. Impact .....	29
5. Changes/Problems. ....	29
6. Products... ..	31
7. Participants & Other Collaborating Organizations .....	32
8. Special Reporting Requirements .....	37
9. Appendix.....	i

## 1. Introduction

The overall goal of the project is to determine how neurofibromin is regulated, and to explore the hypothesis that loss of neurofibromin activity leads to up-regulation of specific receptors. We hope that we will be able to find news of up-regulating neurofibromin in cells in which one allele is defective, and that we will be able to suppress hyperactive Ras signaling in cells in which both alleles are lost. We are building on our earlier discovery, that neurofibromin depends on the adapter protein SPRED1, to function, and we are utilizing the latest technical innovations including CRISPR technology to find genes that regulate neurofibromin SPRED function.

## 2. Keywords

Neurofibromin, Spred1, Spred2, EGFR, mutant EGFR(L858R), Ras-GTP, Ras-MAPK, RTK, neurofibromatosis, RasGAPs, growth factor signaling, therapeutic targeting.

## 3. Accomplishments

### Major Goals of the Project

#### Specific Aim 1

**Major Task 1:** Determine and verify RTKs that interact with Spred1 and Spred2

- Subtask 1: Generation of plasmids and co-IP of exogenous RTKs and Spred1/2
- Subtask 2: Validation of co-IP, endogenous protein co-IP
- Subtask 3: Determine the binding sites and tyrosine phosphorylation pattern on Spred for the RTK/Spred interaction

**Major Task 2:** Determine effect of NF1/Spred loss on Spred-binding RTK signaling

- Subtask 1: Establish good NF1 and Spred1/2 knockdown protocols for indicated cell lines
- Subtask 2: Establish good biological readout assays for cell lines to be tested in response to specific RTKs
- Subtask 3: Determining Ras-MAPK signaling and biological readouts with NF1 or Spred knockdown in response to RTK activation

#### Specific Aim 2

**Major Task 3:** Determine EGFR-mediated changes in Spred and Neurofibromin phosphorylation and interaction partners

- Subtask 1: Purification of Neurofibromin or Spred from cells +/-mutant EGFR
- Subtask 2: Mass spec analysis
- Subtask 3: Mutational analysis of Spred and Neurofibromin phosphorylation sites on Spred/Neurofibromin interaction

- Subtask 4: Knockdown analysis of interacting proteins on Spred/Neurofibromin interaction

**Major Task 4:** Determine targets for increasing Spred/Neurofibromin interaction and NF1 function in NF1<sup>-/+</sup> cells

- Subtask 1: Establish good siRNA and/or drug inhibition protocols for the regulators identified above in relevant NF1<sup>-/+</sup> cell lines
- Subtask 2: Test the effect of siRNA and/or drugs above in Ras-MAPK signaling and functional assays: proliferation, migration, differentiation

### **Specific Aim 3**

**Major Task 5:** Determine novel post-translational modifications and interacting proteins with Neurofibromin

- Subtask 1: Generate cells stably expressing Tet-regulatable full length NF1 and NF1 point mutants
- Subtask 2: Purify NF1 from cells with multiple conditions
- Subtask 3: Mass spec analysis, identify proteins that interact differentially between WT and mutant NF1
- Subtask 4: Determination of effect of differentially interacting proteins on NF1 function

## Major Task 1: Determine and verify RTKs that interact with Spred1 and Spred2

### Subtask 1: Generation of plasmids and co-IP of exogenous RTKs and Spred1/2

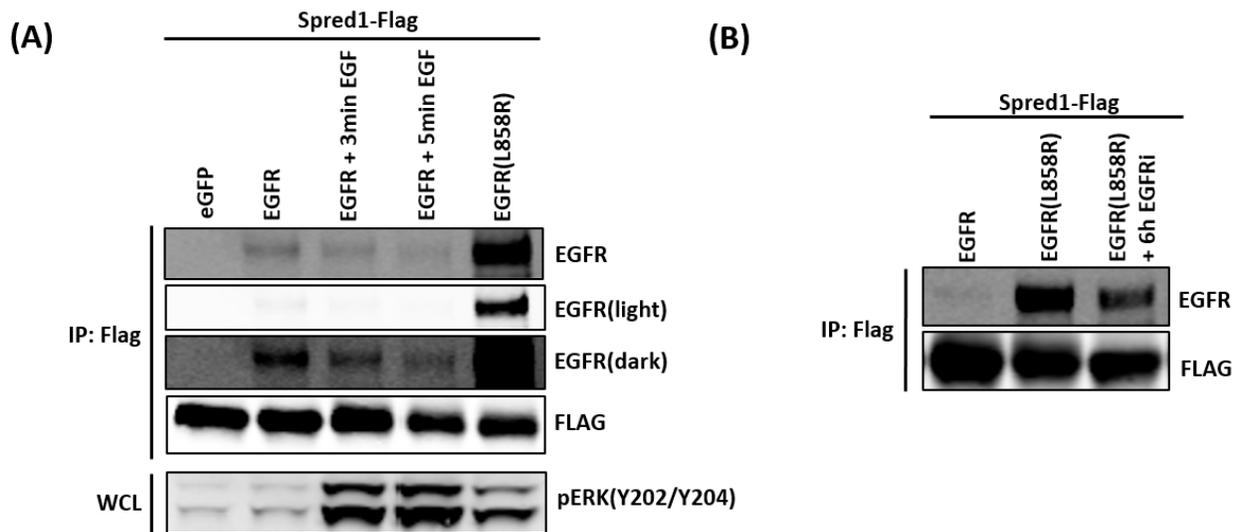
The RTK plasmids that have been generated are shown in Table 1. As a pilot experiment, Spred1 has been shown to bind to the activated form of EGFR(L858R) at a higher affinity than wild-type EGFR as shown in Figure 1. Interestingly, activated EGFR with EGF does not increase Spred1 binding and EGFR inhibition reduces EGFR(L858R)-Spred1 binding. These results are surprising and important for understanding the role of Spred1 on both wild-type and pathogenic RTKs. Spred1 is deleted across multiple cancer types while Spred2 is not. Therefore, given the clinical importance of Spred1, we have prioritized our studies to thoroughly investigate Spred1 before Spred2. We previously identified phosphorylation sites on Spred1 by mass spectrometry following EGFR(L858R) cotransfection. pSpred1(S105) is particularly interesting due to loss of NF1 binding and so additional RTKs have been tested for ability to phosphorylate Spred1 at S105 as shown in Table 1. These samples have been submitted for mass spectrometry analysis and RTK binding is ongoing.

**Table 1**

Candidate RTKs	
EGFR	EGFR(L858R)
CKIT	CKIT(D816V)
FLT3	FLT3(ITD) FLT3(D835Y)
PDGFR $\alpha$	PDGFR $\alpha$ (D842V)
HER2	HER2(V654E)
MET	
BCR-ABL	
CSF1R	

**Table 1.** RTKs which have been cotransfected with Spred1 to determine ability to phosphorylate Spred1

**Figure 1**



**Figure 1.** Spred 1 binds mutant EGFR(L858R) with higher affinity than wild-type EGFR. A) HEK 293T cells were transfected, serum starved for 16 hours, and stimulated with 20ng/ml EGF and EGFR binding accessed by Western blot. (B) Cells were transfected and Spred-NF1 binding was assessed by Western blot. EGFR inhibitor AZD9291 used at 1 $\mu$ M for 6 hours.

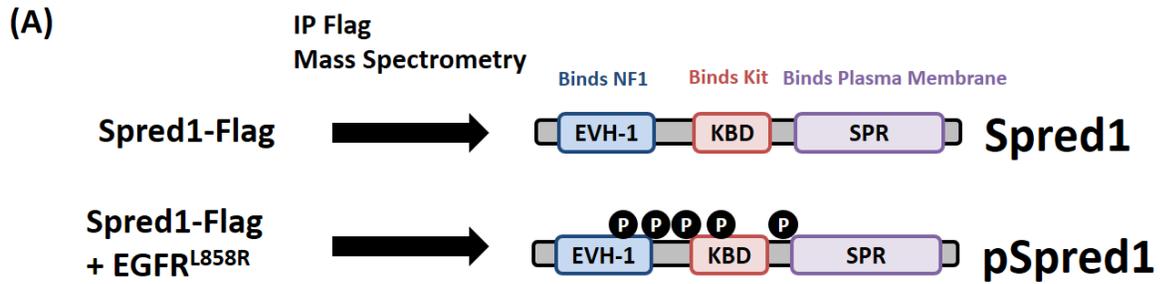
Subtask 2: Validation of co-IP, endogenous protein co-IP

Validation of co-IP and endogenous protein co-IP is ongoing.

Subtask 3: Determine the binding sites and tyrosine phosphorylation pattern on Spred for the RTK/Spred interaction

Flag tagged Spred1 was overexpressed in 293T cells with and without EGFR(L858R). Spred1 was immunoprecipitated and subject to mass spectrometry analysis. Phosphorylation and binding partners with mutant **EGFR** are shown below in Figure 2.

Figure 2



(B)

Phospho Site	Domain	Phospho Motif	Kinase Phospho Motif	Human Genetics
S105	EVH-1	FGLTFQpSPADARA	ERK1, ERK2, GSK-3, CDK5, CK1, CK2, GPKR1, PDH	T102R identified in Legius syndrome patient
S148&S149	Near EVH-1	QANEEDSpSpSSLVKD	CK2, BARK	N/A
S238	KBD	KSIRHVpSFQDEDE	CaMKII, PKA, PKC, CK2, AKT, CLK2, PKCe	N/A
Y292	Near KBD	DSKKSDpYLYSCGD	EGFR, PDGFR	Y292F identified in stomach adenocarcinoma patient

(C)

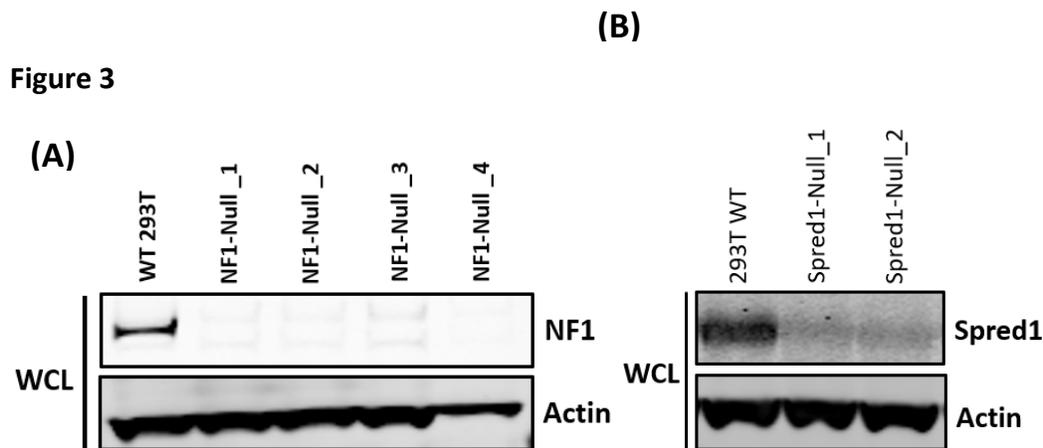
EGFR, ERBB2, ERBB4, CALM1, RPL36, NPM1, HCN3, RPL28, RPL38, NDUFA5, TEX10, NDUFS2, XRCC6, C3orf58, SRP9, FAM91A1, SLIRP, HIGD1A, ZNF444, SUB1, NOB1, EIF3L, PNO1, ASPHD1, MKI67IP, NOP2, TSPYL1, SRSF2, SF3B3, ZNF638, LBR, NSUN2, RRP1, RPUSD4, WDR18, TIMMDC1, FAM132B, ZNF696, ZFPL1, AIFM1, RPL10L, KRT17, KMT2B, SNRPF, DNAJC19, RRP36

**Figure 2.** Identification of Spred1 phosphorylation and binding partners in the presence of mutant EGFR (L8585R). (A) Schematic for Spred1 mass spectrometry analysis. (B) Identification of Spred1 phosphorylation sites downstream of mutant EGFR. Kinase phosphorylation motifs identified using PhosphoMotif Finder and PhosphoSitePlus. (C) Spred1 interacting proteins identified by mass spectrometry.

## Major Task 2: Determine effect of NF1/Spred loss on Spred-binding RTK signaling

Subtask 1: Establish good NF1 and Spred1/2 knockdown protocols for indicated cell lines

NF1-Null and Spred1-Null HEK 293T cells have been generated using CRISPR/Cas9 and single clones have been expanded for biochemical assays (Figure 3). NF1-Null MEFs have also been generated and will be characterized. Spred1-Null HEK 293T and MEFs are currently being generated using CRISPR/Cas9.



**Figure 3.** Clonal NF1-Null and Spred1-Null HEK 293T cells lines were generated using CRISPR/Cas9 with a sgRNA targeting exon 2 of NF1 and exon 1 of Spred1. Following antibiotic selection clones were screened by Western blot.

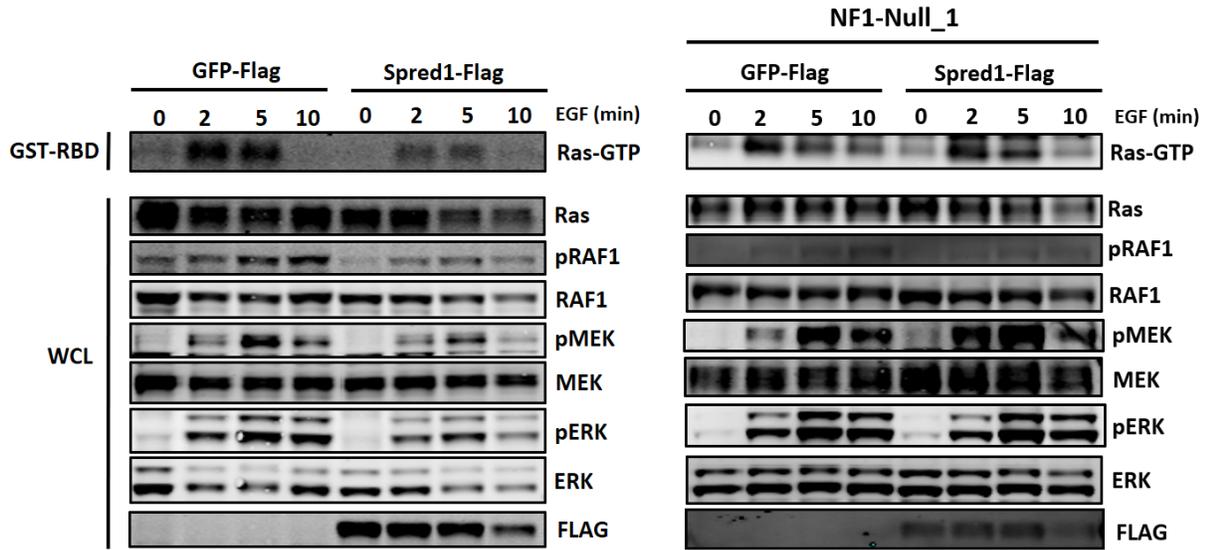
Subtask 2: Establish good biological readout assays for cell lines to be tested in response to specific RTKs

Ras-GTP pulldown assay with GST-RBD, pRAF1, pMEK, and pERK have been optimized as demonstrated in Subtask 3.

Subtask 3: Determining Ras-MAPK signaling and biological readouts with NF1 or Spred knockdown in response to RTK activation

Ras-MAPK signaling in response to EGF has been analyzed in NF1 wild-type and NF1-Null 293T cells (Figure 4). Both the magnitude and duration of Ras-MAPK signaling seems to be enhanced in the NF1-Null 293T cells compared to wild-type. Also, Spred1 overexpression is unable to inhibit Ras-MAPK signaling in NF1-null cells compared to wild-type.

**Figure 4**



**Figure 4.** Spred1 overexpression is unable to suppress RasGTP following EGF stimulation in NF1-Null cells. NF1-Null\_1 cells were transfected, serum starved for 16 hours, and stimulated with 20ng/ml EGF. Downstream signaling was accessed by Western blot.

### Major Task 3: Determine EGFR-mediated changes in Spred and Neurofibromin phosphorylation and interaction partners

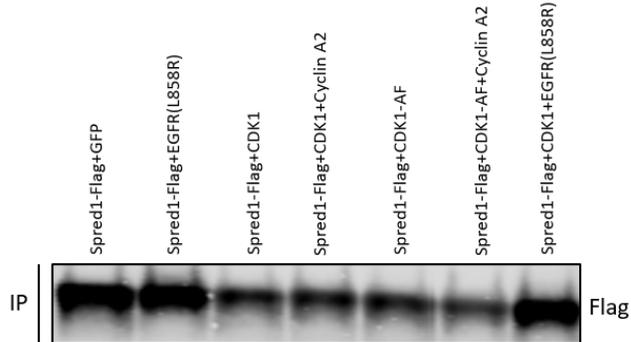
#### Subtask 1: Purification of Neurofibromin or Spred from cells +/-mutant EGFR

See Subtask 2

#### Subtask 2: Mass spec analysis

Neurofibromin and Spred1 have been successfully purified with and without activating EGFR(L858R) and analyzed by mass spec. as shown in Figure 5. For Spred1 see Major Task 1, Subtask 3 and for NF1 see Major Task 5, Subtask 3.

**Figure 5**

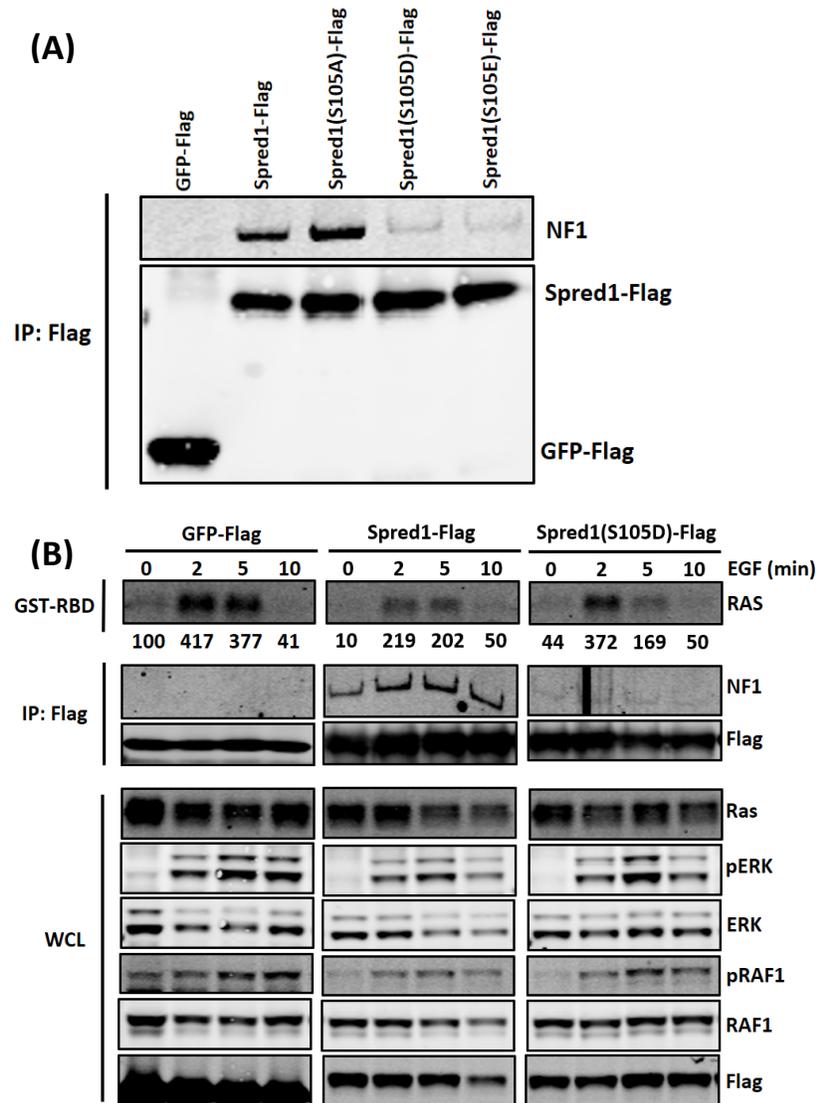


**Figure 5.** Spred1-flag IP from HERK 293T cells for mass spec analysis.

Subtask 3: Mutational analysis of Spred and Neurofibromin phosphorylation sites on Spred/Neurofibromin interaction

Phosphodeficient and phosphomimic mutants of Spred1 phosphorylation sites identified by mass spectrometry were generated and tested for effects on NF1 binding. Spred1(S105) mutants showed the expected binding (Figure 6). Importantly, phosphomimic Spred1(S105D) is unable to bind NF1 and suppress Ras-GTP following EGF stimulation.

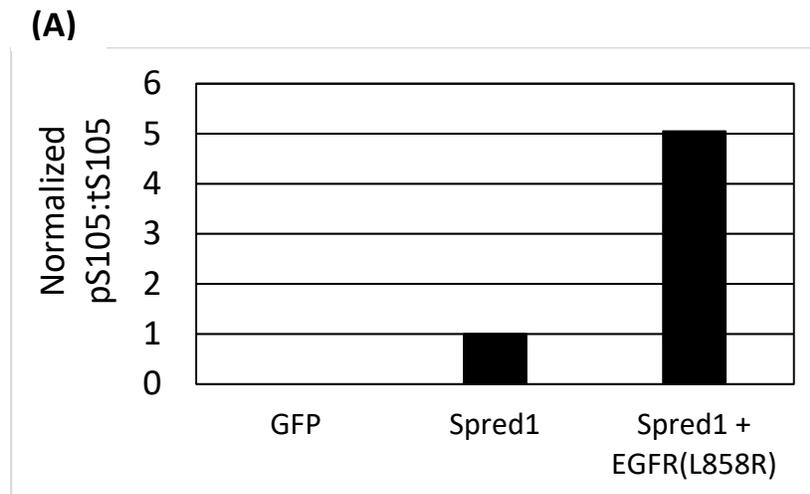
**Figure 6**



**Figure 6.** Phosphomimetic Spred1(S105) is unable to bind NF1 and suppress RasGTP following EGF stimulation. (A) Phosphomimetic and phosphodeficient Spred1 mutants at S105 were generated using site directed mutagenesis and transfected into HEK 293T cells to determine NF1 binding by Flag-IP followed by Western blot. (B) HEK 293T cells were transfected, serum starved for 16 hours, and stimulated with 20ng/ml EGF. Downstream signaling was accessed by Western blot.

Phosphorylated Spred1(S105) has been quantitated and normalized to total Spred1(S105) using transient transfection in 293T cells (Figure 7) and mass spectrometry as previously described. Expression of mutant EGFR(L858R) leads to a 5-fold increase in phospho-Spred1(S105). We will now test candidate kinase inhibitors which reduce phospho-Spred1(S105).

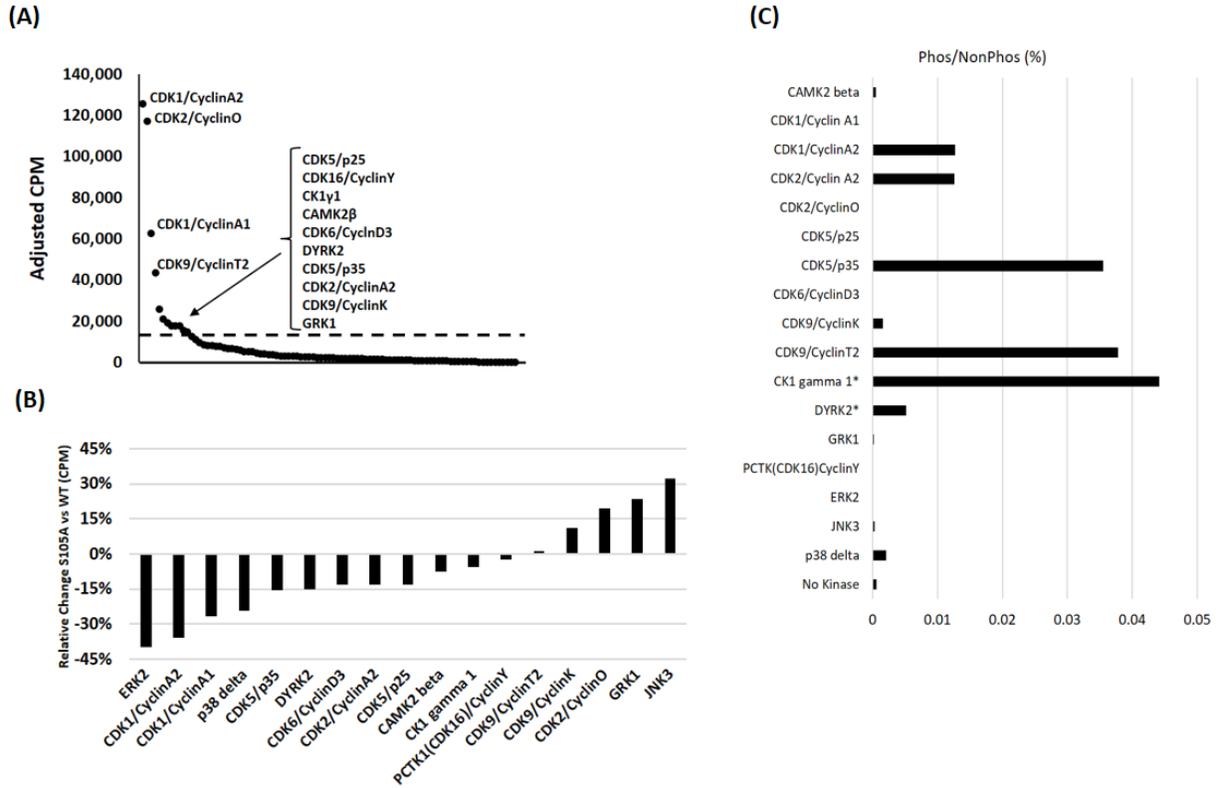
**Figure 7**



**Figure 7.** Design of mass spectrometry experiment to identify Spred1(S105) kinase. Quantitation of relative phosphorylation at the S105 site in SPRED1 was carried out in Skyline v3.0 by quantifying MS1 precursor peak areas of the S105-containing peptides and normalizing them by the sum of abundances of all unmodified peptides detected in the same protein. Transient transfection in HEK 293T was carried out as previously described.

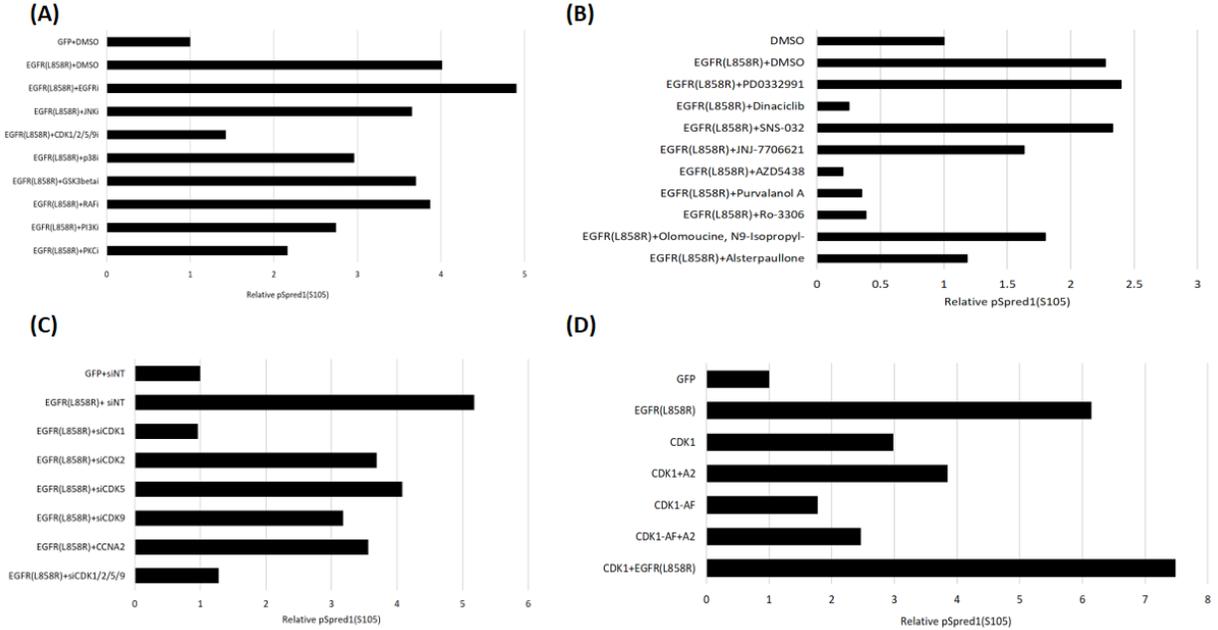
In addition to identifying Spred1(S105) as an important phosphorylation site for NF1 binding and Ras signaling we have also identified CDK1 as the Spred1(S105) kinase. Figure 8 shows results from an in vitro kinase assay with Spred1(S105) peptide, an in vitro kinase assay comparing wild-type Spred1 to Spred1(S105A), and a full length Spred1 in vitro kinase assay followed by mass spec for Spred1(S105) phosphorylation. Figure 9 shows Spred1 phosphorylation at S105 decreases in response to CDK1 inhibitors and increases with CDK1 expression.

**Figure 8**



**Figure 8.** CDK1 phosphorylates Spred1 on S105 in vitro. (A) In vitro kinase assay with Spred1(S105) peptide performed by Kinexus. (B) In vitro kinase assay comparing wild-type Spred1 to Spred1(S105A). (C) In vitro kinase assay with full length Spred1 followed by mass spec for Spred1(S105) phosphorylation.

**Figure 9**

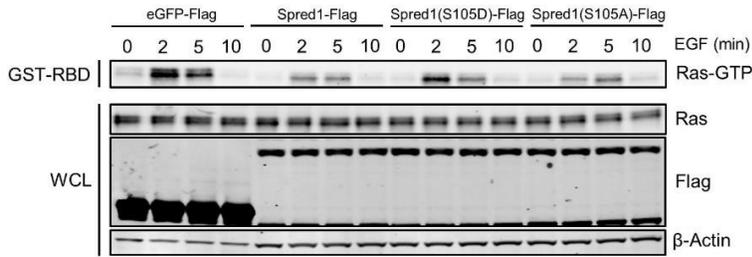


**Figure 9.** CDK1 phosphorylates Spred1 in vivo. (A) Spred1-Flag was overexpressed in HEK 293T cells, immunoprecipitated, and analyzed for Spred1(S105) phosphorylation by mass spec in the presence of selected pathway inhibitors. All inhibitors were used at 2 $\mu$ M for 2 hours. EGFR inhibitor AZD9291, RAF inhibitor LY3009120, GSK3 $\beta$  inhibitor CHIR-99021, CDK1/2/5/9 inhibitor Dinaciclib, PKC inhibitor Sotrastaurin, JNK inhibitor JNK-IN-8, p38 inhibitor LY2228820, and PI3K inhibition GDC-0941. (B) As shown in (A) but with additional CDK1 inhibitors. (C) As shown in (A) but with siRNA knockdown. (D) Spred1-Flag was overexpressed in HEK 293T cells with the addition of CDK1 constructs.

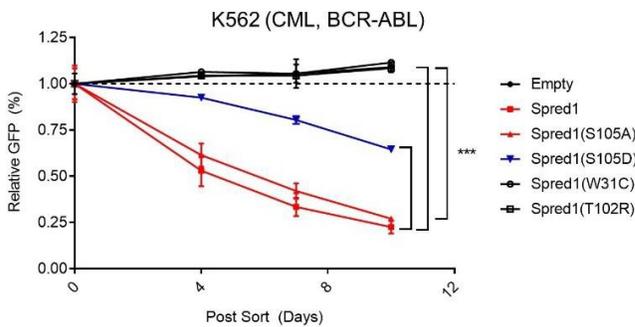
Our collaborators Boettcher et al. recently discovered, through an unbiased whole genome CRISPRa screen, that SPRED1 and NF1 overexpression inhibit K562 proliferation. K562 is a chronic myeloid leukemia (CML) cell line with the BCR-ABL oncogene, is dependent on Ras-GTP for proliferation and may be unable to phosphorylate and inactivate Spred1 on serine 105. Therefore, we expected the K562 cell line would be an ideal model system to test the biological effects of phosphomimetic Spred1(S105). We infected K562 cells with Spred1-IRES-GFP expressing retrovirus and performed a competition assay between infected (GFP+) and uninfected (GFP-) cells (Figure 10B). Representative flow cytometry GFP histograms show similar infection rates and expression levels (Figure 10C). Spred1 wild type infected GFP positive cells were outcompeted by uninfected GFP negative cells while the empty vector controls were not. Phosphomimetic Spred1(S105) infected cells were unable to inhibit proliferation, along with Spred1 Legius syndrome patient mutants W31C and T102R.

**Figure 10**

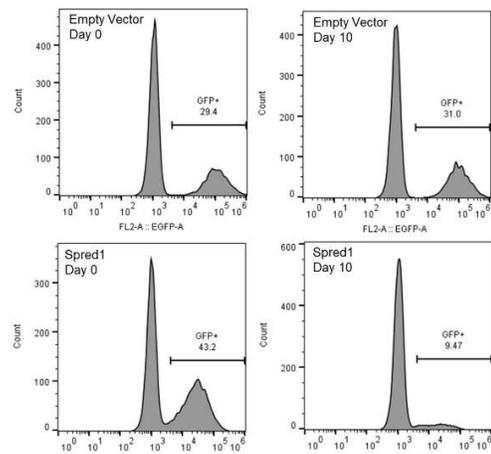
**A**



**B**



**C**

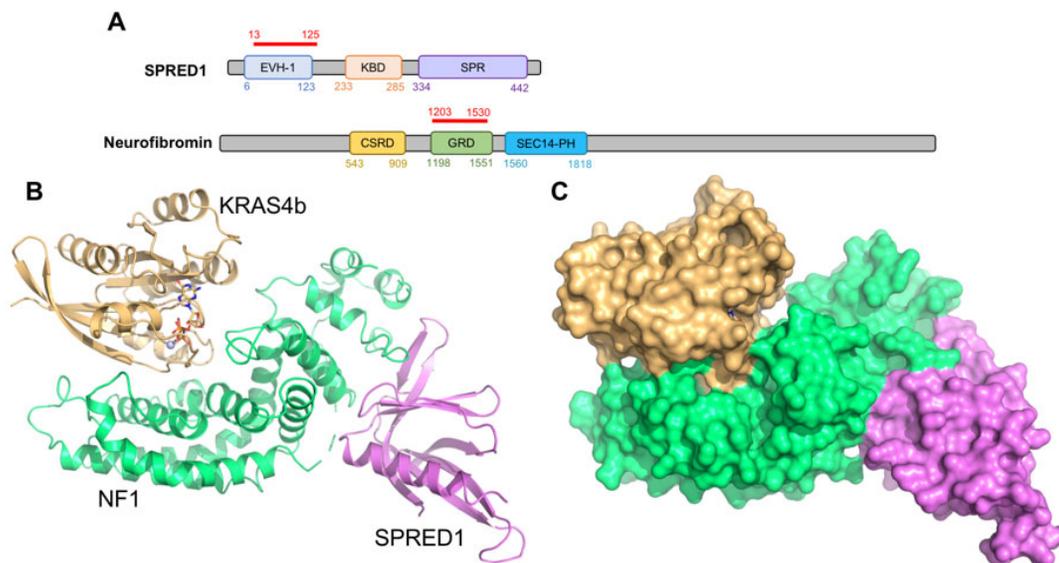


**Figure 10** Phosphomimetic Spred1(S105) alters Ras-GTP signaling following EGF stimulation and K562 proliferation

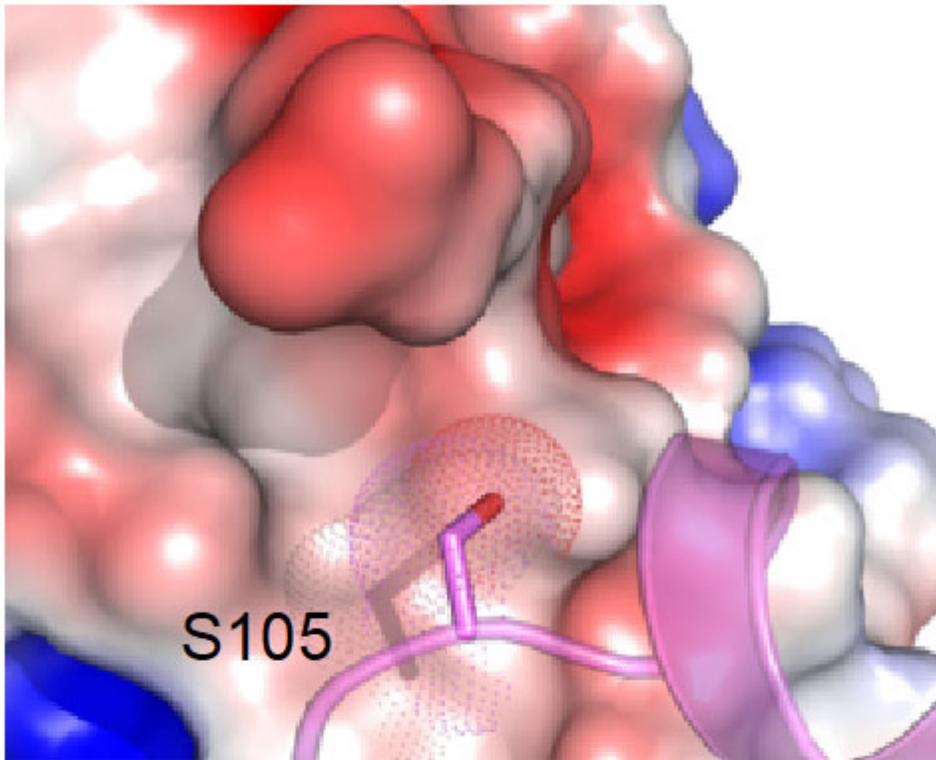
- 1) HEK 293T cells were transfected, serum starved for 16 hours, and stimulated with 10 ng/ml EGF. Downstream signaling was accessed by western blot and Ras-GTP pull-down assay.
- 2) K562 cells were infected with Spred1-IRES-GFP expressing retrovirus. Three days after infection baseline GFP-positive cells were measured by flow cytometry and normalized to 1. GFP positive cells were monitored over time to measure the effect of Spred1 expression on proliferation. The statistical significance of the difference between indicated samples was determined using a two-way ANOVA. \*\*\*P < 0.001.
- 3) Representative flow cytometry GFP histograms.

In collaboration with the structural biology group at Frederick National Labs, we have solved the structure of the KRAS/NF1/SPRED complex, as shown in Figure 11 below.

**Figure 11**



These structures show that RAS and SPRED1's EVH domain bind at distinct sites on the NF1 GAP-related domain, as predicted from our biochemical analysis, published in collaboration with Klaus Scheffzek. Of particular relevance to this project, 105 is directly in the interface between SPRED1 and NF1 (below). The phosphomimetic S105D presents a steric clash which accounts for its reduced binding.



SPRED1 vs NF1

$$K_d = 0.29 \pm 0.03 \mu\text{M}$$

*SPRED1-S105D* vs NF1

$$K_d = 2.63 \pm 0.68 \mu\text{M}$$

## **Major task 4: Determine targets for increasing Spred/Neurofibromin interaction and NF1 function in NF1-/+ cells**

### Subtask 1: Establish good siRNA and/or drug inhibition protocols for the regulators identified above in relevant NF1-/+ cell lines

This sub-task was modified to use CRISPRa technology to identify genes which, when over-expressed in cells, would differentially regulate NF1 -/- cells relative to wild-type.

CRISPRa is a relatively new and efficient screening technique to identify the effect of gene overexpression. Our colleagues Boettcher *et al.*, performed a full genome CRISPRa screen in the BCR-ABL fusion leukemia cell line K562 with imatinib (paper attached) an ABL inhibitor to apply selective pressure, to identify genes that effect proliferation. These cells are an ideal model system since they grow in suspension, allowing for high coverage of a whole genome screen. Genes which when overexpressed inhibited proliferation included NF1, SPRED1, and SPRED2. Importantly, KRAS overexpression increased proliferation, confirming RAS dependence. Further validation screening revealed that Neurofibromin loss negated the toxic effects of SPRED2 overexpression, confirming Spred proteins act through Neurofibromin.

These results show the K562 cell line is an ideal model system for screening genes which regulate Neurofibromin since proliferation is dependent on activity of the fully functional Ras pathway. Additional hits from the Neurofibromin loss candidate screen include the RTK AXL and the AML fusion gene partner NOL4L. When overexpressed in NF1-null cells, but not wild type, proliferation increased, suggesting these genes may suppress proliferation through Neurofibromin. To expand this exciting preliminary data, we performed a whole genome CRISPRa screen comparing wild type to NF1-null K562 cells in the presence of imatinib. We identified many novel genes which may promote or inhibit proliferation through Neurofibromin.

Our CRISPRa screen generated many compelling hits from the Ras pathway, as expected (Table 2-1). The top hit from the CRISPRa screen, which when overexpressed inhibited proliferation in the wild type but not NF1-null cells, was Transforming growth factor beta receptor 2 (TGFB2). TGF- $\beta$  signaling inhibits proliferation and stimulates erythroid differentiation. The TGFB2 ligand TGF- $\beta$ 1 inhibits proliferation of K562 cells, potentially through differentiation into red blood cells (RBCs). Additionally, ABL and MEK inhibition differentiate K562 cells into RBCs, supporting the role of MAPK signaling to prevent differentiation and maintain proliferation.

Importantly, in this whole genome, unbiased screen Spred1 was a top hit, further validating the inhibitory function of Spred1 is dependent on Neurofibromin. Other notable hits were DUSP9, an ERK1/2 phosphatase. In wild type cells, increasing expression of DUSP9 is expected to decrease phosphorylated ERK1/2 and inhibit proliferation. However, NF1-null cells are expected to have elevated Ras-GTP, which may increase phosphorylated ERK1/2 to counterbalance DUSP9 overexpression. Other hits include CBL, a known E3 ubiquitin-protein ligase which degrades BCR-ABL, and tyrosine-protein phosphatase non-receptor type 1 (PTPN1 also known

as PTP1B) an ABL phosphatase. Both CBL and PTPN1 inhibit BCR-ABL signaling in K562 wild type cells, inhibiting proliferation, while NF1 deletion rescues proliferation, likely through elevated Ras-GTP signaling downstream of BCR-ABL. RASA2 and RASA3 are both hits and RasGAPs, suggesting these GAPs may not be functional without Neurofibromin or levels of Ras-GTP are too high from Neurofibromin loss to be reduced. The hit FOXO4 encodes the transcription factor Forkhead box protein O4 which inhibits proliferation, but this inhibition may be dependent on Neurofibromin.

Our CRISPRa screen also identified gene which when overexpressed increased proliferation in wild type K562 cells but not NF1-null. As expected, our top hit was BCR, which leads to increase BCR-ABL expression, a previously validated resistance mechanism to imatinib. However, NF1-null K562 proliferation does not increase with BCR overexpression, suggesting NF1 deletion is sufficient to maximize proliferation, likely through increased Ras-GTP signaling. SOS1, a RasGEF, is also an expected top hit. Additional hits in the Ras pathway include SHOC2, which facilitates Ras/Raf1 binding, and GAB2, a scaffolding protein that links receptor signaling to effectors. Gene Set Enrichment Analysis (GSEA) was performed to identify pathways of interest (Table 2-2). Example enrichment plots for WNT  $\beta$ -Catenin Signaling and KRAS Signaling are shown in Figure 12. Of special interest, TGF- $\beta$  signaling genes were enriched when proliferation was inhibited in wild type cells but not NF1-null cells, supporting the potential role of Neurofibromin in TGF- $\beta$  signaling.

Our CRISPRa screen identified many potentially novel regulators of Neurofibromin which must be validated in additional cell lines. One of the most promising hits was Transforming growth factor beta receptor 2 (TGFB2) which has not been directly linked to Neurofibromin. Additional insight into erythropoiesis could be gained by studying TGFB2 and MAPK signaling. The TGF- $\beta$  pathway could also be validated through treatment of wild type and NF1-null K562 cells with the TGFB2 ligand TGF- $\beta$ 1. Wild type cells may differentiate while NF1-null cells continue to proliferate, with the caveat that receptor expression may be limiting. To circumvent this limitation, K562 cells overexpressing TGFB2 which have previously been generated could be used.

As expected for a Ras-dependent cell line, BCR and SOS1 overexpression increases proliferation in wild type K562 cells but decreases proliferation in NF1-null K562 cells. This data supports the hypothesis of oncogene overdose. NF1-null cells likely have elevated Ras-GTP and were selected for optimal proliferation. But the overexpression of BCR and SOS1 likely leads to super elevated Ras-GTP, which is detrimental to proliferation.

Our CRISPRa screen comparing wild type to NF1-null K562 cell lines generated many expected hits, validating the screen. Many of the hits were unexpected, which could lead to the discovery of novel Ras and/or Neurofibromin effectors. Additional follow-up is needed to validate hits with individual sgRNA in the K562 cell line and move beyond leukemia to other solid cancer cell lines. To strength the hits, a separate CRISPRi screen could be performed using the same library and protocol to generate a list of hits to cross reference as previously demonstrated. We have already deleted NF1 in the K562 cell line containing the stable CRISPRi machinery in anticipation of this screen.

FIGURE 12

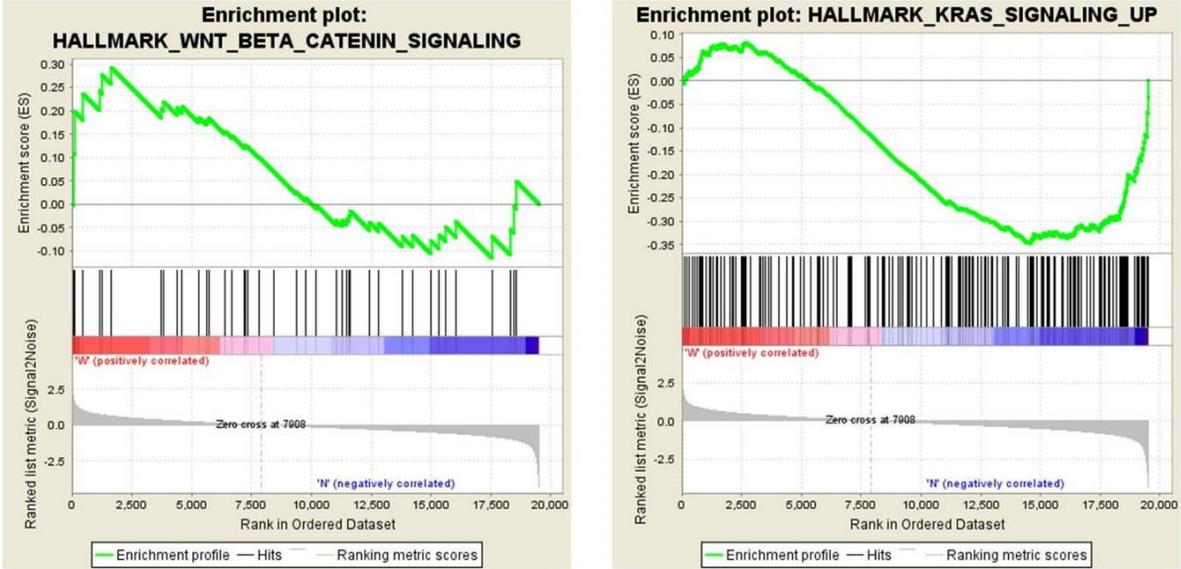


Figure 12. Representative GSEA plots from the CRISPRa screen

**TABLE 2-1**

Overexpression inhibits Proliferation of K562 WT but not NF1-Null					Overexpression promotes proliferation of K562 WT but not NF1-Null				
Hit	Gene	WT	NF1-Null	Difference	Hit	Gene	WT	NF1-Null	Difference
1	TGFBR2	-3.72	0.08	3.81	1	BCR	2.97	-1.30	-4.28
2	DUSP9	-3.61	-0.31	3.30	2	SOS1	3.51	-0.42	-3.92
3	HNF4A	-4.03	-0.78	3.25	3	ABCG2	3.39	0.35	-3.04
4	UBASH3A	-3.46	-0.25	3.20	4	GAB2	2.28	-0.40	-2.68
5	TFAP2A	-3.94	-0.85	3.08	5	RTDR1	1.08	-1.13	-2.20
6	PRDM1	-3.74	-0.70	3.04	6	SLC6A14	2.49	0.34	-2.14
7	IKZF3	-3.85	-0.83	3.02	7	EP300	1.83	-0.31	-2.14
8	TP73	-3.86	-0.90	2.96	8	SLC1A3	1.55	-0.54	-2.09
9	BHLHE40	-2.99	-0.11	2.88	9	SLC4A7	2.31	0.32	-1.99
10	SAMD1	-3.73	-0.86	2.88	10	MAP2K3	2.94	1.00	-1.95
11	DBN1	-3.05	-0.22	2.82	11	GAREM	1.58	-0.30	-1.88
12	UBASH3B	-3.25	-0.52	2.74	12	SLC16A1	1.29	-0.48	-1.78
13	RASA3	-2.71	0.02	2.73	13	SLC38A3	2.19	0.50	-1.69
14	CBL	-2.94	-0.23	2.70	14	PDGFRB	1.62	-0.07	-1.69
15	DUSP9	-3.26	-0.62	2.64	15	PTP4A2	1.27	-0.41	-1.68
16	PTPN1	-2.79	-0.17	2.62	16	PDGFRA	1.37	-0.30	-1.67
17	PTPN9	-3.62	-1.02	2.60	17	SEPT5	0.50	-1.11	-1.61
18	RASA2	-2.88	-0.28	2.60	18	MAP2K6	1.12	-0.41	-1.53
19	POU3F2	-3.76	-1.17	2.59	19	SLC7A2	2.09	0.57	-1.53
20	WT1	-3.75	-1.17	2.58	20	MAML3	1.91	0.39	-1.52
21	CNN1	-3.50	-0.92	2.58	21	RCOR1	1.53	0.04	-1.49
22	PITX1	-2.97	-0.41	2.56	22	MAP3K4	1.37	-0.11	-1.48
23	TFAP2C	-2.61	-0.07	2.54	23	SP3	1.20	-0.28	-1.48
24	FOXO4	-3.11	-0.59	2.52	24	SRM	0.88	-0.59	-1.47
25	MSN	-2.89	-0.40	2.50	25	BCL9L	1.46	-0.01	-1.46
26	EOMES	-2.46	-0.01	2.45	26	BBIP1	1.27	-0.19	-1.46
27	CDKN2C	-3.15	-0.71	2.45	27	MYB	0.81	-0.64	-1.45
28	SPRED1	-2.78	-0.35	2.43	28	GP1BB	0.26	-1.16	-1.43
29	E2F7	-3.24	-0.81	2.42	29	CCNE1	1.66	0.23	-1.43
30	ZBTB1	-3.36	-0.97	2.40	30	SHOC2	1.09	-0.33	-1.41

**Table 2-1.** Hits from CRISPRa screen for novel Neurofibromin effectors. Values are log2.

**TABLE 2-2**

Overexpressed promotes proliferation of K562 WT but not NF1-Null	Overexpression inhibits proliferation of K562 WT but not NF1-Null
HALLMARK_NOTCH_SIGNALING	HALLMARK_UV_RESPONSE_DN
HALLMARK_WNT_BETA_CATENIN_SIGNALING	HALLMARK_PANCREAS_BETA_CELLS
HALLMARK_DNA_REPAIR	HALLMARK_TGF_BETA_SIGNALING
	HALLMARK_HYPOXIA
	HALLMARK_MYOGENESIS
	HALLMARK_IL6_JAK_STAT3_SIGNALING
	HALLMARK_APICAL_JUNCTION
	HALLMARK_PI3K_AKT_MTOR_SIGNALING
	HALLMARK_ESTROGEN_RESPONSE_LATE
	HALLMARK_MTORC1_SIGNALING
	HALLMARK_TNFA_SIGNALING_VIA_NFKB
	HALLMARK_CHOLESTEROL_HOMEOSTASIS
	HALLMARK_KRAS_SIGNALING_UP
	HALLMARK_COMPLEMENT
	HALLMARK_G2M_CHECKPOINT
	HALLMARK_MITOTIC_SPINDLE
	HALLMARK_ESTROGEN_RESPONSE_EARLY
	HALLMARK_APOPTOSIS
	HALLMARK_P53_PATHWAY
	HALLMARK_PEROXISOME
	HALLMARK_KRAS_SIGNALING_DN
	HALLMARK_UV_RESPONSE_UP
	HALLMARK_E2F_TARGETS

**Table 2-2.** Gene Set Enrichment Analysis (GSEA) from the CRISPRa screen.Subtask 2: Test the effect of siRNA and/or drugs above in Ras-MAPK signaling and functional assays: proliferation, migration, differentiation

Once subtask 1 is completed we will test whether blocking CDK1 promotes binding, as expected, and whether increased binding reduces Ras.GTP levels and prevents cells from growing. In addition, we will analyze genes that emerge from our functional CRISPR screen that are candidates for enzymes regulating Spred/neurofibromin interaction.

## Major task 5: Determine novel post-translational modifications and interacting proteins with Neurofibromin

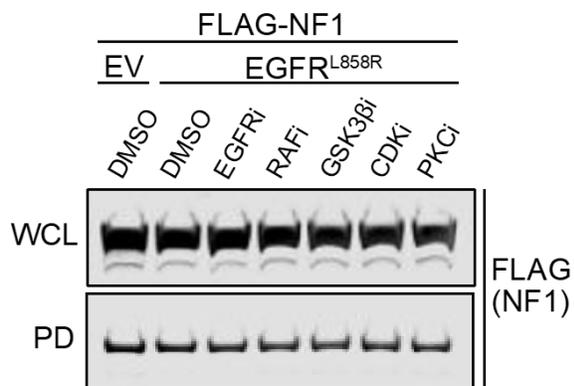
### Subtask 1: Generate cells stably expressing Tet-regulatable full length NF1 and NF1 point mutants

Rather than generate stable cell lines to express NF1, we have optimized transient transfection of NF1-Flag in HEK 293T cells to achieve the same goal as shown in Subtask 2.

### Subtask 2: Purify NF1 from cells with multiple conditions

We have successfully purified NF1 from HEK 293T with multiple conditions as shown below (Figure 13) with more details in Figure 14.

**Figure 13**

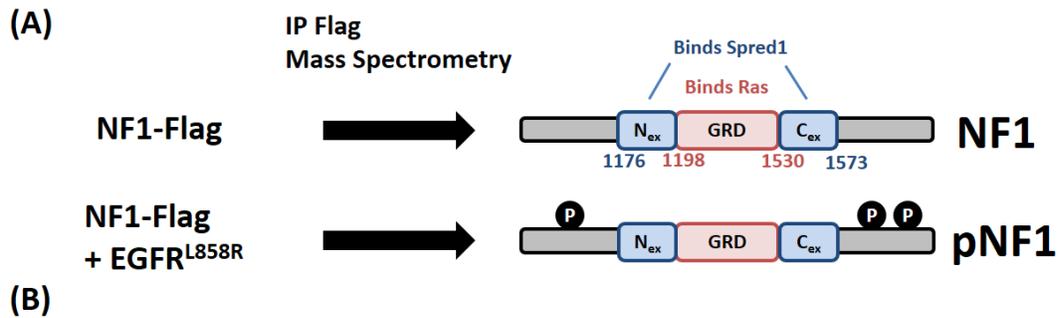


**Figure 13.** Purification of NF1 from HEK 293T cells under multiple conditions.

### Subtask 3: Mass spec analysis, identify proteins that interact differentially between WT and mutant NF1

We have performed mass spectrometry on NF1 as described above with Spred1 to identify phosphorylation sites and binding partners important for function as shown in Figure 14. We have also identified potential NF1 kinases downstream of active EGFR(L858R) by using key pathway inhibitors as shown in Figure 15. The PKC pathway may regulate multiple important phosphorylation sites on NF1 including S665/666, S876/879, S2460, T2510/2514/25115, S2521/2523, S2543, S2597, S2597/2599.

Figure 14



(B)

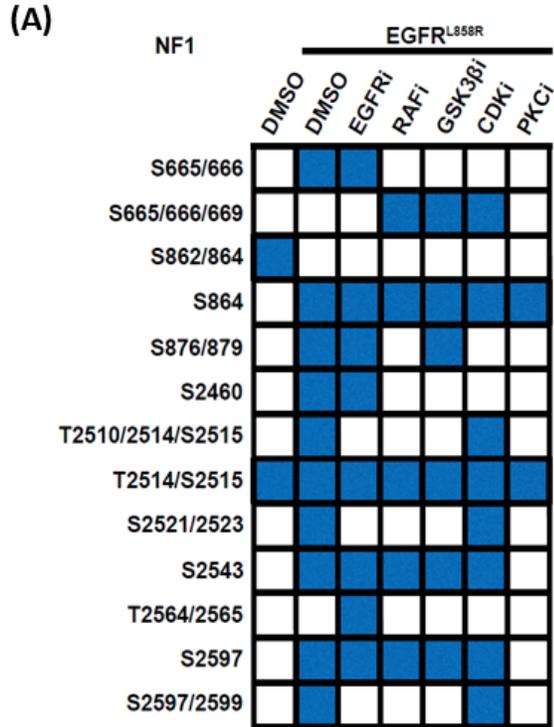
Phospho Site	Domain	Phospho Motif	Kinase Phospho Motif
S864	N/A	SGLATYpSPPMGPV	GRK1, GSK-3, ERK1, ERK2, CDK5
S2597	N/A	PHLRKVpSVSESNV	CaMKII, PKA, PKC, PKCe, CK2, GSK-3, Phosphorylase, MAPKAPK2, GRK1
S2802	N/A	KENVELpSPTTGHC	GSK-3, ERK1, ERK2, CDK5, CK1, CK2

(C)

NONO, RSP16, HNRNPH2, TUBAL3, RPS8, C22orf28, SMARCA4, EEF1A1, EEF1A1P5, ATP5B, TUBB3, TUBB6, HSPA4, RPL6, NDUFA13, EEF1B2, LEPRE1, HSPA4L, HSPH1, RPS5, ZADH2, NDUFA4, DDX53, DHRS7B, RCN2, SMARCA2, DYNC1H1, MRPS27, DDX3Y, XRN2, DSP, CALM1, KHSRP, HNRNPF, RPLPO, RPLPOP6, ATAD3C, ATAD3A, CAD, FKBP8, PPM1B, COX20, HERC2P3, DDX6, KHDRBS1, ALDH3A2, RPS19, NPM1, FAM98A, FXR1,

**Figure 14.** Identification of NF1 phosphorylation and binding partners in the presence of mutant EGFR(L8585R). (A) Schematic for NF1 mass spectrometry analysis. (B) Identification of NF1 phosphorylation sites downstream of mutant EGFR. Kinase phosphorylation motifs identified using PhosphoMotif Finder. (C) NF1 interacting proteins identified by mass spectrometry. To determine which pathways influence NF1 phosphorylation we used a panel of chemical inhibitors and identified NF1 phosphorylation. Multiple inhibitors were found to decrease NF1 phosphorylation sites.

Figure 15



**Figure 15.** Identification of additional NF1 phosphorylation sites and potential kinases induced by mutant EGFR(L858R). All inhibitors were used at 2.5μM for 4 hours. EGFR inhibitor AZD9291, RAF inhibitor LY3009120, GSK3β inhibitor CHIR-99021, CDK1/2/5/9 inhibitor Dinaciclib, PKC inhibitor Sotrastaurin.

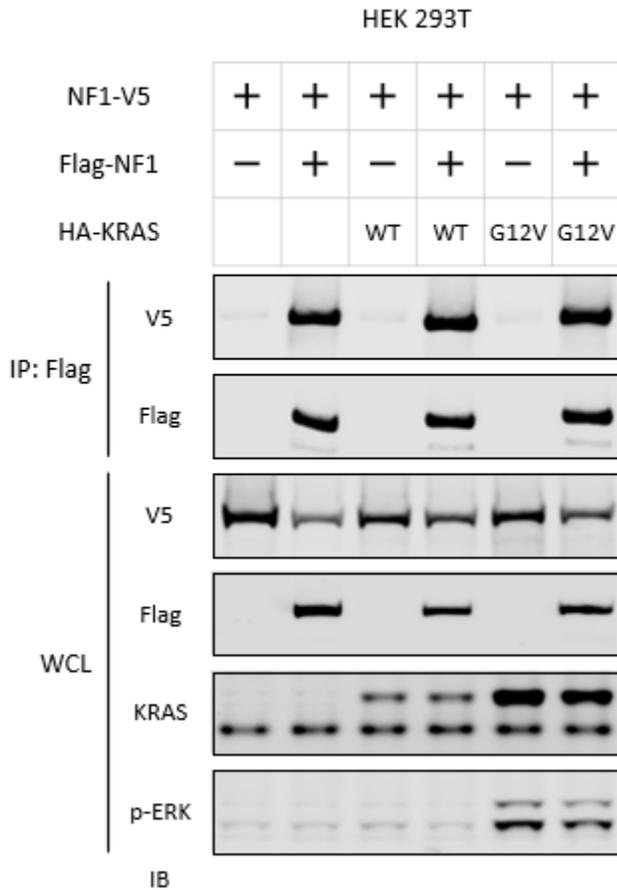
We have also published our findings of the Spred1-binding domain of NF1, which flank the Gap Related Domain (GRD) domain (N<sub>ex</sub> and C<sub>ex</sub>): “*The neurofibromin recruitment factor Spred1 binds to the GAP related domain without affecting Ras inactivation*”. Theresia Duzendorfer-Matt, Ellen L. Mercado, Karl Maly, Frank McCormick, and Klaus Scheffzek, PNAS, 113, 7497-7502, 2016.

Subtask 4: The scope of this sub-task changed when we discovered that neurofibromin is a homo-dimer.

As shown in the figure below (Figure 16), when we co-expressed Flag-tagged neurofibromin with V5 tagged neurofibromin, we were able to immune-precipitate Flag-tagged protein with V5 antibodies and vice versa. We are currently investigating whether dimeric forms of neurofibromin interacts with SPRED1 and whether dimerization affects GAP activity against RAS.

Figure 16

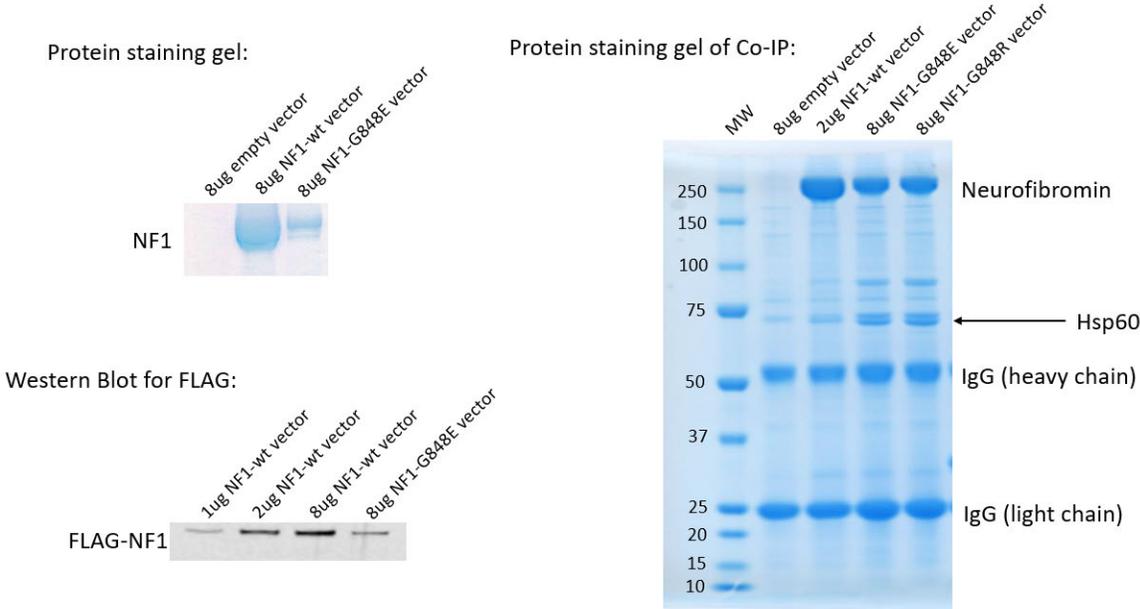
## NF1 forms dimers in 293T cells



We have also analyzed biochemical properties of pathogenic mutants of neurofibromin, focusing initially on missense mutations in a hot spot region (844-848) defined by Koczkowska and colleagues (*Am J Hum Genet.* 2018 Jan 4;102(1):69-87) as a particularly severe mutation. We expressed wild type and codon 848 mutations in 293T cells and showed that they retain similar GAP activity. However, the mutant is expressed at about one-tenth wild type levels, and appears to be unstable. We are testing the hypothesis that this mutant protein can be stabilized by drugs that inhibit proteasomal degradation.

Figure 17

*NF1* mutant expression and interacting proteins



## **Opportunities for training and professional Development**

Graduate student Evan Markegard (funded by the project) was given one-on-one mentorship by Dr. McCormick which included at least bi-weekly meetings either face-to-face or by skype to discuss his progress and goals. Evan has presented his research at lab group meetings approximately four times per year, as well as at the UCSF Helen Diller Comprehensive Cancer Center “Research in Progress” seminar series and he presented at the Cancer Center annual retreat. He also presented a research poster at the American Association of Cancer Research 108<sup>th</sup> Annual Meeting in Washington D.C (April 1-5, 2017). Evan has now completed his training and graduated successfully in 2018.

## **Dissemination of Results**

1. Evan Markegard - Poster Presentation at the 108th Annual meeting of American Association of Cancer Research (April 1-5, 2017).

Abstract 1370: EGFR-mediated Spred1 phosphorylation inhibits NF1 to sustain constitutive Ras/MAPK signaling Evan Markegard, Ellen L. Mercado, Jillian M. Silva, Jacqueline Galeas, Marena I. Trinidad, Anatoly Urisman and Frank McCormick. DOI: 10.1158/1538-7445.AM2017-1370 Published July 2017

2. Evan Markegard - Poster Presentation at the 107th Annual meeting of American Association of Cancer Research (April 16-20, 2016).

Abstract 1874: Oncogenic EGFR signaling inhibits the Spred1-NF1 interaction to sustain constitutive Ras signaling. Evan Markegard, Ellen L. Mercado, Jacqueline Galeas, Marena I. Trinidad, Anatoly Urisman and Frank McCormick. DOI: 10.1158/1538-7445.AM2016-1874 Published 15 July 2016

3. Frank McCormick – Presentation, Helen Diller Comprehensive Cancer Center Seminar Series, February 12, 2016
4. Presentations and discussions at joint lab meetings with the labs of Ophir Klein (Department of Orofacial Sciences, UCSF), Natalia Jura (Department of Cellular and Molecular Pharmacology, UCSF) and Kevin Shannon (Department of Pediatrics, UCSF).
5. Frank McCormick – Presentation, Children’s Tumor Foundation, 2016 Neurofibromatosis Conference, Austin, Texas, June 17-21, 2016
6. Frank McCormick – Presentation at University of Washington, NF Center Research Symposium, St Louis, MO, September 5, 2018

7. Frank McCormick- Presentation at 14<sup>th</sup> Annual Biomedical Research Symposium, St Jude Children's Research Hospital, Memphis, TN, October 12, 2018
8. Frank McCormick – Presentation at UCSF Comprehensive Cancer Center Seminar Series, December 14, 2018
9. Frank McCormick – Seminar at Biocenter-Innsbruck Medical University, Austria, guest of Dr Klaus Scheffzek
10. Dunzendorfer-Matt, T., **Mercado, E.L.**, Maly, K., **McCormick, F.**, Scheffzek, K. 2016. The neurofibromin recruitment factor Spred1 binds to the GAP related domain without affecting Ras inactivation. *Proc Natl Acad Sci U S A.* 113(27): 7947-79502
11. Structure of the SPRED1/neurofibromin complex and its role in Ras regulation. 2018. *In preparation*

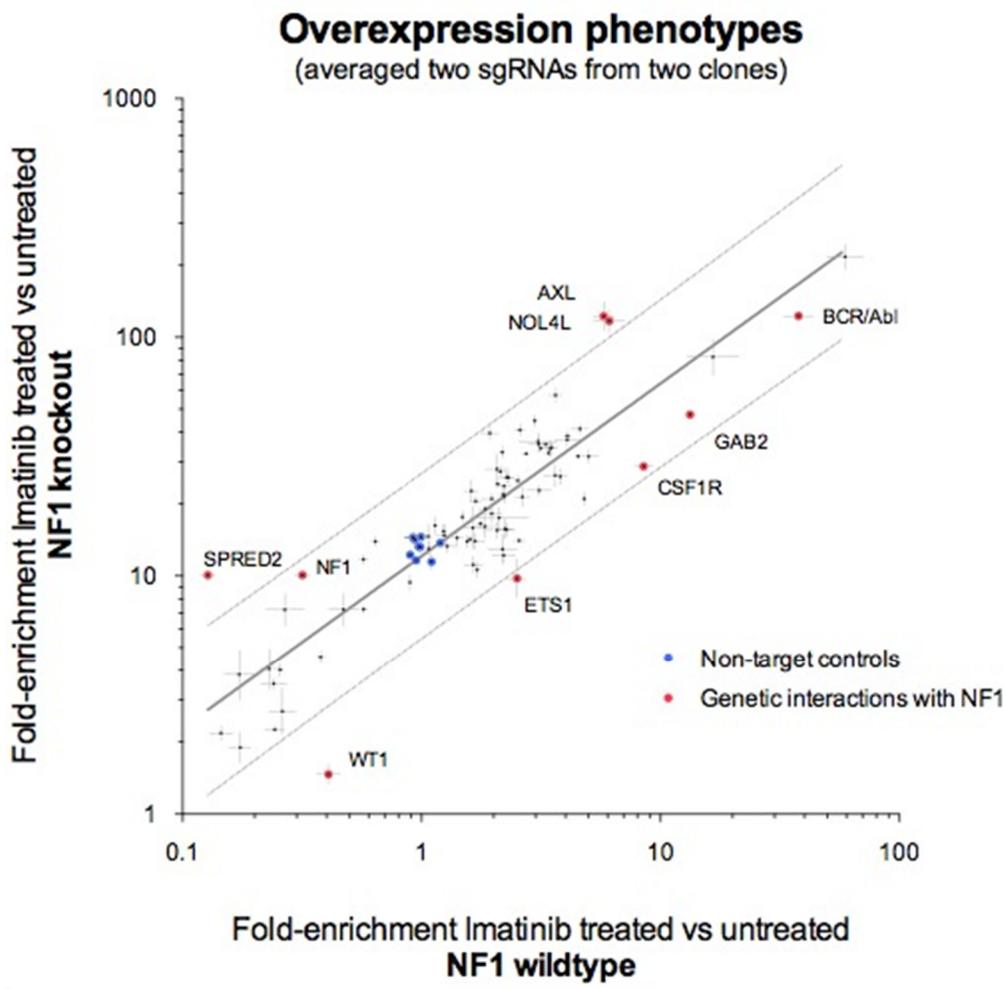
#### **4. Impact**

Nothing to report

#### **5. Changes/Problems**

While the Aims of the project remain unchanged, we will use new technologies to address these Aims. To identify proteins that regulate Neurofibromin and/or Spred, we will use functional screens in place of biochemical, mass spectrometry-based approaches. The latter approach has generated lists of proteins that appear to bind to Neurofibromin when the protein is expressed in cells in culture, but validation of these is challenging. We have been working closely with our UCSF colleagues Drs. Luke Gilbert, Michael Boettcher and Michael McManus to use CRISPRa and CRISPRi systems to identify regulatory proteins (Gilbert et al, *CELL*, 159, 647-661, 2014). We will use biochemical approaches to validate these hits, but strongly feel the primary screen should be functional. As described above, our colleagues performed a full genomic screen for genes that affect how K562 cells depend on imatinib. These cells were used because they grow in suspension, making full genome screens possible at high coverage. One of the strongest hits from the screen was Neurofibromin: over expression of this gene by CRISPRa mediated activation of the promoter, sensitized cells to the drug, whereas suppression made cells drug resistant. KRAS had exactly the opposite effect, as predicted. SPRED genes also score strongly in this system, behaving exactly as Neurofibromin (over-expression caused sensitivity, deletion caused resistance). In addition, we found that loss of Neurofibromin negated the toxic effects of SPRED over-expression, confirming that SPRED proteins (both SPRED1 and SPRED2) act through Neurofibromin. These results show that this system is ideal for screening for genes or drugs that regulate Neurofibromin or SPRED, as the cells are extremely sensitive to the activity of this pathway, and retain wild type, functional genes for the entire Ras pathway downstream of RTK signaling. We therefore expect this system to allow us to obtain candidate regulators, and will use biochemical screens in physiologically relevant cell systems to follow up on these hits, as described in the original proposal. An example of the power of this technology is shown in Figure 18.

Figure 18



**Figure 18.** Screen for genes that differentially affect NF knockout cells vs wild type cells. A small set of genes were over-expressed using CRISPRa technology, in either background, and viability measured by deep sequencing.

In this screen, over-expression of SPRED2 or NF1 had no effect on viability of NF1 deficient cells (score of 10 represents no effect in NF1 KO cells), but both reduced viability of wild type cells (score of 1 represents no effect in wild type cells), as expected. However, over-expression of the receptor tyrosine kinase AXL and the AML fusion gene partner NOL4L greatly increased viability in the NF1 knockout cells, suggesting that these genes normally suppress growth through interaction with neurofibromin. Our efforts to identify tyrosine kinase receptors that are regulated by neurofibromin (described above) will now be expanded to include AXL. We will also use these preliminary data as the basis of a full genome wide screen to identify regulators of Neurofibromin and Spred signaling.

## 6. Products

*Poster Presentation at the 107th Annual meeting of American Association of Cancer Research (April 16-20, 2016).*

Abstract 1874: **Oncogenic EGFR signaling inhibits the Spred1-NF1 interaction to sustain constitutive Ras signaling.** Evan Markegard, Ellen L. Mercado, Jacqueline Galeas, Marena I. Trinidad, Anatoly Urisman and Frank McCormick. DOI: 10.1158/1538-7445.AM2016-1874 Published 15 July 2016

*Poster Presentation at the 108th Annual meeting of American Association of Cancer Research (April 1-5, 2017).*

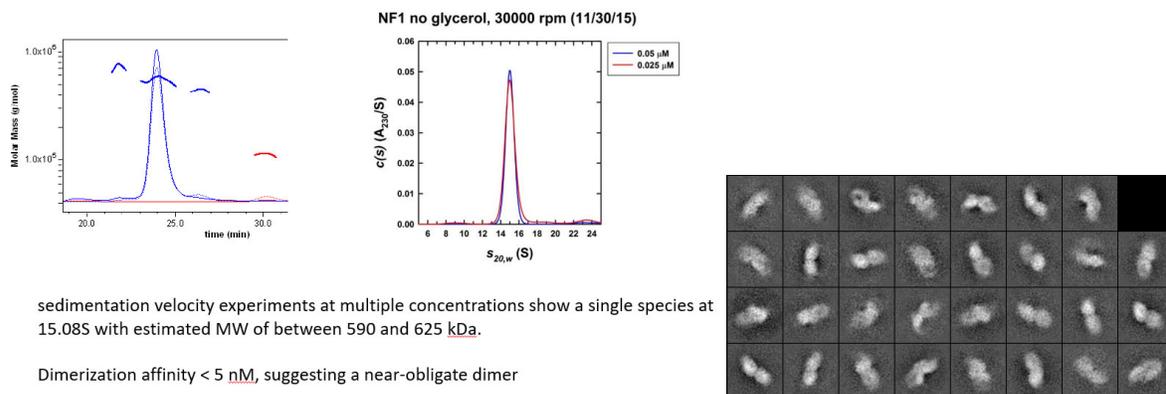
Abstract 1370: **EGFR-mediated Spred1 phosphorylation inhibits NF1 to sustain constitutive Ras/MAPK signaling.** Evan Markegard, Ellen L. Mercado, Jillian M. Silva, Jacqueline Galeas, Marena I. Trinidad, Anatoly Urisman and Frank McCormick. DOI: 10.1158/1538-7445.AM2017-1370 Published July 2017

## 7. Participants and other collaborating organizations

We are collaborating with scientists at the Frederick National Labs to solve the structure of Neurofibromin, alone or in complexes with RAS and/or Spred proteins. Full length Neurofibromin made in baculovirus vectors has been extensively purified and shown to be an obligate dimer, using size-exclusion chromatography, sedimentation analysis and electron microscopy, as shown in Figure 19.

**Figure 19**

### Neurofibromin is a dimer



**Figure 19.** Biophysical characterization of recombinant, full length neurofibromin. Left panel: size exclusion chromatography. Middle panel, sedimentation analysis. Right panel, Negative-stain electron microscopy.

**Individuals working on the project at UCSF**

***No change from last submission.***

Name:	Frank McCormick
Project Role:	PI

Name:	Evan Markegard
Project Role:	Graduate Student

Name:	Osamu Tetsu
Project Role:	Senior investigator

Name:	Ellen Mercado
Project Role:	Post Doc

**Changes in PI support since Award Made:**

Please see below for the Dr. Frank McCormick's (PI) active and pending other support:

ACTIVE

**Funding Number:** 1R35CA197709

PI: McCormick, Frank

**Project Title:** *New Ways of Targeting K-Ras*

**Performance Period:** 04/01/2016-03/31/2023

**Time Commitment:** 6.00 calendar months

**Supporting Agency:** NIH/NCI

**Contracting Officer:** Joanna Watson PhD.

**Level of Funding:**

**Project Goals:** To identify ways of suppressing the function of hyper-active Ras proteins for cancer therapy.

**Specific Aims:** i) Analysis of GTPase Activating Proteins (GAPs) that regulate Ras; ii) analysis of distinct biochemical functions of different K-Ras mutants; iii) analysis of a novel, unique function of K-Ras 4B.

**Overlap:** None

**\*No changes to report\***

**Funding Number:** U01 CA168370

PI: McManus, Michael; McCormick, Frank;

Weissman, Jonathan

**Project Title:** Bay Area Cancer Target Discovery and Development Network

**Performance Period:** 05/01/12-04/30/2017

**Time Commitment:** 0.24 calendar months

**Supporting Agency:** NIH/NCI

**Contracting Officer:** Daniela Gerhard PhD.,

**Level of Funding:**

**Project Goals:** To bridge the gap between the enormous volumes of data generated by the comprehensive molecular characterization of a number of cancer types and the ability to use these data for the development of human cancer therapeutics. Our end goal is to generate game-changing reagents, and data valuable for the development of cancer therapeutics

**Specific Aims:** i) Develop next generation EXPAND libraries targeting cancer specific genetic alterations; ii) Identify recurrently mutated genes that regulate oncogenic pathways and drug responses; iii) Produce genetic interaction maps to uncover pathway relationships between candidate drivers.

**Overlap:** None

***\*Grant has closed out/performance period end date of 4/30/17\****

**Funding Number:** N/A

**PI:** McCormick, Frank

**Project Title:** Preclinical Evaluation of Signaling Pathways Involved in Cancer Malignancy

**Performance Period:** 04/01/2017 – 03/31/2019

**Time Commitment:** 0.12 calendar

**Supporting Agency:** Daiichi-Sankyo Company, Limited

**Contracting Officer:** Kosaku Fujiwara, PhD, VP, Oncology Research Laboratories,  
fujiwara.kosaku.t2@daiichisankyo.co.jp,

**Level of Funding:**

**Project Goals:** Analysis of the Ras pathway in epithelial and mesenchymal cells.

**Specific Aims:** i) Determine how EMT affects MAPK and P13K pathways downstream of oncogenic Ras; ii) identify genes required for survival in epithelial and mesenchymal cells.

**Overlap:** None

***\*New performance period\****

**Funding Number:** N/A

**PI:** McCormick, Frank

**Project Title:** New ways of Treating Pancreatic Cancer Based on Reversing K-Ras-Mediated Stemness

**Time Commitment:** 1.2 calendar months

**Supporting Agency:** Lustgarten Foundation

**Contracting Officer:** Kerri Kaplan.

**Level of Funding:**

**Project Goals:** Preclinical analysis of humanized anti-LIF monoclonal antibodies, using several models of pancreatic cancer, with the intention of moving these antibodies into clinical testing based on the results of these experiments

**Specific Aims:** i) generate human and mouse pancreatic cancer cells with LIF stably knocked down and elucidate the effects of this knock down; ii) validate LIF as a target of intervention for mouse pancreatic cancer in a syngenic mouse model; iii) compare the functions of LIF and IL6 in pancreatic cancer cells; iv) evaluate the therapeutic effects of LIF neutralizing antibody; v) elucidate the signaling pathways specifically driven by LIF and not other IL6 cytokine family members in K-Ras driven pancreatic cancers; vi) pre-clinically evaluate prostratin pro-drug in mouse models of pancreatic cancer; vii) investigate the mechanism of prostratin's tumor suppressing action on pancreatic cancer cell lines.

**Overlap:** None

***\*Grant closed out/performance period end date 12/31/2018\****

**Funding Number:** N/A

**PI:** McCormick, Frank

**Project Title:** Targeting KRAS Mutant Lung Cancers

**Performance Period:** 08/01/2015—07/31/2018

**Time Commitment:** 1.2 calendar months

**Supporting Agency:** Subcontract through Massachusetts General Hospital

**Contracting Officer:** Maida Broudo,

**Level of Funding:**

**Project Goals:** Preclinical evaluation of prostratin and LIF for treating lung adenocarcinoma.

**Specific Aims:** i) Test the specific therapeutic effects of prostratin or its analogs on K-Ras driven lung cancers; 2) Examine the expression patterns of LIF in tumors and stromal cells.

**Overlap:** None

***\*Grant closed out/performance period end date of 7/31/2018\****

PENDING

**McCormick** 7/1/2018 – 6/30/2021 1.20 calendar  
American Association for Cancer Research(AACR), UC Santa Cruz(UCSC)  
Lung Cancer Early Detection Using Long Noncoding RNA  
The major goals of this project are to determine whether long non-coding RNAs can be used for early detection of sub-types of KRAS-driven lung adenocarcinomas.

**McCormick** 4/1/2019 – 3/31/2024 1.80 calendar  
NIH Natl Inst. Child Health & Human Dev.(NIH)  
The role of Lztr1 during embryonic development and its importance in developmental diseases  
The major goals of this project are to identify the molecular mechanisms by which loss of LZTR1 leads to development of Noonan Syndrome and schwannomatosis.

**8. Special Reporting Requirements – None**

## 9. Appendix

### A: Manuscript References

1. Gilbert L.A., et al. 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159:647–661. doi: 10.1016/j.cell.2014.09.029
2. Dunzendorfer-Matt, T., **Mercado, E.L.**, Maly, K., **McCormick, F.**, Scheffzek, K. 2016. The neurofibromin recruitment factor Spred1 binds to the GAP related domain without affecting Ras inactivation. *Proc Natl Acad Sci U S A.* 113(27): 7947-79502
3. Boettcher, M., Tian, R., Blau, J.A., **Markegard, E.**, Wagner, R.T., Wu, D., Mo, X., Biton, A., Zaitlen, N., Fu, H., **McCormick, F.**, Kampmann, M., McManus, M.T. 2018. Dual gene activation and knockout screen reveals directional dependencies in genetic networks. *Nature Biotechnology* 36 (2): 170-178.
4. Koczkowska, M., et al. 2018. Genotype-Phenotype Correlation in NF1: Evidence for a More Severe Phenotype Associated with Missense Mutations Affecting NF1 Codons 844–848. *Am J Hum Genet* 102(1): 69-87.
5. Structure of the SPRED1/neurofibromin complex and its role in Ras regulation. 2018. *In preparation*

## **B: Poster Presentations**

1. **Evan Markegard, Ellen L. Mercado**, Jacqueline Galeas, Marena I. Trinidad, Anatoly Urisman, **Frank McCormick**. Oncogenic EGFR signaling inhibits the Spred1-NF1 interaction to sustain constitutive Ras signaling. [abstract]. In: Proceedings of the 107th Annual Meeting of the American Association for Cancer Research; 2016 Apr 16-20; New Orleans, LA. Philadelphia (PA):AACR; Cancer Res 2016;76(14 Suppl):Abstract nr 1874
2. **Evan Markegard, Ellen L. Mercado**, Jillian M. Silva, Jacqueline Galeas, Marena I. Trinidad, Anatoly Urisman, **Frank McCormick**. EGFR-mediated Spred1 phosphorylation inhibits NF1 to sustain constitutive Ras/MAPK signaling [abstract]. In: Proceedings of the American Association for Cancer Research Annual Meeting 2017; 2017 Apr 1-5; Washington, DC. Philadelphia (PA): AACR; Cancer Res 2017;77(13 Suppl):Abstract nr 1370. doi:10.1158/1538-7445.AM2017-1370.

## **C: Presentations**

1. Frank McCormick – Seminar at Biocenter-Innsbruck Medical University, Austria, guest of Dr Klaus Scheffzeck
2. Frank McCormick – Presentation, Helen Diller Comprehensive Cancer Center Seminar Series, February 12, 2016
3. Presentations and discussions at joint lab meetings with the labs of Ophir Klein (Department of Orofacial Sciences, UCSF), Natalia Jura (Department of Cellular and Molecular Pharmacology UCSF) and Kevin Shannon (Department of Pediatrics, UCSF).
4. Frank McCormick – Presentation, Children’s Tumor Foundation, 2016 Neurofibromatosis Conference, Austin, Texas, June 17-21, 2016
5. Frank McCormick – Presentation at University of Washington, NF Center Research Symposium, St Louis, MO, September 5, 2018
6. Frank McCormick- Presentation at 14<sup>th</sup> Annual Biomedical Research Symposium, St Jude Children’s Research Hospital, Memphis, TN, October 12, 2018
7. Frank McCormick – Presentation at UCSF Comprehensive Cancer Center Seminar Series, December 14, 2018



Published in final edited form as:

Cell. 2014 October 23; 159(3): 647–661. doi:10.1016/j.cell.2014.09.029.

## Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation

Luke A. Gilbert<sup>1,2,3,4,10</sup>, Max A. Horlbeck<sup>1,2,3,4,10</sup>, Britt Adamson<sup>1,2,3,4</sup>, Jacqueline E. Villalta<sup>1,2,3,4</sup>, Yuwen Chen<sup>1,2,3,4</sup>, Evan H. Whitehead<sup>1,3,5,9</sup>, Carla Guimaraes<sup>6</sup>, Barbara Panning<sup>7</sup>, Hidde L. Ploegh<sup>6</sup>, Michael C. Bassik<sup>1,2,3,4,8</sup>, Lei S. Qi<sup>1,3,5,9</sup>, Martin Kampmann<sup>1,2,3,4,\*</sup>, and Jonathan S. Weissman<sup>1,2,3,4,\*</sup>

<sup>1</sup>Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA

<sup>2</sup>Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94158, USA

<sup>3</sup>California Institute for Quantitative Biomedical Research, San Francisco, CA 94158, USA

<sup>4</sup>Center for RNA Systems Biology, University of California, San Francisco, San Francisco, CA 94158, USA

<sup>5</sup>Center for Systems and Synthetic Biology, University of California, San Francisco, San Francisco, CA 94158, USA

<sup>6</sup>Department of Biology, Whitehead Institute for Biomedical Research and Massachusetts Institute of Technology, Cambridge, MA 02142, USA

<sup>7</sup>Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA

### SUMMARY

While the catalog of mammalian transcripts and their expression levels in different cell types and disease states is rapidly expanding, our understanding of transcript function lags behind. We present a robust technology enabling systematic investigation of the cellular consequences of

---

© 2014 Elsevier Inc. All rights reserved.

\*Correspondence: martin.kampmann@ucsf.edu and jonathan.weissman@ucsf.edu.

<sup>8</sup>Present Address: Department of Genetics Stanford University, Palo Alto, CA, 94305, USA

<sup>9</sup>Present Address: Department of Bioengineering, Stanford University, Palo Alto, CA, 94305, USA

<sup>10</sup>Co-first author

### AUTHOR CONTRIBUTIONS

BA, LAG, MAH, MK, JSW were primarily responsible for the conception, design, and interpretation of the experiments and wrote the manuscript. BA, YC, LAG, MAH, MK, JEV, and EHW conducted experiments. LAG cloned dCas9 chimeras and sgRNA expression constructs, constructed cell lines, carried out tiling screens, and conducted validation experiments. MAH designed libraries, carried out tiling and genome-scale screens, and analyzed screen data. BA constructed the inducible cell line and conducted all inducible experiments. CG and HLP contributed to CTx-DTA studies. BP contributed to XIST studies. MCB and LSQ contributed to the conception and interpretation of the experiments.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

repressing or inducing individual transcripts. We identify rules for specific targeting of transcriptional repressors (CRISPRi), typically achieving 90–99% knockdown with minimal off-target effects, and activators (CRISPRa) to endogenous genes via endonuclease-deficient Cas9. Together they enable modulation of gene expression over a ~1000-fold range. Using these rules, we construct genome-scale CRISPRi and CRISPRa libraries, each of which we validate with two pooled screens. Growth-based screens identify essential genes, tumor suppressors and regulators of differentiation. Screens for sensitivity to a cholera-diphtheria toxin provide broad insights into the mechanisms of pathogen entry, retro-translocation and toxicity. Our results establish CRISPRi and CRISPRa as powerful tools that provide rich and complementary information for mapping complex pathways.

---

## INTRODUCTION

Dramatic advances in sequencing technology have catalogued a universe of transcribed loci—greatly exceeding the number of canonical protein-coding open reading frames (ORFs)—which collectively are responsible for carrying out the instructions encoded by the genome (Djebali et al., 2012). A central challenge now is to understand the biological role of these transcripts and how quantitative differences in their expression define cellular states in normal development and in disease. Despite intense efforts, the function of many protein-coding genes remains poorly defined. Even less is known about the biological roles of most non-canonical transcripts such as enhancer RNAs, upstream antisense RNAs, lncRNAs, or other intergenic RNAs (Cech and Steitz, 2014). Efforts to address this deficiency in our knowledge would be greatly aided by techniques that are capable of dynamically and precisely controlling the expression of individual transcripts.

One way to explore the function of genes is to disrupt their expression through repression. The dominant tool for programmed knockdown of mRNAs is RNA interference (RNAi) (Chang et al., 2006). However, RNAi has pervasive problems with off-target effects, which can be especially confounding in the context of large-scale screens (Adamson et al., 2012; Jackson et al., 2003; Sigoillot et al., 2012). Additionally, because RNAi is mediated by cytoplasmic argonaute proteins, gene silencing through this approach is best suited to depletion of cytosolic mRNA targets.

An alternative emerging strategy is the use of programmable genome editing methods that permanently delete or modify DNA using designable, sequence-specific endonucleases such as zinc finger, transcription activator-like effector (TALE) nucleases, or CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9 (CRISPR-associated protein 9) proteins (Gaj et al., 2013; Sander and Joung, 2014). A series of elegant studies recently exploited the readily programmable nature of Cas9, in which the specificity is determined by a short guide (sg)RNA, to enable genome-scale loss-of-function screens (Koike-Yusa et al., 2014; Shalem et al., 2014; Wang et al., 2014). These studies established CRISPR-mediated cutting as a powerful screening technology complementary to RNAi and haploid mutagenesis screens (Carette et al., 2009). Nonetheless, screening approaches based on genome editing are currently focused on loss-of-function studies involving irreversible frameshift disruptions, limiting their utility for the study of essential genes and long

noncoding RNAs. Additionally, double-stranded DNA breaks can be cytotoxic (Huang et al., 1996; Jackson, 2002). Finally, indels formed from error-prone DNA repair are often short and in-frame, which could limit the ability to disable all of the alleles of a gene.

A programmable DNA binding protein that can recruit an effector domain to turn transcription on and off in a dynamic and quantitative manner offers, in principle, a more flexible tool for interrogating the many transcripts in complex genomes. Pioneering experiments with designed chimeric zinc finger and TALE proteins fused to transcription effector domains demonstrate that such an approach can modulate transcription of endogenous genes (Beerli et al., 1998, 2000; Zhang et al., 2011). However, as each transcript target requires a unique fusion protein, expanding these methods to genome-scale is arduous.

Recently, we and others have used catalytically inactive Cas9 (dCas9) fusion proteins guided by gene-specific sgRNAs to localize effector domains to specific DNA sequences to either repress (CRISPRi) or activate (CRISPRa) transcription of target genes (Gilbert et al., 2013; Sander and Joung, 2014). To date, a small number of sgRNAs have been tested, leaving unanswered whether CRISPRi/a is a feasible strategy for globally interrogating gene function and, if so, how best to target a gene to activate or repress transcription while minimizing off-target effects.

Here, we describe the development and application of a method for high-specificity, genome-scale modulation of transcription of endogenous genes in human cells using CRISPRi/a. To accomplish this, we first performed a saturating screen in which we tested the activity of every unique sgRNA broadly tiling around the transcription start sites (TSSs) of 49 genes known to modulate cellular susceptibility to ricin (Bassik et al., 2013). From this, we extracted distinct rules for regions where either CRISPRi or CRISPRa maximally changes the expression of endogenous genes, as well as rules for predicting off-target effects, providing an algorithm to design two genome-scale libraries targeting each gene with 10 sgRNAs. We validated these libraries by screening for genes that control cell growth and response to a chimeric cholera/diphtheria fusion toxin (CTx-DTA) (Guimaraes et al., 2011). These experiments demonstrate that our CRISPRi/a screening platform is robust, showing high reproducibility and activity with undetectable intrinsic toxicity. More generally, we establish that transcriptional repression is inducible, reversible, and can target essential genes. We demonstrate that we can use CRISPRi and CRISPRa to control transcript levels for endogenous genes across a high dynamic range. We also provide extensive evidence that properly designed CRISPRi reagents are highly specific. As such, these methods represent transformative tools for defining transcript function across the breadth of transcripts encoded by the human genome.

## RESULTS

### A High-Throughput Tiling Screen Defines Rules for CRISPRi Activity at Endogenous Genes

CRISPRi can repress transcription by directly blocking RNA polymerase activity (dCas9) or through effector domain-mediated transcriptional silencing (dCas9-KRAB) (Gilbert et al.,

2013; Qi et al., 2013). In order to better understand and optimize CRISPRi activity, we used a pooled high-throughput screen to define rules that determine CRISPRi repression of endogenous genes. We targeted 49 genes that we had previously shown to modulate cellular susceptibility to the AB toxin ricin (Bassik et al., 2013). The extent of gene repression for these genes typically has a monotonic relationship with the ricin resistance phenotype, allowing us to use a ricin resistance score calculated by monitoring sgRNA frequencies in a pooled screen to indirectly measure transcriptional repression.

Using massively parallel oligonucleotide synthesis, we generated a library of sgRNAs that tile the DNA in a 10-kilobase window around the TSS of these 49 genes (54,810 total sgRNAs) (Bassik et al., 2009) (Figure 1A). We also included 1,000 negative control sgRNAs derived from scrambled sequences corresponding to the same windows.

We packaged this tiling library of sgRNAs into lentiviral particles and transduced K562 human myeloid leukemia cells stably expressing dCas9 or a dCas9-KRAB fusion protein, which we have previously described (Gilbert et al., 2013). We harvested populations of cells expressing the library either at the outset of the experiment, after growth under standard conditions, or following ricin treatment. We then counted the frequency of each sgRNA in the library in each sample using deep sequencing to determine how each sgRNA in the library modulates cell growth and cellular susceptibility to ricin phenotypes. We defined these phenotypes quantitatively as  $\gamma$  and  $\rho$ , respectively (See Figure S1A and (Kampmann et al., 2013)).

Many sgRNAs potently repress gene expression, as evidenced by their impact on ricin sensitivity (Figure 1B and Figure S2A). Plotting this data for all 49 genes showed that active sgRNAs cluster around or just downstream from the TSS of each gene for dCas9-KRAB and dCas9, respectively (Figure 1C). We saw that strong CRISPRi activity is obtained by targeting dCas9-KRAB to a window of DNA from -50 to +300 bp relative to the TSS of a gene, with a maximum in the ~50–100 bp region just downstream of the TSS (Figure 1C–D). This suggested that optimal activity leverages the combined activity of dCas9 interference along with repression from the KRAB domain. We also observed that sgRNAs with protospacer lengths of 18–21 base pairs were significantly more active than sgRNAs containing longer protospacers (Figure S2B). Nucleotide homopolymers had a strongly negative effect on sgRNA activity (Figure S2D). However, neither the DNA strand that was targeted nor the sgRNA GC content across a broad range strongly correlated with sgRNA activity (Figure S2C and S2E).

To evaluate the feasibility of genome-scale genetic screens based on CRISPRi, we compared the strength of phenotypes obtained with CRISPRi to our previously published shRNA data. We applied the rules described above to randomly subsampled data from our sgRNA tiling library, picking sets of 10 or 24 sgRNAs. We calculated a normalized phenotype z-score by dividing mean phenotypes for each gene by the standard deviation of sgRNA phenotypes from the non-targeting control set (Figure S1B). We see significant ricin phenotypes for each of the 49 genes. Moreover, in virtually every case the normalized ricin phenotype z-score or p-value is stronger (in many cases far stronger) than seen with a comparably-sized shRNA library (generated by sub-sampling our published data) (Figure 1E and Figure S2F).

## CRISPRi Transcriptional Silencing is Highly Sensitive to Mismatches between the Target DNA Site and the sgRNA

To assess CRISPRi off-target activity at endogenous genes, we selected a set of 30 sgRNAs from our tiling library (6 sgRNAs/gene targeting 5 genes). For each of these sgRNAs, we tested the activity of a series of derivative sgRNAs with a variable number and position of mismatches (Figure 2). This experiment allowed us to measure the relative amount of gene repression for sgRNAs with or without mismatch base pairing targeting the same DNA locus. We found that even a single mismatch at the 3' end of the protospacer decreased CRISPRi activity on average, while combinations of mismatches that pass our off-target filter abolished activity (Figure 2, Figure S3, and Extended Experimental Procedures). From this analysis, we concluded that properly designed CRISPRi sgRNAs have minimal off-target transcriptional repression activity.

## A High-Throughput Tiling Screen Defines Rules for CRISPRa Activity at Endogenous Genes

We recently developed an improved CRISPRa method, termed sunCas9, in which expression of a single sgRNA with one binding site is sufficient to robustly activate transcription (Tanenbaum et al.) In the sunCas9 system, a single dCas9 fusion protein bound to DNA recruits multiple copies of the activating effector domain, thus amplifying our ability to induce transcription (Figure 3A).

To define rules for optimal CRISPRa sgRNA design, we used our tiling library, which targets genes capable of modulating cellular sensitivity to ricin. We previously showed for several of the genes in this tiling library that knockdown and plasmid overexpression resulted in opposite ricin phenotypes (Bassik et al., 2013). For example, knockdown of *SEC23B* sensitized cells to ricin, whereas *SEC23B* overexpression desensitized cells to ricin. These observations suggested that we should be able to observe reversed phenotypes in this tiling screen arising from CRISPRa activity.

We transduced K562 cells stably expressing the sunCas9 system (Figure 3A) with the sgRNA tiling library and screened for ricin phenotypes as described for CRISPRi above. Analysis of data for individual genes or averaged data for all 49 genes demonstrated that many sgRNAs for each gene affected ricin resistance (Figure 3B and Figure S4A–B). Our negative control sgRNAs showed very little activity and were not correlated between biological replicate screens, suggesting that CRISPRa activity is specific. We observed a peak of active sgRNAs for CRISPRa at –400 to –50 bp upstream from the TSS (Figure 3B). This activity pattern fits with a model in which each VP16 domain can bind the mediator complex and recruit basal transcription machinery, activating transcription when spaced appropriately from a TSS (Mittler et al., 2003). With this system, we have shown we can turn on genes that are poorly expressed and increase the expression of well-expressed genes (Figure 3E). Overall, our CRISPRi/a tiling screens provide rules for how CRISPRi/a controls expression of endogenous genes.

## An Allelic CRISPRi/a Series of Transcript Activation and Repression Shows that Protein Abundance Dynamically Modulates the Cellular Response to Ricin

For many genes, we do not know how the relative abundance of the encoded protein relates to its function. We observed a marked anti-correlation in our ricin screens between CRISPRa phenotypes and CRISPRi phenotypes for individual genes (Figure 3C). As the genes targeted by our tiling library were selected based on a knockdown phenotype, all genes showed phenotypes in the CRISPRi screen, but only a subset showed phenotypes in the CRISPRa screen.

To validate results from both the CRISPRi and CRISPRa tiling screens, we selected an allelic series of sgRNAs by phenotype from the screen and re-tested each sgRNA individually (38 sgRNAs targeting 4 genes). Our results show that our CRISPRi/a screens produced reliable phenotype scores, robustly reproduced upon retesting, and that CRISPRi/a can activate and repress the transcription of endogenous genes over a wide dynamic range (up to ~1000-fold) (Figure 3D–E), enabling systematic interrogation of how gene dosage controls cellular functions of interest.

### A Robust and Highly Specific Genome-Scale CRISPRi Screening Platform

The results of our tiling CRISPRi screen established our ability to pick active sgRNAs with low off-target activity and provided a set of rules enabling us to design a robust genome-scale sgRNA library. We chose a library size of 10 sgRNAs/gene for the following reasons. Over half of the sgRNAs conforming to these rules gave clear ricin phenotypes. For a library with 10 sgRNAs/gene, 94% of the genes would thus have 2 or more highly active sgRNAs. Finally, computational sub-sampling of the phenotypic data from our tiling library data to 10 sgRNAs/gene and calculation of z-scores for hit genes indicated that a library with 10 sgRNAs/gene would reliably detect hit genes (Figure 1E).

We synthesized and cloned a genome-scale CRISPRi sgRNA library targeting 15,977, human protein-coding genes (10 sgRNAs/TSS, targeting 20,898 TSS) with 11,219 non-targeting control sgRNAs for a total of 206,421 sgRNAs (Table S2). To evaluate this library, we first screened for genes essential for cell growth in K562 cells. Briefly, K562 cells stably expressing dCas9-KRAB were transduced in replicate with the entire genome-scale library, and each replicate was grown for 10 days at a minimum library coverage of 3,750 cells/sgRNA in a single spinner flask.

To characterize our screening methodology and library design, we examined the correlation between screen replicates and our negative control distributions. Individual sgRNAs reproducibly showed dramatic depletion (up to 256-fold) over a 10-day screen, demonstrating that individual sgRNAs can have profound effects on cell growth (Table S2) (Figure 4A). The distribution of our negative-control sgRNAs was very narrow with little correlation between replicates (Spearman  $R = 0.036$ ), suggesting that the off-target activity of these controls is very low (Figure 4A). Indeed, 99.7% of our negative controls had no detectable activity. The observed specificity is consistent with our previously published RNA-seq data (Gilbert et al., 2013).

To further explore the prevalence of off-target effects, we examined two classes of genes not expected to show any on-target activity in our screen: olfactory receptors and genes on the Y chromosome. The sgRNAs that target these genes were designed and picked in the same manner as the rest of library; however, olfactory receptors should not be expressed in this cell type and, as K562 cells are derived from a female donor, sgRNAs that target genes on the Y chromosome lack a DNA target. As with the negative controls, these genes show no phenotype on average and exhibit very little correlation between replicates (Spearman  $R = 0.057$  for olfactory genes and  $-0.052$  for Y-targeting) (Figure 4A). We also observed no evidence of non-specific toxicity due to expression of dCas9-KRAB and our sgRNA library in K562 cells suggesting that dCas9 bound to the genome is not toxic under these conditions (Figure 4B). Thus, CRISPRi is highly specific and non-toxic.

To identify hit genes in this screen, we used a metric of average growth phenotype ( $\gamma$ ) for the top three sgRNAs for each gene (see Experimental Procedures and Table S3). Among the top hits were genes involved in essential cellular functions, including translation, transcription and DNA replication (Figure 4C and Figure S5A) (Huang et al., 2009a, 2009b), thus validating our approach as a screening platform.

### A Genome-Scale CRISPRa Screening Platform

The results of our CRISPRa tiling screen established our ability to confidently measure gene phenotypes resulting from inducing expression with single sgRNAs. As with the CRISPRi tiling screens, our data enabled the development of a set of rules that allowed construction of a genome-scale CRISPRa library. Many of these rules overlapped with those of CRISPRi (e.g., sgRNA length and sequence preferences). A key difference is that the optimal window for targeting sgRNAs for CRISPRa lies upstream of the TSS (Figure 3B). We therefore constructed an independent CRISPRa library, designing 10 sgRNAs between 400 to 50 base pairs upstream of each TSS for 15,977 human genes, along with 5,968 non-targeting control sgRNAs, for a total of 198,810 sgRNAs.

We evaluated our CRISPRa platform in a screen for genes that affect cell growth when induced in K562 cells constitutively expressing the sunCas9 system. Replicate screens were conducted as described above. The magnitude of growth defects seen in our CRISPRa screen was comparable to that of the above CRISPRi screen, although fewer sgRNAs caused a growth phenotype (Figure S5B and Table S2). We analyzed control sgRNAs with no genomic target or Y chromosome targets and found minimal phenotypes, which lacked substantial correlation between experimental replicates (Spearman  $R = 0.155$  and  $R = 0.010$ , respectively), indicating that the phenotype distribution observed in non-targeting controls was primarily a result of stochastic noise rather than off-target effects. Furthermore, the fraction of cells expressing sgRNAs and the sunCas9 system was stable over the course of the experiment, indicating that there was no general toxicity associated with the CRISPRa platform (Figure 4B). These data suggest that CRISPRa is also specific and non-toxic.

### Defining Regulators of Survival and Differentiation in Human Cells by CRISPRa

We then investigated the genes whose induction caused cells to deplete over the course of our CRISPRa screen. We scored genes by the average  $\gamma$  of the three most active sgRNAs as

above, and compared these phenotypes to those observed in the CRISPRi screen (Figure 4D and Table S3). The results from the two screens had little overlap, suggesting that few genes are both essential and toxic upon induction, and that wild-type expression levels of genes are generally optimal for K562 growth. Whereas CRISPRi hits are naturally limited to expressed genes, CRISPRa hits included genes across a broad range of endogenous expression levels (Figure S5D). We observed that the majority of genes that inhibited growth in the CRISPRa screen fell into three overlapping classes.

The first class was tumor suppressor genes: 18 of the top 50 genes, including six of the top seven, are known to have potent tumor suppressor activity (Vogelstein et al., 2013; Zhao et al., 2013). These genes include p53-related protein *TP73*, cell cycle inhibitors *CDKN1C* (p57) and *CDKN1A* (p21), apoptotic factors *BAK1* and *BCL2L11* (BIM), and chromatin remodeling factor *ARID1A* (Figure 4D and Table S4). Gene set enrichment analysis (GSEA) confirmed this observation, highlighting several genes important in the intrinsic pathway of apoptosis or in chronic myeloid leukemia (CML) homeostasis consistent with the origin of K562 cells as a clonal isolate from a CML blast crisis (ATCC) (Figure 4E). Similarly, top gene ontology annotations included “positive regulation of apoptosis” and “regulation of cell cycle” (Figure S5C). While tumor suppressors are classically considered to be mutated early in cancer progression (Vogelstein et al., 2013), these results demonstrate that many potential tumor suppressor genes remain functional but down-regulated, and suggest that CRISPRa can be used to pinpoint intact pathways and vulnerabilities in tumor cells.

Transcription factor families with well-established roles in tissue development and differentiation represent another class of growth hits, accounting for 16 of the top 50 genes (K562 cells have known potential to undergo erythroid differentiation). These genes include CCAAT/Enhancer-binding proteins (CEBP), Homeobox genes, Forkhead box genes, Ikaros family zinc finger proteins, and hematopoietic differentiation factor *SPI1* (PU.1) (Figure 4D and Table S4) (Spitz and Furlong, 2012). This observation is reflected in enriched annotations relating to multicellularity, cell differentiation, and development (Figure S5C).

The complementary nature of the CRISPRi and CRISPRa screens is nicely illustrated by results from two gene pairs (*SPI1/GATA1* and *CEBPA/CEBPG*) in which one member of each pair inhibits the function of the other. *GATA1* and *CEBPG* were strong hits in the CRISPRi screen, consistent with their roles as inhibitors of myeloid differentiation. By contrast, both *SPI1* and *CEBPA* were robust hits in our CRISPRa activation screen. These observations are consistent with the inhibitory functions of *SPI1* and *CEBPA*: silencing of *CEBPA* leads to de-repression of *CEBPG* (Alberich-Jordà et al., 2012) and the protein encoded by *SPI1* (PU.1) is a direct binding partner of GATA-1 and inhibits its transcriptional activity (Zhang et al., 2000).

Finally, several hit genes have key roles in mitosis. *PLK4* controls centrosome duplication, and overexpression of the gene in U2OS cells leads to increased centriole number (Habadanck et al., 2005). The proteins encoded by *KIF18B* and *KIF2C* form a complex that destabilizes microtubules during mitosis (Tanenbaum et al., 2011).

Overall, the results from our paired CRISPRi/a growth screens demonstrate that complementary information can be obtained by loss- and gain-of-function genetic screens, and highlight the utility of the platform for future studies into tumor biology and cell differentiation.

### **Dynamically Controlling Gene Expression with CRISPRi**

The ability to reversibly tune the expression of select transcripts would be a powerful tool for evaluating transcript function. To evaluate the applicability of CRISPRi to this purpose, we cloned a lentiviral expression construct that places an optimized KRAB-dCas9 fusion protein under the control of a doxycycline-inducible promoter (Figure 5A–B). Induced expression of KRAB-dCas9 robustly depletes transcript levels from sgRNA-targeted genes (Figure 5C and S5E). To further assess dynamic control of CRISPRi, we inducibly repressed several genes identified in our genome-scale CRISPRi growth screen (Figure S5G). Cells that express sgRNAs targeting these essential genes showed almost no growth phenotype in the absence of doxycycline, but rapidly and robustly disappeared from the population upon addition of doxycycline (Figure 5D). Additionally, gene repression and resulting phenotypes were reversible (Figure 5C and S5E–F), indicating that KRAB-dCas9 does not create a permanently repressive chromatin state at targeted promoters.

To test our ability to dynamically control expression of essential genes on a larger scale, we cloned a sublibrary targeting 426 manually curated genes (10 sgRNAs/TSS or 4,923 targeting sgRNAs plus 750 non-targeting controls). This library was transduced into K562 cells stably expressing our inducible KRAB-dCas9 fusion protein, and cell growth effects were then evaluated in the presence and absence of doxycycline. Only 4 sgRNAs were depleted strongly in the absence of doxycycline; however, with induction of KRAB-dCas9, many sgRNAs were strongly depleted (Figure 5E). Negative control sgRNAs again produced a narrow distribution of phenotypes with little correlation between biological replicates with or without doxycycline. Additionally, we found no evidence that targeted KRAB-dCas9 generally impedes cell growth (Figure 5F). Taken together, these results demonstrate CRISPRi is non-toxic, inducible and reversible.

### **A Genome-Scale CRISPRi Screen Reveals Pathways and Complexes that Govern Response to Cholera and Diphtheria Toxin**

To test the performance of our CRISPRi approach for detecting genes controlling a more complex cellular phenotype, we performed a genome-scale CRISPRi screen for genes that modulate sensitivity to a chimeric toxin composed of the diphtheria toxin catalytic A subunit covalently linked to cholera toxin (CTx-DTA, Figure 6A). This chimera had been previously developed to provide a growth readout for cholera intoxication (Guimaraes et al., 2011). Some aspects of both cholera and diphtheria toxin entry and toxicity are well characterized but open questions remain. The cell surface receptor for cholera toxin is the GM1a ganglioside (Van Ness et al., 1980). After endocytosis, the toxin traffics via the Golgi to the endoplasmic reticulum (ER), from which it retro-translocates into the cytosol, possibly through the ER-associated degradation (ERAD) machinery. Once in the cytosol, the DTA moiety ADP-ribosylates the diphthamide residue in Elongation Factor 2, halting translation and killing the cell (Figure 6A).

K562 cells expressing the CRISPRi sgRNA library and dCas9-KRAB were either grown under standard conditions or treated with several pulses of CTx-DTA over the course of 10 days. We observed highly correlated enrichment and depletion of many sgRNAs between replicates, indicating that CRISPRi can identify genes that modulate both resistance and sensitivity to a selective pressure (Table S2).

We ranked genes by the average phenotype of their three strongest sgRNAs (Table S3, Figure 6B, and Figure S6). GSEA revealed that KEGG pathways enriched for top protective hit genes were “Infection with *Vibrio cholerae*” and “Glycosphingolipid biosynthesis, ganglio-series” (Figure S7B), while gene sets for top sensitizing genes included “ribosome” and “proteasome” (Figure S7B). Since the diphtheria toxin catalytic subunit inhibits translation, depletion of the ribosome can be expected to sensitize cells to the toxin. Disruption of the proteasome also sensitizes cells to CTx-DTA, suggesting that the cytosolic toxin is a substrate for proteasomal degradation. Taken together, the unbiased GSEA analysis provides support for the high specificity in hit gene identification by our CRISPRi approach.

We further defined the 50 hits with the strongest protective effect and the 50 hits with the strongest sensitizing effect as “top hits” (all of these are far outside of the range seen with otherwise matched negative control sgRNAs). We characterized these genes by assigning them to cellular pathways and protein complexes according to their previously characterized roles (Figure 6B and Figure S6). Our CRISPRi screen identified a protective effect of knockdown for all top hits recovered in the previously published haploid mutagenesis screen (white stars in Figure 6B). The two top pathways identified by haploid mutagenesis as modulating cellular sensitivity to CTx-DTA are the diphthamide biosynthetic pathway (required to generate eEF-2-diphthamide, the target of diphtheria toxin) and the ganglioside biosynthetic pathway (required to produce GM1a, the cell-surface receptor for cholera toxin). Our screen also identified many additional core components of each pathway. While knockdown of all hits in the diphthamide biosynthesis pathway had a protective effect, the results for ganglioside biosynthesis genes showed a more complex pattern: knockdown of enzymes involved in the production of GM1a were protective, whereas knockdown of enzymes that catalyze the production of other gangliosides (including GM1b) was sensitizing. These results provide genetic confirmation that GM1a is the relevant cell-surface receptor for CTx-DTA and more broadly illustrate the value of being able to reliably detect both sensitizing and protective genes to dissect biological pathways.

Many of the top hits are components of cellular pathways and protein complexes previously identified in experiments to be important for retrograde trafficking and retro-translocation of other toxins such as ricin and Shiga toxin (Bassik et al., 2013; Smith et al., 2009). Retro-translocation of the catalytic chain of CTx has been proposed to be mediated by the ER-associated degradation (ERAD) pathway, although this pathway was not identified in previous genetic screens. Consistent with this proposed role for the ERAD machinery, knockdown of members of the ERAD E3 ubiquitin ligase complex, *SYVN1* (encoding Hrd1) and *SEL1L* (the mammalian homolog of yeast Hrd3) rendered cells resistant to CTx-DTA. Factors that mediate cytosolic degradation of ERAD substrates (in particular *UBXN4*, also

known as UBXD2 or erasin, and the proteasome) appeared as sensitizing hits, suggesting that they may reduce cytosolic levels of the toxin's catalytic subunit in WT cells.

To validate the suggested role of the identified ERAD factors in toxin retro-translocation from the ER to the cytosol, we quantified the amount of CTx chains in the cytosol and membrane fractions. As expected, *SEL1L* knockdown resulted in a dramatic reduction of cytosolic CTx-A1, whereas levels in the membrane fraction were much less affected (Figure 7A–C). By contrast, knockdown of *B4GALNT1*, an enzyme required for the synthesis of the CTx receptor GM1a, resulted in a nearly complete absence of CTx chains from both the cytosolic and the membrane fraction (Figure 7A–C).

An open question in CTx biology is how the toxin traverses the Golgi network (Wernick et al., 2010). Our screen revealed that COG and GARP complexes, which tether late endosomes to the trans-Golgi network or modulate intra-Golgi retrograde transport (Bonifacino and Rojas, 2006) are critical host factors for CTx-DTA. These and other complexes and pathways we identify here (Figure 6B), including several involved in RNA processing, had not previously been linked to cholera toxin biology, highlighting the potential of CRISPRi as a discovery platform. Importantly, many top hits—even those not previously implicated in cholera or diphtheria pathogenesis—were tightly clustered in well-defined protein complexes and pathways. For several of these, the vast majority of components were hits, suggesting that CRISPRi screens can approach saturation.

### Potent Phenotypes and Knockdown Levels Achieved by the Genome-Scale CRISPRi Library

To validate the results from this screen, we re-tested sgRNAs that putatively modulate cellular response to CTx-DTA in mechanistically diverse ways. For each sgRNA, we quantified both the ricin phenotypes as well as the change in abundance of the targeted transcript by qPCR. Our re-test experiments were highly correlated with data from the primary screen (Figure 7D). In our validation experiments for the tiling ricin screen and the genome-scale CTx-DTA screen, the activities of 71 out of 72 sgRNAs were robustly confirmed and were highly correlated ( $R^2=0.879$ ) with the results obtained in the primary screen. Finally, analysis of mRNA levels by qPCR data showed robust repression, with ~80–99% knockdown for each sgRNA and at least 90% for every gene (Figure 7E).

### A Genome-Scale CRISPRa Screen of Cholera-diphtheria Toxin Complements and Extends CRISPRi Results

To further explore the biological insights gained from CRISPRa screening, we performed a genome scale CRISPRa screen for genes that modulate sensitivity to CTx-DTA (Table S2–3). As with the CRISPRi screen, GSEA revealed the specificity of the detected hits (Figure S7B). For some of the top hits, CRISPR-mediated transcriptional activation and repression caused opposite phenotypes (e.g. enzymes in ganglioside biosynthesis, Figure 6C), similar to what we observed for genes controlling ricin sensitivity (Figure 3C).

CRISPRa also revealed additional and highly complementary information as illustrated by analysis of glycosphingolipid biosynthesis pathways. Induction of enzymes in the neolacto

branch of sphingolipid biosynthesis protected cells from CTx-DTA (Figure 6C and Figure S7A,C). This pathway is a parallel branch to the ganglioside branch, which produces the CTx-DTA receptor GM1a. Our findings suggest that upregulation of the neolacto branch diverts the common precursor lactosylceramide away from the ganglioside branch. Similarly, upregulation of the sulfatide-generating enzyme *GAL3ST1* has a protective effect, presumably by diverting ceramide from the sphingolipid to the cerebroside pathway (Figure 6C). These results highlight the capacity of CRISPRa to complement CRISPRi by querying the consequences of upregulating pathways that may otherwise be inactive.

### Effective Knockdown of non-coding RNAs

Finally, we investigated whether CRISPRi was able to repress the transcription of long non-coding RNAs (lncRNAs), a class of transcripts that have been difficult to systematically perturb by other methods (Bassett et al., 2014). Using our CRISPRi library design algorithm, we selected and cloned up to three sgRNAs each targeting six characterized lncRNAs (*GAS5*, *H19*, *MALAT1*, *NEAT1*, *TERC*, *XIST*) (Geisler and Coller, 2013) with good evidence of expression in K562 cells. We transduced the sgRNAs into cells expressing dCas9-KRAB and quantified the amount of transcript knockdown by qPCR. We achieved >80% knockdown for all but one of the lncRNA genes tested (Figure 7F). Overall, more than 50% of the sgRNAs yielded >85% knockdown. We confirmed the strong repression of *XIST* by RNA fluorescence *in situ* hybridization (FISH) and observed no residual expression along the X chromosome (Figure 7G). These results demonstrate that CRISPRi can effectively repress lncRNA expression, enabling future systematic studies of non-coding gene function.

## DISCUSSION

Here, we establish CRISPRi and CRISPRa as robust tools for systematically manipulating transcription of endogenous genes in human cells. We demonstrate that CRISPRi/a can be used to rapidly screen for both loss-of-function and gain-of-function phenotypes in a pooled format. We identify both known and unexpected genes that control growth of K562 cells or that modulate sensitivity to a toxin (CTx-DTA). We also show that we can use CRISPRi/a to create allelic series of gene expression, spanning a broad range from ~100-fold repression to ~10-fold induction, allowing us to define how the abundance of a protein or transcript relates to its function.

A key feature of CRISPRi is the low incidence of off-target effects, as evidenced by the near-absence of activity for three large and distinct classes of negative control sgRNAs in our genome-scale CRISPRi library. This feature simplifies validation and interpretation of screening results. The observed specificity likely stems from two distinct properties of our system. First, CRISPRi/a complexes bound outside a narrow window around the TSS largely fail to modulate transcription; this dramatically shrinks the sequence space across the genome where off-target binding could produce significant off-target activity. Second, CRISPRi activity is highly sensitive to mismatches between the sgRNA and target DNA, suggesting that off-target binding of dCas9 observed in ChIP-seq experiments is too transient to impact transcription (Duan et al., 2014; Kuscu et al., 2014; Wu et al., 2014).

CRISPRa screening provides a new approach for exploring the diversity of transcripts across complex genomes. Gene activation has been used to dissect the limiting component of a biochemical process, identify the molecular target of a drug, or activate key rate-limiting steps in a pathway (Davis et al., 1987; Rine et al., 1983). Recently, a combinatorial cDNA overexpression screen identified genes that, when co-expressed, reprogram fibroblasts into pluripotent stem cells (Takahashi and Yamanaka, 2006). CRISPRa should greatly accelerate similar searches for combinations of factors with emergent properties. In addition, CRISPRa will likely provide insight into cellular pathways where redundancy hampers loss-of-function genetic approaches. CRISPRa will also enable the exploration of cellular states in which otherwise inactive pathways are induced, and thereby reveal functional coupling within complex cellular networks and suggest potential therapeutic strategies.

Our ability to control transcription with high specificity simplifies the analysis and validation of high-throughput screening data. The genome-scale CRISPRi and CRISPRa libraries described here contain 10 sgRNAs per TSS. The resulting library size allows each to be screened in a population of 200 million cells, which can be easily grown in a single spinner flask. Furthermore, the observed high specificity and an improved understanding of rules governing sgRNA activity should enable us to create even more compact sgRNA libraries. Additionally, an sgRNA library designed to activate or repress a broader range of transcripts in the human genome could reveal the function of many non-canonical RNAs encoded in the human genome. As most non-coding transcripts are nuclear and lack an open reading frame, methods that directly modulate transcription are optimally suited for interrogating the function of these RNAs (Derrien et al., 2012).

Systematic genetic interaction (GI) maps are powerful tools for revealing gene functions within pathways or complexes (Bassik et al., 2013; Boone et al., 2007; Costanzo et al., 2010). A CRISPRa GI map or a combined CRISPRi/a GI map could yield rich novel biology and help elucidate how networks of proteins dictate cellular function. More generally, quantitative methods of turning on and off one or multiple transcripts represent a critical tool for understanding how expression of the genes encoded in our genomes controls cell function and fate.

## EXPERIMENTAL PROCEDURES

### CRISPRi/a Libraries

**Tiling libraries**—sgRNAs were designed targeting 49 genes (see Figure 1E) previously identified in shRNA screens as having ricin resistance phenotype. All possible sgRNAs within a 10kb window around the gene TSS and meeting certain criteria were included (see Extended Experimental Procedures). Negative controls were designed based on scrambled sequences from these 10kb windows and filtered by the same criteria as targeting sgRNAs.

**Genome-scale CRISPRi/a libraries**—Genes were selected from the entire set of protein coding genes, although a subset of genes with a RPKM of 0 in a K562 cell RNA-seq expression data set were excluded. sgRNAs conforming to rules including low predicted off-targets and minimal length (see Figure S2 and Extended Experimental Procedures) were selected from a window of  $-50$  to  $+300$ bp (CRISPRi) or  $-400$  to  $-50$ bp (CRISPRa) with

respect to the TSS. Negative controls we designed in the same way based on scrambled sequence derived from the same window of several hundred genes.

**Library Cloning**—Oligonucleotides encoding sgRNAs designed as described above were synthesized as pooled libraries. These were then cloned into lentiviral vectors for expression from a U6 promoter (see Extended Experimental Procedures).

### Cell Line Construction

For constitutive and inducible CRISPRi screens, polyclonal cells expressing dCas9/KRAB fusion proteins driven from an SFFV or TRE3G promoter, respectively, were generated by viral transduction. For CRISPRa screens, a clonal cell line expressing dCas9-GCNx10 and a scFV-sfGFP-VP64 fusion was generated (See Extended Experimental Procedures).

### Growth and Toxin Screens

Cells were grown at minimum library coverage of 1,000 for tiling screens and 3,750 for genome-scale screens. For growth screens cells were grown in spinner flasks and harvested at 0 and 10 days after puromycin selection. For toxin screens, cells were treated with pulses of ricin or CTx-DTA (Bassik et al., 2013; Guimaraes et al., 2011) and harvested when sufficient selective pressure relative to untreated cells had been applied. Briefly, DNA was isolated, the cassette encoding the sgRNA was amplified by PCR, and relative sgRNA abundance was determined by next generation sequencing as previously described (Bassik et al., 2013; Kampmann et al., 2014).

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

The authors thank M. Tanenbaum, J. Tsai, K. Kostova, J. Zalatan, W. Lim, S. Weissman, J. Doudna, and R. Vale for unpublished reagents, technical advice and helpful discussion. L.S.Q. acknowledges support from the UCSF Center for Systems and Synthetic Biology. This work was supported by NIH P50 GM102706 (J.S.W.), NIH P50 GM081879 (L.S.Q., E.H.W.), NIH U01 CA168370 and NIH R01 DA036858 (J.S.W.), as well as the Howard Hughes Medical Institute (J.E.V., Y.C., M.K., M.A.H., J.S.W.). B.A. is an HHMI Fellow of the Damon Runyon Cancer Research Foundation (DRG-[2182-14]). L.A.G. is a Fellow of the Leukemia and Lymphoma Society. M.A.H. is supported by the UCSF Medical Scientist Training Program. M.K. is supported by NCI/NIH Pathways to Independence Award K99CA181494.

### References

- Adamson B, Smogorzewska A, Sigoillot FD, King RW, Elledge SJ. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat Cell Biol.* 2012; 14:318–328. [PubMed: 22344029]
- Alberich-Jordà M, Wouters B, Balastik M, Shapiro-Koss C, Zhang H, Di Ruscio A, DiRuscio A, Radomska HS, Ebralidze AK, Amabile G, et al. C/EBP $\gamma$  deregulation results in differentiation arrest in acute myeloid leukemia. *J Clin Invest.* 2012; 122:4490–4504. [PubMed: 23160200]
- Bassett AR, Akhtar A, Barlow DP, Bird AP, Brockdorff N, Duboule D, Ephrussi A, Ferguson-Smith AC, Gingeras TR, Haerty W, et al. Considerations when investigating lncRNA function in vivo. *eLife.* 2014; 3:e03058. [PubMed: 25124674]

- Bassik MC, Lebbink RJ, Churchman LS, Ingolia NT, Patena W, LeProust EM, Schuldiner M, Weissman JS, McManus MT. Rapid creation and quantitative monitoring of high coverage shRNA libraries. *Nat Methods*. 2009; 6:443–445. [PubMed: 19448642]
- Bassik MC, Kampmann M, Lebbink RJ, Wang S, Hein MY, Poser I, Weibezahn J, Horlbeck MA, Chen S, Mann M, et al. A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. *Cell*. 2013; 152:909–922. [PubMed: 23394947]
- Beerli RR, Segal DJ, Dreier B, Barbas CF. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc Natl Acad Sci U S A*. 1998; 95:14628–14633. [PubMed: 9843940]
- Beerli RR, Dreier B, Barbas CF. Positive and negative regulation of endogenous genes by designed transcription factors. *Proc Natl Acad Sci U S A*. 2000; 97:1495–1500. [PubMed: 10660690]
- Bonifacino JS, Rojas R. Retrograde transport from endosomes to the trans-Golgi network. *Nat Rev Mol Cell Biol*. 2006; 7:568–579. [PubMed: 16936697]
- Boone C, Bussey H, Andrews BJ. Exploring genetic interactions and networks with yeast. *Nat Rev Genet*. 2007; 8:437–449. [PubMed: 17510664]
- Carette JE, Guimaraes CP, Varadarajan M, Park AS, Wuethrich I, Godarova A, Kotecki M, Cochran BH, Spooner E, Ploegh HL, et al. Haploid genetic screens in human cells identify host factors used by pathogens. *Science*. 2009; 326:1231–1235. [PubMed: 19965467]
- Cech TR, Steitz JA. The noncoding RNA revolution—trashing old rules to forge new ones. *Cell*. 2014; 157:77–94. [PubMed: 24679528]
- Chang K, Elledge SJ, Hannon GJ. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods*. 2006; 3:707–714. [PubMed: 16929316]
- Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JLY, Toufighi K, Mostafavi S, et al. The genetic landscape of a cell. *Science*. 2010; 327:425–431. [PubMed: 20093466]
- Davis RL, Weintraub H, Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell*. 1987; 51:987–1000. [PubMed: 3690668]
- Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res*. 2012; 22:1775–1789. [PubMed: 22955988]
- Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, et al. Landscape of transcription in human cells. *Nature*. 2012; 489:101–108. [PubMed: 22955620]
- Duan J, Lu G, Xie Z, Lou M, Luo J, Guo L, Zhang Y. Genome-wide identification of CRISPR/Cas9 off-targets in human genome. *Cell Res*. 2014
- Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*. 2013; 31:397–405. [PubMed: 23664777]
- Geisler S, Collier J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol*. 2013; 14:699–712. [PubMed: 24105322]
- Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*. 2013; 154:442–451. [PubMed: 23849981]
- Guimaraes CP, Carette JE, Varadarajan M, Antos J, Popp MW, Spooner E, Brummelkamp TR, Ploegh HL. Identification of host cell factors required for intoxication through use of modified cholera toxin. *J Cell Biol*. 2011; 195:751–764. [PubMed: 22123862]
- Habedanck R, Stierhof YD, Wilkinson CJ, Nigg EA. The Polo kinase Plk4 functions in centriole duplication. *Nat Cell Biol*. 2005; 7:1140–1146. [PubMed: 16244668]
- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009a; 4:44–57. [PubMed: 19131956]
- Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009b; 37:1–13. [PubMed: 19033363]

- Huang LC, Clarkin KC, Wahl GM. Sensitivity and selectivity of the DNA damage sensor responsible for activating p53-dependent G1 arrest. *Proc Natl Acad Sci U S A*. 1996; 93:4827–4832. [PubMed: 8643488]
- Jackson SP. Sensing and repairing DNA double-strand breaks. *Carcinogenesis*. 2002; 23:687–696. [PubMed: 12016139]
- Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol*. 2003; 21:635–637. [PubMed: 12754523]
- Kampmann M, Bassik MC, Weissman JS. Integrated platform for genome-wide screening and construction of high-density genetic interaction maps in mammalian cells. *Proc Natl Acad Sci U S A*. 2013; 110:E2317–E2326. [PubMed: 23739767]
- Kampmann M, Bassik MC, Weissman JS. Functional genomics platform for pooled screening and generation of mammalian genetic interaction maps. *Nat Protoc*. 2014; 9:1825–1847. [PubMed: 24992097]
- Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera MDC, Yusa K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol*. 2014; 32:267–273. [PubMed: 24535568]
- Kuscu C, Arslan S, Singh R, Thorpe J, Adli M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol*. 2014; 32:677–683. [PubMed: 24837660]
- Mittler G, Stühler T, Santolin L, Uhlmann T, Kremmer E, Lottspeich F, Berti L, Meisterernst M. A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. *EMBO J*. 2003; 22:6494–6504. [PubMed: 14657022]
- Van Ness BG, Howard JB, Bodley JW. ADP-ribosylation of elongation factor 2 by diphtheria toxin. Isolation and properties of the novel ribosyl-amino acid and its hydrolysis products. *J Biol Chem*. 1980; 255:10717–10720. [PubMed: 7000782]
- Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013; 152:1173–1183. [PubMed: 23452860]
- Rine J, Hansen W, Hardeman E, Davis RW. Targeted selection of recombinant clones through gene dosage effects. *Proc Natl Acad Sci U S A*. 1983; 80:6750–6754. [PubMed: 6316322]
- Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol*. 2014; 32:347–355. [PubMed: 24584096]
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*. 2014; 343:84–87. [PubMed: 24336571]
- Sigoillot FD, Lyman S, Huckins JF, Adamson B, Chung E, Quattrochi B, King RW. A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens. *Nat Methods*. 2012; 9:363–366. [PubMed: 22343343]
- Smith RD, Willett R, Kudlyk T, Pokrovskaya I, Paton AW, Paton JC, Lupashin VV. The COG complex, Rab6 and COPI define a novel Golgi retrograde trafficking pathway that is exploited by SubAB toxin. *Traffic Cph Den*. 2009; 10:1502–1517.
- Spitz F, Furlong EEM. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet*. 2012; 13:613–626. [PubMed: 22868264]
- Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006; 126:663–676. [PubMed: 16904174]
- Tanenbaum ME, Macurek L, van der Vaart B, Galli M, Akhmanova A, Medema RH. A complex of Kif18b and MCAK promotes microtubule depolymerization and is negatively regulated by Aurora kinases. *Curr Biol CB*. 2011; 21:1356–1365.
- Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A versatile protein tagging system for signal amplification in single molecule imaging and gene regulation. *Revis*.
- Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science*. 2013; 339:1546–1558. [PubMed: 23539594]

- Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. *Science*. 2014; 343:80–84. [PubMed: 24336569]
- Wernick NLB, Chinnapen DJF, Cho JA, Lencer WI. Cholera toxin: an intracellular journey into the cytosol by way of the endoplasmic reticulum. *Toxins*. 2010; 2:310–325. [PubMed: 22069586]
- Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol*. 2014; 32:670–676. [PubMed: 24752079]
- Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*. 2011; 29:149–153. [PubMed: 21248753]
- Zhang P, Zhang X, Iwama A, Yu C, Smith KA, Mueller BU, Narravula S, Torbett BE, Orkin SH, Tenen DG. PU. 1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood*. 2000; 96:2641–2648. [PubMed: 11023493]
- Zhao M, Sun J, Zhao Z. TSGene: a web resource for tumor suppressor genes. *Nucleic Acids Res*. 2013; 41:D970–D976. [PubMed: 23066107]

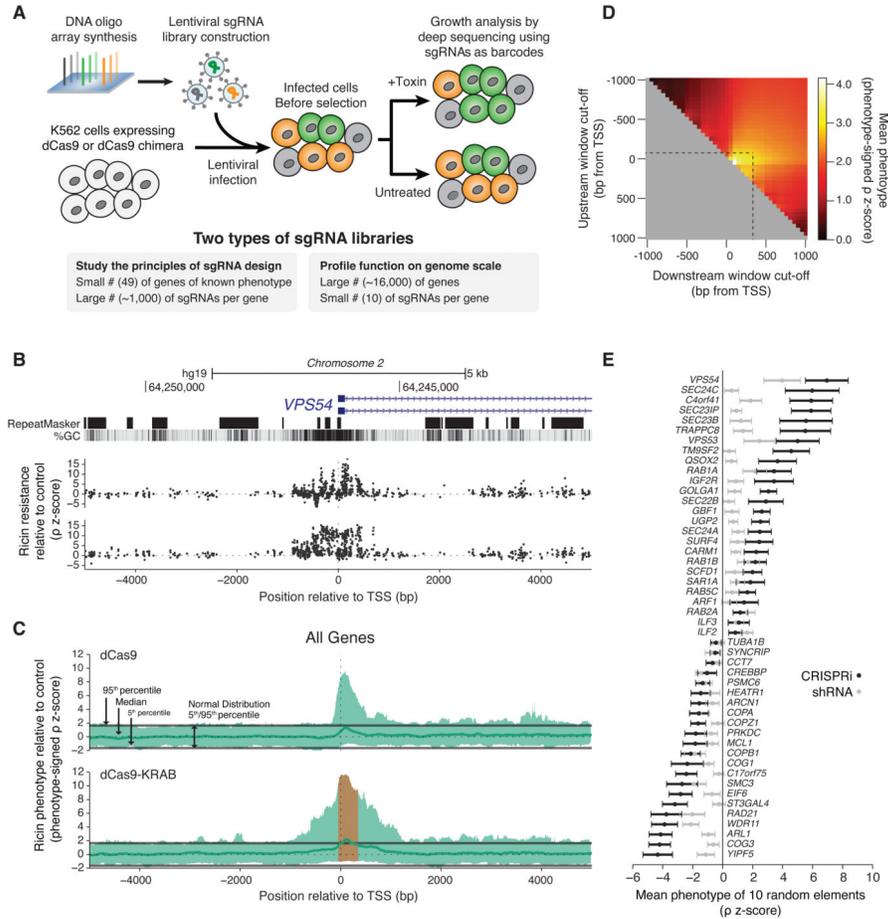
**HIGHLIGHTS**

CRISPRi and CRISPRa provide complementary information for mapping complex pathways

CRISPRi/a expression series (up to ~1000 fold) reveal how gene dose controls function

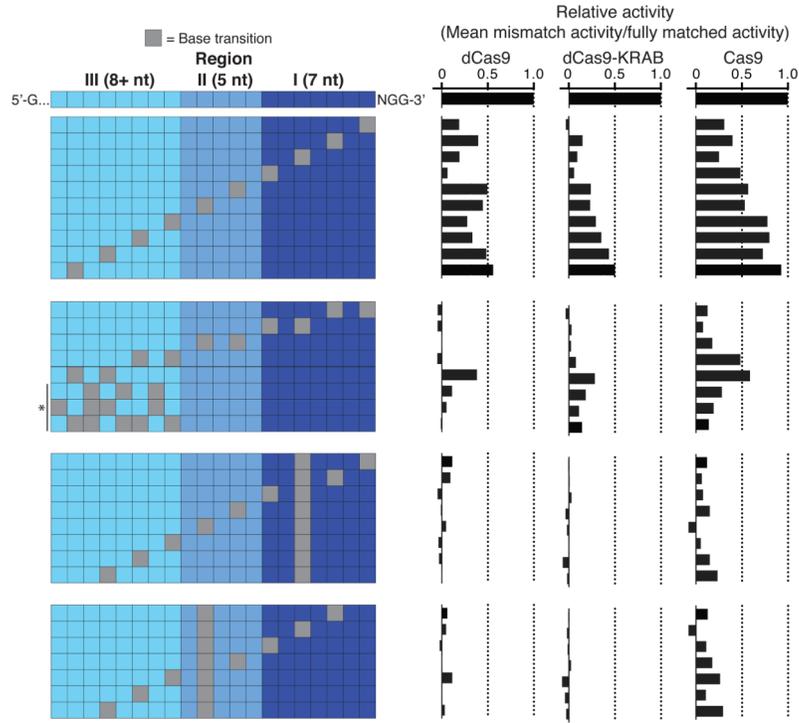
CRISPRi provides strong (typically 90–99%) knockdown with minimal off-target effects

Genome-scale screens elucidate pathways controlling cholera/diphtheria toxicity



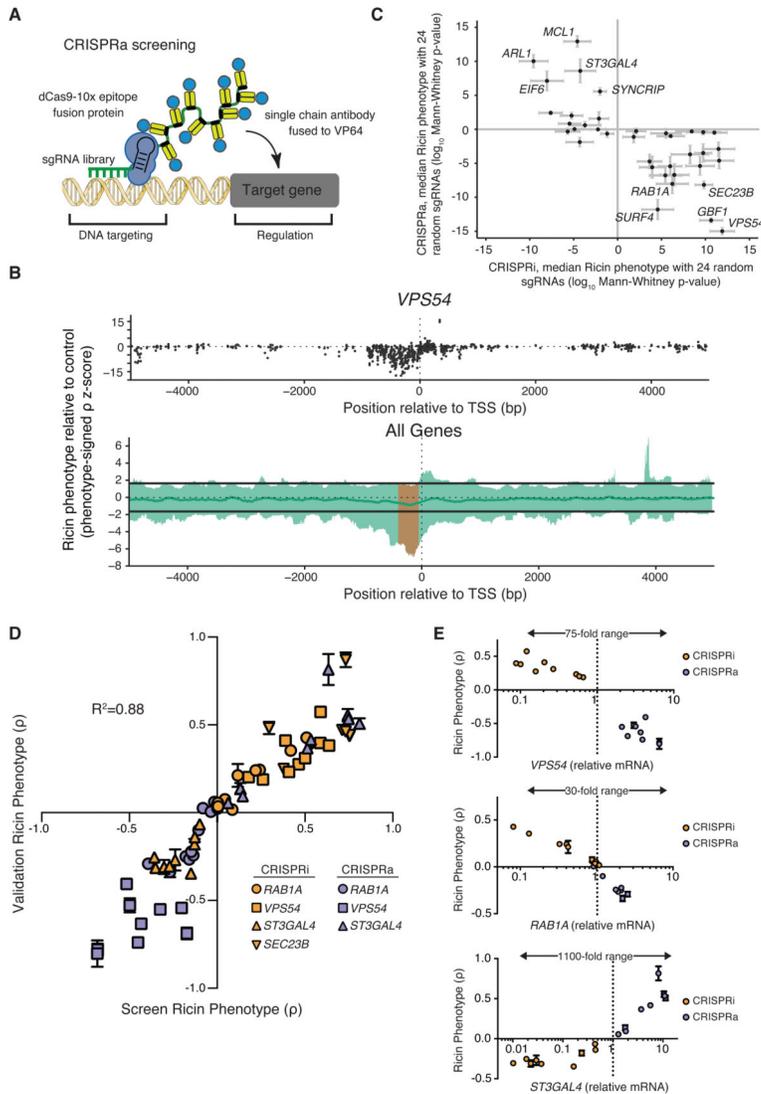
**Figure 1. A Tiling sgRNA Screen Defines Rules for CRISPRi Activity at Endogenous Genes in Human Cells**

(A) Massively parallel determination of growth or toxin resistance phenotypes caused by sgRNAs in mammalian cells expressing dCas9 or dCas9 fusion constructs. (B) UCSC tracks showing the genomic organization, GC content, and repetitive elements around the TSS of a representative gene, *VPS54*, across a 10kb window targeted by the tiling sgRNA library. sgRNA ricin resistance phenotypes (as z-scores, see Figure S1 and Experimental Procedures) in dCas9 and dCas9-KRAB expressing K562 cells are depicted in black on the top and bottom, respectively. See also Figure S2A for more examples. (C) Sliding-window analysis of all 49 genes targeted in a tiling sgRNA library. Green line: median sgRNA activity in a defined window for all genes. Orange region: observed average window of maximum CRISPRi activity. Data displayed as a phenotype signed z-score, excluding all guides longer than 22bp. (D) CRISPRi activity for all 49 genes in defined windows relative to the TSS of each gene. (E) Ricin resistance phenotypes, comparing CRISPRi sgRNAs selected by our rules to RNAi, for genes previously established to cause ricin resistance phenotypes when knocked down by RNAi. Mean phenotype-signed z-score for 10 subsampled sgRNAs or shRNAs. See also Figure S2F.



**Figure 2. CRISPRi Activity is Highly Sensitive to Mismatches Between the sgRNA and DNA sequence**

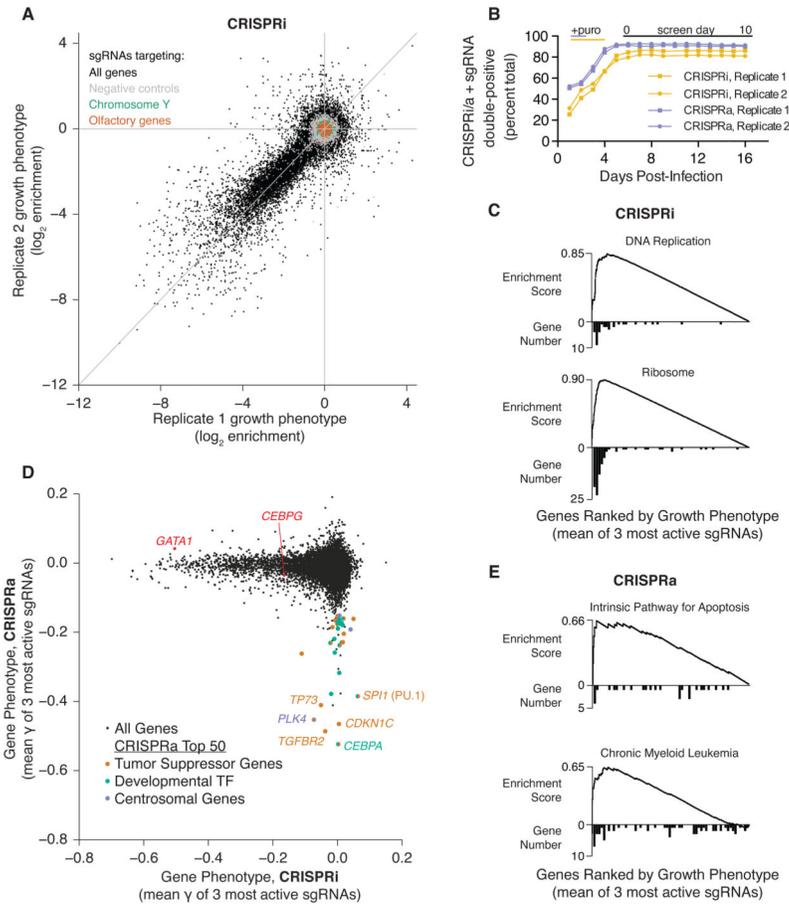
On- and off-target activity of dCas9, dCas9-KRAB and Cas9 for sgRNAs with a varying number and position of mismatches. Off-target activity of sgRNAs with mismatches is displayed as percent of the on-target activity for the corresponding sgRNA without mismatches. Asterisk indicates sgRNAs with 3, 4, or 5 mismatches randomly distributed across region 3 of the sgRNA sequence. Data is displayed for each mismatch position as the mean of all sgRNAs with that mismatch; see Figure S3 for individual sgRNA activities. sgRNAs were included in the analysis only if the fully matched guide was highly active (phenotype-signed z-score  $\geq 4$ ); N=5 for dCas9, N=11 for dCas9-KRAB, and N=10 for Cas9.



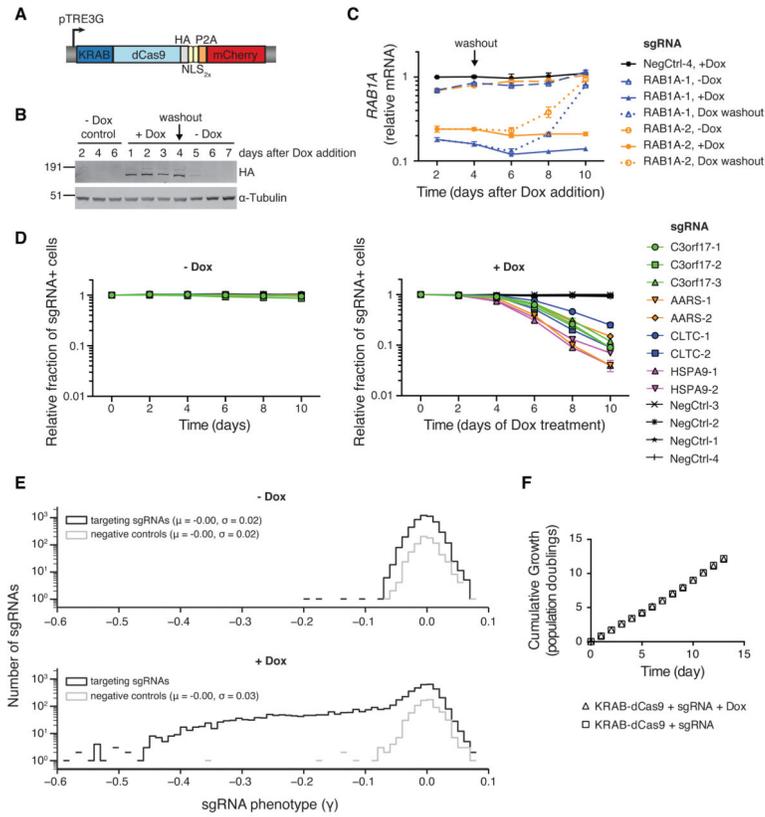
**Figure 3. A Tiling sgRNA Screen Defines Rules for CRISPRa Activity at Endogenous Genes in Human Cells**

(A) A schematic of the dCas9-SunTag + scFV-VP64 + sgRNA system for CRISPRa. (B) Activity of sgRNAs in K562 cells stably expressing each component of CRISPRa, as a function of the distance of the sgRNA site to the TSS of the targeted gene (Phenotype-signed z-scores; therefore, negative values represent opposite results than from knockdown). Top, sgRNAs targeting *VPS54*; Bottom, sliding-window analysis of all 49 genes targeted by our tiling library in green. Green line, median activity; orange, window of maximal activity. Guides longer than 22bp were excluded. See also Figure S4. (C) CRISPRa phenotypes and CRISPRi (dCas9-KRAB) phenotypes are anti-correlated for select genes. For each gene, a p-value is calculated using CRISPRi/a sgRNA activity relative to a negative control distribution for 24 sub-sampled sgRNAs. (D) CRISPRi knockdown and CRISPRa activation of the same gene can have opposing effects on ricin resistance in both primary screens and single sgRNA validation experiments (mean  $\pm$  standard deviation of 3 replicates). (E) Modulation of expression levels for 3 genes by CRISPRi and CRISPRa as quantified by

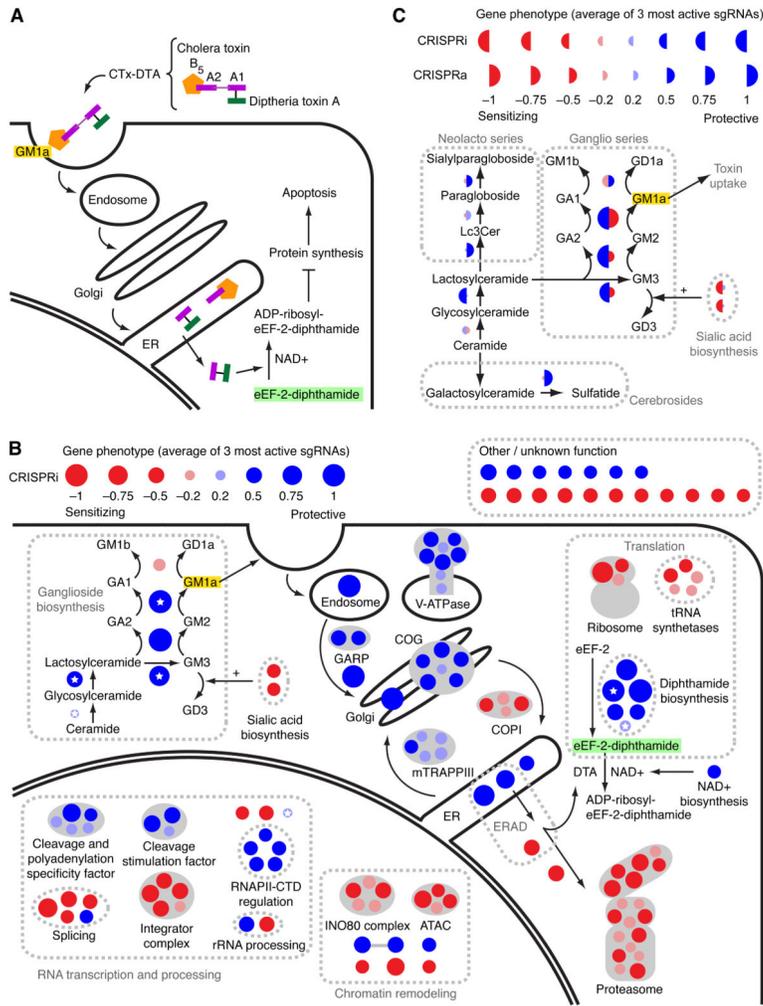
qPCR plotted against the ricin resistance phenotype (mean  $\pm$  standard deviation of 3 replicates) measured for each sgRNA.



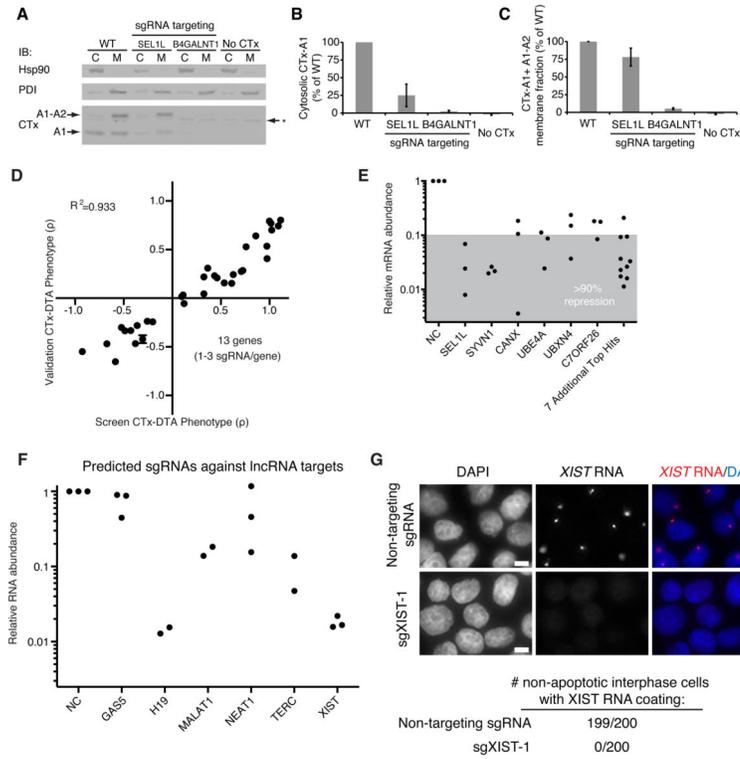
**Figure 4. Genome-Scale CRISPRi and CRISPRa Screens Reveal Genes Controlling Cell Growth** (A) sgRNA phenotypes from a genome-scale CRISPRi screen for growth in human K562 cells (black). Three classes of negative control sgRNAs are color-coded: non-targeting sgRNAs (grey), sgRNAs targeting Y-chromosomal genes (green) and sgRNAs targeting olfactory genes (orange). (B) Co-expression of sgRNAs and dCas9-KRAB or dCas9-SunTag + scFV-VP64 is not toxic in K562 cell lines over 16 days. (C) Gene set enrichment analysis (GSEA) for hits from the CRISPRi screen. A histogram of gene distribution is shown under the GSEA curve. (D) CRISPRi versus CRISPRa gene phenotypes for genome-scale growth screens (black). For the 50 genes in the CRISPRa screen with the most negative growth phenotype, each gene was annotated and labeled based on evidence of activity as a tumor suppressor (orange), developmental transcription factor (green), or in regulation of the centrosome (purple). Two additional CRISPRi hit genes that are discussed in the text are labeled in red. See Table S4 for annotations and references. (E) GSEA for hits from the CRISPRa growth screen. A histogram of gene distribution is shown under the GSEA curve.



**Figure 5. CRISPRi Gene Silencing is Inducible, Reversible, and Non-Toxic**  
**(A)** Expression construct encoding an inducible KRAB-dCas9 fusion protein. **(B)** Western blot analysis of inducible KRAB-dCas9 in the absence, presence, and after washout of doxycycline. **(C)** Relative *RAB1A* expression levels (as quantified by qPCR) in inducible CRISPRi K562 cells transduced with *RAB1A*-targeting sgRNAs in the absence, presence, and after washout of doxycycline. Mean  $\pm$  standard error of technical replicates (N=2) normalized to control cells (assayed in the presence of doxycycline) from the day 2 time point. **(D)** Competitive growth assays performed with inducible CRISPRi K562 cells transduced with the indicated sgRNAs in the presence and absence of doxycycline. Data is represented as the mean  $\pm$  standard deviation of replicates (N=3). See also Figure S5G. **(E)** A CRISPRi sublibrary screen for effects on cell growth was performed with inducible CRISPRi K562 cells in the presence and absence of doxycycline. **(F)** Cumulative growth curves from the sublibrary screen represented in (E) show no bulk changes to growth caused by induction of KRAB-dCas9. Mean  $\pm$  standard deviation of replicate infections each screened in duplicate.



**Figure 6. Genome-Scale CRISPRi and CRISPRa Screens Reveal Known and New Pathways and Complexes Governing the Response to a Cholera-Diphtheria Fusion Toxin (CTX-DTA)**  
**(A)** Model for CTx-DTA binding, retrograde trafficking, retro-translocation and cellular toxicity. **(B)** Overview of top hit genes detected by the CTx-DTA screen. Dark red and blue circles: Top 50 sensitizing and protective hits, respectively. Light red and blue circles: further hits that fall into the same protein complexes or pathways as top 50 hits. Circle area is proportional to phenotype strength. White stars denote genes identified in a previous haploid mutagenesis screen (Guimaraes et al., 2011). See also Figure S6 for hit gene names. **(C)** CRISPRi and CRISPRa hits in sphingolipid metabolism. Display as in (B), except that the left and right sides of each circle represent the phenotypes in the CRISPRi and CRISPRa screens, respectively.



**Figure 7. CRISPRi Strongly Represses Gene Expression of Both Protein-Coding and Non-Coding Genes, Resulting in Reproducible Phenotypes**

(A–C) Cells expressing a negative control sgRNA or an sgRNA targeting *SEL1L* or *B4GALNT1* were incubated with cholera toxin and fractionated to quantify cholera toxin present in the cytosolic and membrane fractions by Western blot. *B4GALNT1* repression blocks toxin uptake whereas *SEL1L* repression prevents toxin retro-translocation from the membrane fraction to the cytosol. (D) Validation of CTx-DTA screen phenotypes with single sgRNA re-test experiments. Data is represented as the mean  $\pm$  standard deviation of replicates (N=3). (E) CRISPRi knockdown of 13 hit genes (28 sgRNAs; same sgRNAs as 7D) identified in the CTx-DTA screen was quantified by qPCR. The gray shaded region denotes sgRNAs showing at least 90% knockdown for each gene. (F) CRISPRi knockdown of 6 lncRNA genes was quantified by qPCR. 2–3 sgRNAs computationally predicted to target each gene were cloned and transduced into K562 cells expressing dCas9-KRAB. (G) K562 cells expressing dCas9-KRAB were transduced with either a non-targeting sgRNA or an sgRNA targeting the *XIST* locus (sgXIST-1). The cells were then stained with DAPI and an RNA FISH probe for the *XIST* transcript. 200 non-apoptotic interphase cells in each condition were scored for *XIST* RNA coating. *XIST* is undetectable in cells transduced with sgXIST-1. Scale bar represents 5 $\mu$ m.



Published in final edited form as:

*Nat Biotechnol.* 2018 February ; 36(2): 170–178. doi:10.1038/nbt.4062.

## Dual gene activation and knockout screen reveals directional dependencies in genetic networks

Michael Boettcher<sup>1</sup>, Ruilin Tian<sup>2</sup>, James A. Blau<sup>1</sup>, Evan Markegard<sup>3</sup>, Ryan T. Wagner<sup>1</sup>, David Wu<sup>1</sup>, Xiulei Mo<sup>4</sup>, Anne Biton<sup>5,6</sup>, Noah Zaitlen<sup>5</sup>, Haiyan Fu<sup>4</sup>, Frank McCormick<sup>3</sup>, Martin Kampmann<sup>2</sup>, and Michael T. McManus<sup>1,\*</sup>

<sup>1</sup>Department of Microbiology and Immunology, University of California San Francisco Diabetes Center, WM Keck Center for Noncoding RNAs, University of California, San Francisco, San Francisco, California, USA

<sup>2</sup>Institute for Neurodegenerative Diseases, Department of Biochemistry and Biophysics, University of California, San Francisco and Chan Zuckerberg Biohub, San Francisco, California, USA

<sup>3</sup>Helen Diller Family Comprehensive Cancer Center, Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, California, USA

<sup>4</sup>Department of Pharmacology and Emory Chemical Biology Discovery Center, Emory University School of Medicine, Atlanta, Georgia 30322, USA

<sup>5</sup>Department of Medicine, Lung Biology Center, University of California, San Francisco, San Francisco, California, USA

<sup>6</sup>Centre de Bioinformatique, Biostatistique et Biologie Intégrative (C3BI, USR 3756 Institut Pasteur et CNRS), Paris, France

### Abstract

Understanding the direction of information flow is essential for characterizing how genetic networks affect phenotypes. However, methods to find genetic interactions largely fail to reveal directional dependencies. We combine two orthogonal Cas9 proteins from *Streptococcus pyogenes* and *Staphylococcus aureus* to carry out a dual screen in which one gene is activated while a second gene is deleted in the same cell. We analyse the quantitative effects of activation and knockout to calculate genetic interaction and directionality scores for each gene pair. Based on the

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

\*Correspondence to: michael.mcmanus@ucsf.edu.

#### Author Contributions

The project was conceived and directed by M.B. and M.T.M. Screen optimisation was performed by M.B. and D.W. Libraries were designed by J.A.B. with guidance from M.B. and cloned by M.B. Orthogonal vectors and cell lines were created by M.B. All screens were performed by M.B., with A.B. assisting in CRISPRa screen analysis. R.T. developed the computational pipelines and the statistical framework for data analysis for screens with guidance from M.K. R.T. also selected the best-performing sgRNAs for arrayed validation. M.B. and R.T.W. conducted and analysed arrayed validation experiments. E.M. and R.T.W. performed western blot analyses. X.M. and H.F. conducted and analysed TR-FRET experiments. M.B. and M.T.M. wrote the manuscript with critical input from R.T., M.K., N.Z. and F.M. All authors read and approved the final manuscript.

#### Competing Financial Interests

The authors declare no competing financial interests.

results from over 100,000 perturbed gene pairs, we reconstruct a directional dependency network for human K562 leukemia cells and demonstrate how our approach allows the determination of directionality in activating genetic interactions. Our interaction network connects previously uncharacterised genes to well-studied pathways and identifies targets relevant for therapeutic intervention.

---

Genetic interaction mapping approaches compare single gene loss-of-function phenotypes against combinations of loss-of-function to identify aggravating or alleviating interactions<sup>1</sup>. However, studies to address the directionality of genetic interactions have been largely limited to lower eukaryotes<sup>2-6</sup>, despite the fact that elucidating human pathway directionality is key to properly interpreting functional genetic dependencies. Such basic information offers rational approaches for therapeutic intervention, precision medicine, and evading drug resistance in human cancers. At the most fundamental level, knowing the directional flow of genetic information is critical to properly reconstruct networks and assemble a cohesive picture of gene function.

The recently discovered bacterial CRISPR phage-defence system has remarkably advanced RNA interference and related gene perturbation technologies<sup>7</sup>. A growing CRISPR toolbox offers a diversity of approaches to perform a highly parallel functional interrogation of every single gene in the human genome<sup>8,9</sup>. However, whereas single perturbation (e.g. knockout or overexpression) approaches have proven highly successful to systematically attribute function to individual mammalian genes, they typically do not provide a deeper understanding of how these genes function together in complex genetic signalling networks.

To reconstruct directional regulatory networks in human cells, we developed an orthogonal CRISPR system comprising two Cas9 enzymes derived from different species. This system allows the simultaneous and asymmetric activation of one gene and deletion of a second gene in the same cell. When compared to conventional symmetrical loss-of-function experiments in which the function of both interaction partners is lost, our orthogonal asymmetric platform allowed us to determine whether the activated gene functionally depends on, or can compensate for the loss of a deleted gene. Using this platform, we identified directional genetic interactions between genes whose activation or ablation altered the fitness of human chronic myeloid leukaemia (CML) cells. We demonstrate that the orthogonal screening approach can quantify loss- and gain-of-function phenotypes from the same cell, and that it is suitable to systematically identify genetic interactions between cancer relevant genes. We reconstruct a substantial number of directional dependencies, connecting previously uncharacterised genes to well-studied pathways.

## Results

### CRISPRa screen identifies cancer pathway genes

CML is a leukaemia characterised by a reciprocal translocation between chromosome 9 and 22. This translocation creates the *BCR-ABL* fusion oncogene, a constitutively active tyrosine kinase oncogene that causes myeloid precursor cells to divide in an uncontrolled fashion<sup>10</sup>. Application of *BCR-ABL* tyrosine kinase inhibitors (e.g. imatinib) have

revolutionised treatment for this cancer, and decades of study have yielded fundamental information on the genes critical for *BCR-ABL* dependent signalling. We thus chose CML to benchmark our method to identify directional genetic interactions, using the K562 CML cell line to systematically quantify genes that function as negative or positive regulators of cancer cell fitness.

To ascertain CRISPR screening conditions capable of identifying the full gamut of imatinib dependent phenotypes, we characterised K562 cell response to a broad range of imatinib drug concentrations. We found that K562 cells respond to a wide range of imatinib concentrations (10 – 1,000 nM), and that CRISPR mediated activation (CRISPRa) of the imatinib efflux transporter *ABCB1* using the SunTag system<sup>11</sup> can result in an approximately 2-fold increase in the IC<sub>50</sub> after 3 days of treatment (Supplemental Fig. 1a). However 3 days of drug treatment did not provide the dynamic range needed to identify genes displaying weaker phenotypes in a screen. To optimise conditions, we analysed the influence of repeated imatinib treatment cycles at IC<sub>50</sub> on cell viability. We observed increased cell viability for *ABCB1* overexpressing cells, 31.5-fold (sgABCB1-1) and 23.5-fold (sgABCB1-2) over negative controls (sgNTC), after three cycles of 100 nM imatinib (day 9) (Supplemental Fig. 1b). These results show that repeated exposure to low imatinib doses allows for much greater enrichment of cells with activated resistance genes than a single treatment.

To systematically identify genes whose activation can alter imatinib drug response, we created an ultra-complex, genome-scale sgRNA library consisting of over 260,000 total sgRNAs targeting every coding and over 4,000 non-coding Refseq annotated (hg19) transcripts in the human genome. Quality-controlled sgRNA libraries (Supplemental Fig. 2) were introduced into K562 CRISPRa target cells<sup>11</sup> followed by 14 days of imatinib treatment with escalating doses of imatinib ranging from 100 nM (IC<sub>50</sub>) to 300 nM (IC<sub>80</sub>) (Fig. 1a). Abundance of sgRNA encoding sequences was determined via next generation sequencing (NGS), comparing the beginning (baseline) and endpoint (day 14) of the screen (Supplemental Table 1). NGS read count ratios of the top 25% most enriched/disenriched sgRNAs were normalized to define an enrichment score ( $\tau$ ) for each gene (Supplemental Table 2). Activation phenotypes were found to be highly reproducible ( $r > 0.98$ ) between technical screen replicates (Supplemental Fig. 3). From a total of 26,700 targeted transcripts, we observed that the activation of 332 genes significantly (FDR < 0.05,  $p < 0.001$ ) altered the fitness of imatinib treated K562 cells, with 57% (188 genes) causing significant depletion (blue) and 43% (144 genes) driving cell enrichment (Fig. 1b).

A key advantage of the gain-of-function approach used here, as opposed to more commonly employed loss-of-function approaches, is that genes exhibiting no- to very low-expression can also be investigated. We found that out of the 332 candidate genes, 21% were not expressed in K562 cells (FPKM < 10<sup>0</sup>) indicating that imatinib responsive genes could be identified from the full spectrum of endogenous gene expression levels (Supplemental Table 2 and Supplemental Fig. 4). This approach allows for the study of genes which may have functional relevance in other cell types, i.e. different types of cancers including non-CML leukemias.

To assess the quality of the screening data on a global level, we executed a gene set enrichment analysis (GSEA)<sup>12,13</sup> using the above 332 target genes identified in the imatinib CRISPRa screen. GSEA identified the strongest gene enrichment in leukaemia and other cancer-related KEGG signalling pathways (Supplemental Fig. 5a), illustrating the ability to identify positive and negative regulators of cancer cell survival pathways. A graphic summary of the CRISPRa screen results, assembled into their relevant oncogenic pathways is shown in Supplemental Fig. 5b. The three strongest hits, namely *ABCB1*, *ABCG2* and *BCR-ABL* are well known to be overexpressed in CML patients with high tolerance to imatinib<sup>14</sup>. Additionally, we identified *BCR-ABL* binding partners *CBL* and *CRKL*<sup>15</sup>, and downstream effectors *SOS1*, *SOS2*, *GAB2*, *RAF1*, *MYC*, *PIM1*, *PIM2* and *STAT5B*<sup>16</sup>, the c-Abl phosphatase *PTPN12*<sup>17</sup>, the Ras-GAPs *NFI*, *RASA1* and *RASA3*<sup>18</sup>, the cell cycle regulators *CDK6*<sup>19</sup> and *CCND3*<sup>20</sup> and receptor tyrosine kinases having well documented roles in imatinib resistance, specifically *PDGFRB*<sup>21</sup>, *FGFR1*<sup>22</sup>, *CSF1R*<sup>23</sup> and *AXL*<sup>24</sup>.

To evaluate the reproducibility of the CRISPRa screen data, we tested three sgRNAs per gene, against 20 of the 332 significant candidate genes individually in an arrayed 96 well plate validation assay. Genes selected for validation included the five most significantly enriched candidate genes *BCR-ABL*, *ABCB1*, *SLC6A14*, *CDK6* and *MYC* as well as 15 genes whose activation produced less significant phenotypes (Supplemental Table 3). The sgRNAs targeting these 20 genes were selected based on CRISPRa screen enrichment (Supplemental Table 1). The results showed a high degree of quantitative reproducibility when compared to screen enrichment data ( $r=0.78$ ), displaying a wide dynamic enrichment range over several orders of magnitude (Supplemental Fig. 6). In addition to the large number of aforementioned genes with well-established roles in leukaemia and imatinib resistance, we identified and validated a set of candidate genes with uncharacterised roles in cancer therapy resistance (Fig. 1c), including numerous solute carriers, the non-coding RNAs *PVT1* and *LOC101928865*, as well as *BBX*, *NOL4L* and *ZC3HAV1* for which upregulation following sgRNA expression was further confirmed via qRT-PCR (Supplemental Fig. 7). In total, these experiments yielded a highly reproducible list of target genes, some having well-established functions in cancer pathways, while others are completely uncharacterised. These results gave us the opportunity to study the functional relationships between the genes; hence, we sought to develop an orthogonal CRISPR platform that could illuminate genetic interactions and directional dependencies for drug resistance.

### The orthogonal CRISPR system

To enable the scalable investigation of directional dependencies, we conceptualized an orthogonal CRISPR system that would allow the simultaneous activation and deletion of two genes in the same cell. To test this concept, we developed a K562 cell line harbouring the *Streptococcus pyogenes* based SunTag CRISPRa system and Cas9 nuclease from *Staphylococcus aureus* (SaCas9). These two Cas9 proteins have different PAM requirements and structural studies have shown that each enzyme recognises different constant regions of the cognate sgRNA<sup>25,26</sup>. These observations suggest that each Cas9 enzyme is not likely to cross-react with the cognate sgRNA engineered for the other Cas9 enzyme.

To test whether both CRISPR systems can work in parallel to activate and delete genes in the same cell without cross-interference, we designed orthogonal sgRNA expression constructs to activate the imatinib efflux transporter *ABCG2* in combination with a non-target control sgRNA, or alternatively, in combination with an sgRNA that deletes *ABCG2* (Fig. 2a). As expected, *ABCG2* protein levels increased following the expression of a CRISPRa sgRNA against *ABCG2* (Fig. 2b, left panel). In contrast, *ABCG2* expression is almost completely abolished when the same CRISPRa sgRNA is used in combination with an sgRNA that deletes *ABCG2* via SaCas9 nuclease (Fig. 2b, right panel). A small residual population of *ABCG2* expressing cells can be observed, which we suspect likely represents cells harbouring non-edited or in-frame indels of the *ABCG2* gene (Fig. 2b, right panel). To compare drug resistance profiles of these sgRNA constructs, we analysed cell enrichment following 11 days of imatinib treatment (Fig. 2c). As shown in Figure 1c, *ABCG2* activation confers imatinib resistance; however, resistance is almost completely reversed in cells that concomitantly express an sgRNA that deletes *ABCG2* (Fig. 2c). Taken together, these data demonstrate the ability of the orthogonal CRISPR platform to simultaneously functionally activate and delete genes in the same cell.

Since our orthogonal system is based on two completely independent CRISPR systems, it opens the door to combinations of any two CRISPR-based technologies, such as transcriptional silencing<sup>27</sup> or targeted DNA methylation<sup>28</sup>, which represents a substantial advance compared to the only other orthogonal CRISPR-based method published to date by Dahlman et al.<sup>29</sup>, which achieves gene activation and knockout using ‘catalytically dead’ sgRNAs engineered to bind the MS2:P65:HSF1 (MPH) activation complex in combination with a catalytically active wt CRISPR Cas9 nuclease from *S.pyogenes*. In contrast, our approach is based on Cas enzymes from two different bacterial species – *S.pyogenes* and *S.aureus* – both of which recognize distinct sgRNAs and PAMs. Consequently, the co-expressed sgRNAs for either CRISPR system do not compete for common protein factors or target sites within the same cell.

### Systematic quantification of genetic interactions

To establish a high-throughput screen using the above described orthogonal CRISPR platform, we created an orthogonal dual sgRNA library composed of selected combinations of CRISPRa and SaCas9 nuclease sgRNAs. This library combined activating sgRNAs targeting 87 enriched or depleted candidate genes from the primary screen (2 sgRNAs/gene for a total of 174 sgRNAs) and knockout sgRNAs targeting 1,327 genes (8 sgRNAs/gene for a total of 11,594 sgRNAs). The knockout sgRNA population targeted all KEGG annotated cancer-relevant signalling pathway genes. The final dual orthogonal sgRNA library contained a total of over 2 million sgRNA or 100,000 gene combinations (annotated library sequences are provided in Supplemental Tables 4 and 5) targeting well-established and dark matter genes.

To promote rigour and reproducibility, the complex sgRNA combination expression library was transduced into two independently derived clonal lines of orthogonal K562 cells. The clonal lines were screened in parallel, in two separate bioreactors in the presence of escalating doses of imatinib. After 19 days, cells from both bioreactors were harvested and

sgRNA representation was compared between baseline cells (day 0) and imatinib-treated cells (day 19, see Methods for details). Before calculating genetic interactions, it was imperative that we first calculated the effects of single gene activation-only and knockout-only phenotypes. To facilitate this calculation, we included a large number of non-target control sequences in the combination library; 18 non-target controls in the CRISPRa position and nearly 900 in the SaCas9 nuclease position (see Supplemental Tables 4 and 5). Dissecting these two populations of control vectors allowed the clean evaluation of replicate performance for both single gene activation ( $r=0.96$ , Fig. 3b) and single gene knockout phenotypes ( $r=0.98$ , Fig. 3c). Both single gene activation and single gene knockout phenotypes included negative and positive regulators of cell fitness in the presence of imatinib. These results were highly reproducible between replicates (Fig. 3b and c), and CRISPRa values correlated ( $r=0.9267$ ) with values from the initial CRISPRa screen (Supplemental Fig. 8).

Notably, phenotypic measurements derived from all possible combinations of sgRNAs in the activation and knockout position with one another were also found to be highly reproducible between clonal screen replicates ( $r=0.94$ ), allowing a quantitative comparison of all possible combinations of sgRNAs in their full operationally functional orthogonal context (Fig. 3d). Taken together these data confirm: 1) the ability for both Cas9 systems to work in parallel to produce activation and knockout phenotypes in the same cell, and 2) the suitability of our NGS analysis pipeline to accurately quantify phenotypes from combinatorial gene perturbations.

### Deducing directional dependencies

In genetic interactions where a gene activates its partner, gene activation and knockout produce opposing phenotypes ( $\tau_{act}$  and  $\tau_{ko}$ ) and the double perturbation phenotype ( $\tau_{act+ko}$ ) can lie in the full spectrum between both individual perturbation phenotypes (Fig. 3e). To systematically identify and quantitate directional genetic dependencies from the screen, we determined genetic interaction (GI) scores from individual and combinatorial  $\tau$  values and based on those, defined a single directionality score  $\Psi$  (Fig. 3f). In essence,  $\Psi$  displays a negative value when enrichment scores  $\tau$  from gene activation and knockout have opposing signs, as would be expected in activating interactions. For interactions with negative  $\Psi$  scores, the  $\tau_{act}$  and GI scores were multiplied to determine whether the activated gene functions downstream (positive value) or upstream (negative value) of the deleted gene (Fig. 3g).

To maintain rigor, we assigned directionality only in reproducible genetic interactions (GI scores that exceeded a 1x standard deviation in both clonal cell line replicates) that exhibited a negative  $\Psi$  score. A summary of all calculated  $\tau$ , GI and  $\Psi$  values is shown in Supplemental Table 6. Based on the determined GI and  $\Psi$  scores, we derived a directional genetic interaction network *de novo*. We assembled the network from the most significant and reproducible directional and non-directional interactions determined by the orthogonal screen (Fig. 3h; Supplemental Table 7 and Supplemental Figure 9). As explained above, directionality among genetic interactions can be inferred only when the activated gene displayed the opposite phenotype of the knocked-out gene; but not if activation and

knockout both resulted in the same phenotype (Supplemental Fig.10). The resulting directional-edge model connects a total of 70 cancer-centric nodes, via 137 gene:gene interactions determined from both clonal replicates, for 26 of which directionality could be directly inferred. For clarity, interactions for all 70 individual nodes are depicted in Supplemental Figure 11.

To quantitatively evaluate the orthogonal screen data, we validated the performance and calculated pathway directionalities in several independent assays. First, a sample set of predicted directional interactions between selected genes with negative  $\Psi$  scores which passed the cut-off in clonal replicate 2 (Supplemental Table 6) were re-tested in an arrayed validation assay using the same orthogonal clonal line. Single as well as combinatorial activation and knockout  $\tau$  values from the arrayed validation experiments were determined and used to derive  $GI_v$  and  $\Psi_v$  scores (Fig. 4a). Control single gene activation and knockout phenotypes validated for all re-tested genes. Activation of *SPRED2*, *WT1* and *TFAP2A* had a sensitizing effect to imatinib treatment, while deletion of *PTPN1*, *NF1*, *MAP4K5* and *RASA2* caused cells to steadily enrich in the culture over time (Fig. 4b and Supplemental Table 8). Figure 4b shows single and double perturbation  $\tau$  values of twelve gene:gene combinations, determined on day 14 of the arrayed validation along with calculated  $GI_v$  and  $\Psi_v$  scores.

Overall,  $GI_v$  scores were found to be in good agreement ( $r=0.72$ ) with  $GI$  scores determined by the orthogonal screen (Supplemental Table 9). In five cases the activated gene was unable to execute its sensitizing function following the deletion of its interaction partner ( $\tau_{act} \times GI =$  negative: *SPRED2-NF1*, *WT1-PTPN1* and *TFAP2A-PTPN1/-NF1/-MAP4K5*), supporting a model for an upstream function of the activated gene. In contrast, we observed three instances where the activated gene could compensate for the loss of its interaction partner ( $\tau_{act} \times GI =$  positive: *SPRED2-PTPN1/-RASA2* and *WT1-MAP4K5*), supporting a model for a downstream function of the activated gene. Out of the total of twelve tested combinations, ten were predicted by the orthogonal screen to show a directional interaction, of which the aforementioned eight were confirmed by our arrayed validation while two interactions (*WT1-NF1* and *TFAP2A-RASA2*) did not reproduce (Fig. 4b and Supplemental Table 9). The inability to validate those two interactions might be explained by the markedly different experimental conditions between the orthogonal screen in a 14 L agitated bioreactor with precisely controlled culture conditions versus validation in a 96-well plate. Moreover,  $GI$  and  $\Psi$  scores from the screen were calculated based on multiple sgRNAs for gene activation and knockout whereas validation was performed with one selected sgRNA in either position.

Based on the validated interactions, we reconstructed a Ras-centric high-confidence directional genetic interaction model with  $\Psi_v$  scores calculated from the validation data (Fig. 4c). This model is further supported by our findings that *SPRED2* cannot sensitise *NF1*-deleted cells to imatinib treatment, despite showing a similar increase in mRNA levels following its activation (Supplemental Fig. 12a and b). At this point it is important to mention that the relative changes in gene expression, as detected by qRT-PCR, do not necessarily translate into equivalent phenotypes. In other words, although the significant increase in *SPRED2* mRNA levels following its CRISPRa mediated activation might seem

modest (approx. 2 fold), the detected phenotype at 8 days after imatinib treatment is a remarkably significant 3-fold decrease in cell numbers (Supplemental Data Figure 12). Moreover, *SPRED2* overexpression in HEK293T cells lowered Ras-GTP levels only in the presence of *NFI*, confirming that the ability of *SPRED2* to suppress Ras activity depends on *NFI* (Supplemental Figure 12c). These observations build upon previous observations that double knockdown of *SPRED1/2* leads to increased Ras-GTP levels due to disruption of an *NFI/SPRED2* complex<sup>30</sup>.

### Exploiting genetic vulnerabilities for cancer therapy

Given the potential to discover genetic dependencies of therapeutic relevance, we investigated the observed interaction between *NFI* and the TAM receptor tyrosine kinase AXL<sup>31</sup> (Fig. 3h and Supplemental Table 7). Targeting AXL-mediated signaling pathways can lead to regained drug sensitivity and improved therapeutic efficacy, defining AXL as a promising target for cancer therapeutics<sup>32,33</sup>. However, a key issue for therapeutic intervention is the selection of appropriate biomarkers and potential synergistic drug targets for combination-based regimes. To evaluate therapeutic applicability and potential synergies, we applied R428 (a specific AXL inhibitor which is currently being evaluated in clinical trials<sup>34</sup>) to a population of Cas9 *NFI*-knockout sgRNA treated cells as well as control cells. We found that *NFI*-knockout cells were highly sensitive to R428, whereas *NFI*-wildtype control cells did not show a significant response to 8 days of treatment with 500 nM R428 (Fig. 5a). Given that the *NFI*-knockout sgRNA treated cells contain sub-populations of non- and in-frame edited cells (Supplemental Fig.13) we anticipate that these observations likely underestimate R428 drug sensitivity. Moreover, these observations were also extended to lung epithelial cells using RNAi mediated knockdown of *NFI* in BEAS-2B cells, which displayed significantly increased drug sensitivity ( $p=3 \times 10^{-3}$  at 500 nM R428,  $p=6 \times 10^{-5}$  at 1000 nM R428) when compared to matched control cells (Supplemental Fig. 14). Finally, we confirmed that *NFI*-knockout K562 cells are more resistant to treatment with imatinib, but that these cells can be re-sensitised to imatinib by R428 treatment (Fig. 5b).

To explore the nature of the selective AXL dependency of *NFI*-deficient cells, we quantitated phosphorylated AXL kinase (p-AXL) levels in control untreated wildtype cells and *NFI*-knockout sgRNA treated cells. *NFI*-knockout sgRNA cells displayed markedly higher levels of p-AXL than *NFI*-wildtype cells, indicating that these cells had accumulated higher levels of AXL activity and p-AXL levels were reduced upon R428 treatment in both, *NFI*-wildtype and knockout cells (Fig. 5c). To further investigate the interaction between *NFI* and AXL, we performed a homogeneous time-resolved Förster resonance energy transfer (TR-FRET) assay where the stringent proximity (<10 nm) based idiosyncrasy allows the detection of direct physical interactions<sup>35</sup>. These experiments provided additional support that *NFI* and AXL physically interact with each other in a cell based assay (Fig. 5d). Additionally, we show that both proteins bind to all three Ras isoforms N-Ras, H-Ras and K-Ras (Supplemental Figure 15), supporting a model where *NFI* deficient cells become increasingly dependent on AXL signalling, and that these cells can be selectively targeted by the AXL inhibitor R428. Given the high recurrence of *NFI* mutations and AXL activation in a variety of human cancers, our data provide an informed basis for therapeutic intervention.

## Discussion

Inferring the direction of genetic interactions has been a long-standing challenge. While previously described genetic interaction studies are based on simple dual loss-of-function<sup>36–39,40–43</sup>, the orthogonal approach combines the power of CRISPR mediated activation of one interaction partner with the functional loss of a second gene in the same cell. Here we establish the full methodology and reagents necessary to conduct highly parallel directional CRISPR screens in human cancer cells, including stable CRISPRa-SaCas9 nuclease cell lines, dual sgRNA libraries, and a barcode-free next generation sequencing strategy to quantify sgRNA combinations in orthogonal screens. As a general concept, our described inference of directionality strategy is readily applicable to numerous other dual activation/inhibition expression embodiments, notwithstanding other transcriptional, post-transcriptional, and post-translational regulatory modules.

Approaches to construct quantifiable directional models for genetic interactions have been limited and there have been no established technologies to efficiently specify directionality within pathways. This is particularly a problem in fields such as cancer biology where a major ongoing focus is to identify synergistic genetic vulnerabilities that provide a sound basis for the design of rational polytherapies to help prevent drug resistance. Here, we provide a comprehensive dataset consisting of single and combinatorial gain- and loss-of-function phenotypes in CML cells, and a high-confidence network of genetic interactions that will help researchers to build hypotheses to further understand why some patients respond well to tyrosine kinase inhibitors like imatinib, whereas others acquire resistance. In many cases, directional dependencies need to be considered when designing a treatment plan for patients harbouring multiple genetic lesions, and the described orthogonal platform offers a fresh new approach to uncovering key dependencies in pathways critical for human gene function and disease.

## Methods

### Vector maps

For the single sgRNA (sgLenti), dual sgRNA (sgLenti-orthogonal) and SaCas9 nuclease vector, vector maps are provided in Genbank format (Supplemental 1–3) and have been deposited along with the plasmids at Addgene

### CRISPRa and orthogonal K562 cell lines

K562 CRISPRa cells<sup>11,46</sup> were kindly provided by Luke Gilbert and cultured in RPMI 1640 medium, supplemented with 10% fetal bovine serum and 1x Anti-Anti (Gibco). Via lentiviral transduction, *S. aureus* Cas9 under the control of an EF1 $\alpha$  promoter, was introduced into K562 CRISPRa cells (see Supplemental 3 for vector map). Successfully transduced cells were selected with hygromycin (200  $\mu$ g/mL) and single clones were expanded for 14 days. To test functionality of the expanded clonal orthogonal lines, cells were transduced with sgRNAs to activate the imatinib efflux transporter *ABCG2* via CRISPRa (5'-GCCACTGCGTTCAGCTCTGG-3') or to knock it out in combination with SaCas9 (5'-CATCTGCTATCGAGTAAACTG-3'). Four weeks post introduction of the

SaCas9 expression cassette, clonal lines were screened for functionality of both CRISPR systems (CRISPRa and SaCas9 nuclease) via flow cytometry analysis of >10,000 cells stained with CD338 (*ABCG2*) antibodies (Miltenyi, 130-104-960). Out of a total of 28 screened orthogonal lines, all 28 retained the functional CRISPRa system but only four lines displayed stable function of SaCas9 nuclease. Two of those lines were used for the orthogonal CRISPR screen.

### CRISPRa and orthogonal sgRNA library design

For the initial CRISPRa screen, a genome-scale sgRNA library consisting of over 260,000 total sgRNAs targeting every coding, and over 4,000 non-coding, Refseq annotated (hg19) transcripts in the human genome, as well as every unique protein coding isoform with up to 12 sgRNAs, plus 7,700 non-target control sequences (NTC).

The promoter regions for coding transcripts targeted windows 25 to 500bp upstream of the Refseq-annotated transcription start sites. SgRNAs were designed against targets in the promoters that are of the format (N)<sub>20</sub>NGG, and selected sgRNAs must pass the following off-targeting criteria: 1) the 11bp-seed must not have an exact match in any other promoter region, and 2) if there is an exact off-target seed match, then the rest of the sgRNA must have at least 7 mismatches with the potential off-target site. Regions outside a window of 25 to 500 bp upstream of the TSS were not considered for off-targeting since the employed CRISPRa system was shown to work only in proximity to the TSS of genes<sup>46</sup> and to not further limit the number of designable sgRNAs for the narrow on-target space. After all sgRNAs that pass off-targeting criteria were generated, up to 12 sgRNAs/transcript were selected that were nearest to the transcription start sites. All sgRNA sequences are shown in Supplemental Table 1. In addition to the sgRNA sequence, every plasmid contained a unique 20 nt barcode sequence (see Supplemental 1 for vector map). This sequence allowed the distinction between sgRNAs expressed from different plasmids and hence in different sub-populations of cells and was used to bin cells into mutually exclusive barcode bins to create technical screen replicates after sequencing.

For the orthogonal genetic interaction screen, a focused nuclease-active *S. aureus* Cas9 library was generated targeting 1327 genes. For the selected genes, sgRNAs targeting coding exons were generated using Cas-Designer<sup>47</sup>, generating sgRNAs that were adjacent to the PAM sequence 'NHGRRT' (H = A, C, or T), which allows for targeting with *S. aureus* Cas9 but not with *S. pyogenes* Cas9. Potential off-targets against the human genome were identified using Cas-OFFinder<sup>48</sup>. To score sgRNA sequences by Cas-OFFinder, sgRNAs that have perfect-seed off-targets and 5 mismatches or less in potential off-target regions were penalised. The 20% of sgRNAs with the highest off-target penalties and bottom 20% of sgRNAs with the lowest out-of-frame scores from Cas-Designer were eliminated. From the resulting list of sgRNAs, up to 8 sgRNAs/gene were selected, targeting the most 5' constitutive exons for each gene.

### CRISPRa and orthogonal sgRNA library cloning

For the CRISPRa library, the designed 20 nt target specific sgRNA sequences were synthesised as a pool, on microarray surfaces (CustomArray, Inc.), flanked by overhangs

compatible with Gibson Assembly<sup>49</sup> into the pSico based sgLenti sgRNA library vector (see Supplemental 1 for vector map). The synthesised sgRNA template sequences were of the format: 5'-GGAGAACCACCTTGTTGG-(N)<sub>20</sub>-GTTTAAGAGCTATGCTGGAAAC-3'. Template pools were PCR amplified using Phusion Flash High-Fidelity PCR Master Mix (ThermoFisher Scientific) according to the manufacturers protocol with 1 ng/uL sgRNA template DNA, 1 uM forward primer (5'-GGAGAACCACCTTGTTGG-3'), 1 uM reverse primer (5'-GTTTCCAGCATAGCTCTTAAAC-3') and the following cycle numbers: 1x (98C for 3 min), 15x (98C for 1 sec, 55C for 15 sec, 72C for 20 sec) and 1x (72C for 5 min). PCR products were purified using Minelute columns (Qiagen). The library vector sgLenti was prepared by restriction digest with AarI (Thermo-Fischer) at 37C overnight, followed by 1% agarose gel excision of the digested band and purification via NucleoSpin columns (Macherey-Nagel). Using Gibson Assmblly Master Mix (NEB), 1000 ng digested sgLenti and 100 ng amplified sgRNA library insert were assembled in a total 200 uL reaction volume. The reaction was purified using P-30 buffer exchange columns (Biorad) that were equilibrated 5x with H<sub>2</sub>O and the total eluted volume was transformed into three vials of Electromax DH5α (ThermoFisher). E.coli were recovered, cultured overnight in 500 mL LB (100 ug/mL ampicillin) and used for Maxiprep (Qiagen). In parallel, a fraction of the transformation reaction was plated and used to determine the total number of transformed clones. The coverage was determined to be 70x clones per sgRNA ensuring even representation of all library sgRNA sequences and their narrow distribution (Supplemental Fig. 2). Fidelity of sgRNA sequences was confirmed with a more than 90% perfect Bowtie alignment rate and narrow distribution of sgRNA sequences, with read counts for 87% of sgRNA sequences falling within a single order of magnitude.

For orthogonal CRISPR libraries, CRISPRa sgRNA pools of 174 sgRNA against 87 selected target genes (2 sgRNAs/gene) plus 18 non-target control sgRNAs were cloned into position 1 of the AarI-digested plasmid sgLenti-orthogonal exactly as described for the CRISPRa library. Off target analysis using Cas-OFF finder<sup>48</sup> showed that out of the total 192 sgRNAs, only two had an additional perfectly matched genomic target site which was outside of the defined relevant CRISPRa off-target space (25 to 500 nt upstream of the TSS) while the rest had exactly one target site. In addition, 6 sgRNAs had off-target sites with 1 mis-match and 65 sgRNAs had off-target sites with 2 mis-matches outside the defined off-target space. A full summary of CRISPRa sgRNA sequences with the number and nature of determined off-target sites is shown in Supplemental Table 4.

Following amplification in E.coli, library plasmids with the first position cloned were digested with BfuAI (NEB) to allow cloning of SaCas9 sgRNAs into the second position. To remove undigested orthogonal sgRNA library plasmid from the pool, the purified (Nucleospin, Macherey-Nagel) BfuAI digested plasmid was subsequently digested with AscI for which restriction sites exist in the stuffer sequences in both sgRNA positions 1 and 2. BfuAI/AscI digested plasmid was extracted from 1% Agarose gel (Nucleospin, Macherey-Nagel).

Synthesised SaCas9 sgRNA template sequences (12,500 total, 8 sgRNAs/gene) were of the format: 5'-GAAAGGACGAAACACCGTG-(N)<sub>22</sub>-GTTTTAGTACTCTGGAAACAGAATCT-3'. PCR amplification of the SaCas9 template

pool was performed as described above using primer sequences: 5'-GAAAGGACGAAACACCGTG-3' and 5'-AGATTCTGTTTCCAGAGTACTAAAAC-3' and the purified PCR product was cloned into BfuAI digested sgLenti-orthogonal (see Supplemental 2 for vector map) via Gibson Assembly as described above. The resulting orthogonal sgRNA library was transformed into Electromax cells at 30x coverage as described above and the plasmid sgRNA library pool was purified (Qiagen Plasmid Maxi kit). From the resulting plasmid pool, sgRNA sequences were recovered via PCR as described below and sequenced for quality control. At a read depth of 94x, 2.389 million out of the total possible 2.394 million combinations (>99%) were read at least once, with less than 5% of the library elements read 20 or less times.

### Lentivirus production

HEK293T cells were seeded at 65,000 cells per ccm in 15 cm dishes in 20 mL media (DMEM, 10% fetal bovine serum) and incubated overnight at 37C, 5% CO<sub>2</sub>. The next morning, 8 ug sgRNA library plasmid, 4 ug psPAX2 (Addgene #12260), 4 ug pMD2.G (Addgene #12259) and 40 uL jetPRIME (Polyplus) were mixed into 1 mL serum free OptiMEM (Gibco) with 1x jetPRIME buffer, vortexed and incubated for 10 min at RT and added to the cells. 24 h later, 40U DNaseI (NEB) were added to each plate in order to remove untransfected plasmid and at 72h post-transfection, supernatant was harvested, passed through 0.45 um filters (Millipore, Stericup) and aliquots were stored at -80C.

### Genome-wide and orthogonal CRISPR screens

Imatinib selection conditions for all screens were optimized by activating the imatinib efflux transporter *ABCB1* using sgABCB1-1 (5'-CAGGAACAGCGCCGGGGCGT-3') and sgABCB1-2 (5'-AGCATTTCAGTCAATCCGGGC-3') (Supplemental Figure 1). K562 CRISPRa/orthogonal cells were transduced with lentivirally packaged sgRNA libraries at MOI=0.3 and 500x coverage. The low MOI was used to reduce the frequency of multiple-infected cells; thus, only one gene was activated in each cell. Cells were then cultured in RPMI with 10% FBS and 1x Anti-Anti (Gibco) in a 37C incubator with 5% CO<sub>2</sub>. 48h post transduction, cells were selected with puromycin (2 ug/mL) for 96h. Following selection, aliquots of 300 million cells each, were frozen down in FBS with 10% DMSO for later analysis via NGS (see below). Fully selected cells (300 million) were transferred into a 14 liter CelligenBlu bioreactor (Eppendorf) and sub-cultured at 37C, pH=7.4 and 2% oxygen. Coverage at cell level was kept above 1000x throughout the entire screen and the culture was diluted with fresh medium when cell density reached 1 mio/mL.

For the genome-wide CRISPRa screen: 14 days post transduction, aliquots of 300 mio cells from the beginning of the screen were frozen down (baseline sample) as described above and an IC<sub>50</sub> concentration of 100 nM imatinib (Sigma) was added to the bioreactor vessel. Imatinib was refreshed on day 17 (IC<sub>60</sub> = 150 nM) and day 19 (IC<sub>80</sub> = 300 nM) after initial sgRNA library transduction and cells for the analysis of the final time point were harvested on day 28. For the orthogonal genetic interaction screen: Puromycin selected cells (2 ug/mL) at 8 days post transduction (2.5 billion per sample) were frozen down as described above and 100 nM imatinib (IC<sub>50</sub>) were added to the bioreactor vessel. Imatinib concentrations were increased throughout the screen to the IC<sub>60</sub> concentration of 150 nM (day 10), the IC<sub>80</sub>

of 300 nM (day 13 and 15) and finally the IC<sub>90</sub> of 500 nM (day 17). On day 19 2.5 billion cells per sample were harvested for downstream analysis via NGS as described below.

### Genomic DNA (gDNA) extraction

Cell pellets from baseline and imatinib treated samples were resuspended in 20 mL P1 buffer (Qiagen) with 100 ug/mL RNase A and 0.5% SDS followed by incubation at 37C for 30 min. After that, Proteinase K was added (100 ug/mL final) followed by incubation at 55C for 30 min. After digest, samples were homogenised by passing them three times through a 18G needle followed by three times through a 22G needle. Homogenised samples were mixed with 20 mL Phenol:Chlorophorm:Isoamyl Alcohol (Invitrogen #15593-031), transferred into 50 mL MaXtract tubes (Qiagen) and thoroughly mixed. Samples were then centrifuged at 1,500g for 5 min at room temperature (RT). The aqueous phase was transferred into ultracentrifuge tubes and thoroughly mixed with 2 mL 3M sodium acetate plus 16 mL isopropanol at RT before centrifugation at 15,000g for 15 min. The gDNA pellets were carefully washed with 10 mL 70% ethanol and dried at 37C. Dry pellets were resuspended in H<sub>2</sub>O and gDNA concentration was adjusted to 1 ug/uL. The degree of gDNA shearing was assessed on a 1% agarose gel and gDNA was sheared further by boiling at 95C until average size was between 10–20 kb.

### PCR recovery of sgRNA sequences from gDNA

Multiple PCR reactions were prepared to allow amplification of the total harvested gDNA from a 1000x cell coverage for each sample. For the first round of two nested PCRs, the total volume was 100 uL containing 50 ug sheared gDNA, 0.3 uM forward (5′ - ggcttgattctataactctgatagca-3′) and reverse (5′ - cggggactgtggcgatgtg-3′) primer, 200 uM each dNTP, 1x Titanium Taq buffer and 1 uL Titanium Taq (Clontech). PCR cycles were: 1x (94C - 3 min), 16x (94C - 30 sec, 65C - 10 sec, 72C - 20 sec), 1x (68C - 2 min). All first round PCRs were pooled and a fraction was used as template for the second round PCR. The total volume of the second round PCR was 100 uL containing 2 uL pooled first round PCR, 0.5 uM forward (5′ - AATGATACGGCGACCACCGAGATCCACAAAAGGAAACTCACCTAAC-3′) and reverse (5′ - CAAGCAGAAGACGGCATAACGAGAT-(N)<sub>6</sub>-GTGACTGGAGTTCAGACGTG-3′) primer where (N)<sub>6</sub> is a 6 nt index for sequencing on the Illumina HiSeq platform, 200 uM each dNTP, 1x Titanium Taq buffer and 1 uL Titanium Taq (Clontech). PCR cycles were: 1x (94C - 3 min), 16x (94C - 30 sec, 55C - 10 sec, 72C - 20 sec), 1x (68C - 2 min). The resulting PCR product (344 bp) was extracted from a 1% agarose gel. For the orthogonal genetic interaction screen, conditions for the first round PCR were slightly modified to: total reaction volume 80 uL containing 20 ug sheared gDNA and the second round PCR product was 887 bp.

Gel extracted bands from the primary CRISPRa screen were submitted for sequencing on an Illumina HiSeq 2500 platform using paired end 50 kits with the custom sequencing primer 5′ - GAGACTATAAGTATCCCTTGGAGAACCACCTTGTGG-3′ for reading the sgRNA sequence and the Truseq Illumina reverse primer to read out 20 nt random barcode sequences used for generation of technical screen replicates (separation of sgRNA reads into three groups with mutually exclusive barcode sequence bins). For orthogonal dual sgRNA

library analysis, single end 50 kits were used and read cycles were split, 25 cycles for Read1 with the sequencing primer above (reading the *S.pyogenes* sgRNA) and 25 read cycles for the ‘Illumina indexing read’ with the custom indexing primer 5’-TTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGTG-3’ (reading the *S.aureus* sgRNA).

### Data analysis

Total read counts of sgRNA sequences from each NGS sample were collapsed and quantified via alignment to the sgRNA library reference sequences using Bowtie 2.0<sup>14</sup>. Data analysis was conducted similarly as described previously<sup>38,44</sup>. Briefly, for the primary CRISPRa screen, the frequency of sgRNAs was determined by deep sequencing and the average read count of three technical replicates was used. The phenotype  $\tau$  was calculated to quantify the effect of an sgRNA on cell growth in the presence of imatinib. Specifically,  $\tau$  values were calculated as:

$$\tau_x = \log_2 \left( \frac{\left( \frac{N_t^x}{N_{t0}^x} \right)}{\left( \frac{N_t^{NTC}}{N_{t0}^{NTC}} \right)} \right)$$

where  $N^x$  denotes the frequency of sgRNA  $x$  and  $N^{NTC}$  denotes the frequency of non-targeting control sgRNAs at baseline ( $t_0$ ) or after imatinib treatment ( $t$ ). Gene-level phenotypes were calculated by averaging the phenotypes of the top 25% most extreme sgRNAs targeting this gene. The statistical significance for each gene is determined by comparing the set of  $\tau$  values for sgRNAs targeting it with the set of  $\tau$  values for non-targeting control sgRNAs using the Mann-Whitney U test, as described previously<sup>44</sup>. To correct for multiple hypothesis testing, we first performed random sampling with replacement among the set of  $\tau$  values for non-targeting control sgRNAs and calculated p values for each sampling. Then, we calculated the false discover rate (FDR) based on the distribution of P values for all genes in the library and for non-targeting controls generated above. The P-value cutoff was chosen based on an FDR < 0.05.

For the orthogonal double-sgRNA screen, combinations of non-targeting control sgRNAs served as negative control, combinations of one non-targeting control sgRNA and one targeted sgRNA were used to determine single-sgRNA phenotypes and combinations of two targeted sgRNAs were used to calculate double phenotypes. Raw read counts used for analysis are shown in Supplemental Table 10. We then implemented a series of filtering steps on the sgRNA level. First of all, on the *SaCas9* nuclease side, p values were calculated for each gene as described above. Only the sgRNAs targeting genes that have significant editing phenotypes (P value < cutoff) were retained. Subsequently, GI scores were calculated using the ‘force-fit’ definition for genetic interactions on the sgRNA level and sgRNAs were further filtered by GI correlation as described previously<sup>44</sup>. On the CRISPRa side, if two sgRNAs targeting the same gene have low correlation, the gene was excluded for further

analysis. After the filtering process, gene-level phenotypes and GI scores were calculated by averaging all double-sgRNAs targeting the same gene-gene combinations.

### Directional genetic interaction network model

Genetic interactions whose GI scores exceeded a 1x standard deviation consistently in both clonal screen replicates were used to construct a GI network (Supplemental Table 7). To quantify directionality in these reproducible genetic interactions, a directionality score ( $\Psi$ ) was calculated as

$$\psi = \tau_{activation} \times \tau_{knockout} \times GI^2$$

resulting in a negative  $\Psi$  when gene activation and knockout had opposing phenotypes. Negative  $\Psi$  values below a negative 1x standard deviation of all calculated  $\Psi$  values were used to infer the direction of genetic interactions. The network analysis software platform Cytoscape<sup>50</sup> was used to visualise the genetic interaction model. Where applicable, directionality in GIs was indicated by arrow shaped edges and line shaped edges indicate significant GIs for which directionality could not be inferred. Nodes were coloured according to gene function with blue symbolizing genes that act to decrease and red to increase cell fitness of imatinib treated cells.

### Arrayed competitive growth validation experiments

Individual CRISPRa or orthogonal dual sgRNA sequences for validation experiments were sub-cloned into the same vector as the respective libraries. For that purpose oligonucleotides encoding the sgRNA sequence as well as the reverse complementary sgRNA sequence were synthesised with compatible 4 nucleotide 5'-overhangs for cloning into the SpCas9 (top strand: TTGG, bottom strand: AAAC) or SaCas9 (top strand: GCTG, bottom strand: AAAC) position of the target vector respectively (for vector preparation see above). Oligonucleotides were adjusted to 100  $\mu$ M and reverse complementary strands were mixed, heated to 99°C and cooled down to 4°C at a ramp rate of  $-0.1^\circ\text{C}/\text{sec}$  in a thermocycler. Annealed oligonucleotide double strands were diluted 1:200 and 1  $\mu$ L was mixed with 50 ng digested vector, 1  $\mu$ L 10x T4 ligase buffer (NEB) and 0.5  $\mu$ L T4 ligase (2000U/ $\mu$ L, NEB) in a total volume of 10  $\mu$ L. Following incubation for 30 min at room temperature, 1  $\mu$ L ligation reaction was transformed into 20  $\mu$ L chemically competent DH5 $\alpha$  E.coli, plated on LB-amp (100  $\mu$ g/mL) agar plates and incubated at 37°C overnight. Individual clones were picked and sgRNA sequences were validated via Sanger sequencing.

All library vectors co-expressed mCherry which was used to track the abundance of sgRNA expressing cell populations in growth competition assays. For this purpose, sgRNA expressing cells were mixed with parental - mCherry-negative - cells at ratios between 1:1 and 1:3 in 96-well plates before repeated treatment with imatinib, R428 or no drug for indicated time periods. Enrichment or depletion of the mCherry positive (sgRNA expressing) cell population, indicating an increase or decrease of fitness over time following sgRNA expression and could conveniently be followed via FACS quantification of the mCherry-positive (sgRNA expressing) versus mCherry negative (parental, no sgRNA

expressing) population. For screen validation experiments,  $\tau$  values for each sgRNA were calculated equivalent to screen  $\tau$  values:

$$\tau_x = \log_2 \left( \frac{\left( \frac{N_t^x}{N_{t0}^x} \right)}{\left( \frac{N_{NTC}^x}{N_{t0}^{NTC}} \right)} \right)$$

where N represents the fraction of sgRNA expressing (mCherry-positive/mCherry-negative) baseline cells ( $t_0$ ) or indicated time points (t) and X represents a given sgRNA while NTC represents a scrambled non-target control sgRNA. Each value was quantified from three technical replicates. Where indicated, fold-enrichment or fold-change values were calculated as 2 to the power of  $\tau$  ( $2^\tau$ ). The sgRNA sequences used for validation of candidate genes from the primary CRISPRa screen are shown in Supplemental Table 3.

For validation of genetic interactions, dual sgRNA expression constructs were cloned into sgLenti-orthogonal using CRISPRa sgRNAs for *SPRED2* (5'-GATTCGGAGCCAGACGGTCG-3'), *WT1* (5'-GGACTCACTGCTTACCTGAA-3'), *TFAP2A* (5'-AGGGGAATGTGGCGGAATTG-3') and non-target control (5'-CCCTGCCGTCCTCTACGAAT-3') and *SaCas9* nuclease sgRNAs for *NFI* (5'-TTGTCTTTGGGTGTATTAGCAA-3'), *MAP4K5* (5'-AGCAGGACTACGAACTCGTCCA-3'), *PTPN1* (5'-ACTTTCTTGATATCAACGGAAG-3'), *RASA2* (5'-CCCACTAGAGAACTGTTGCAT-3') and non-target control (5'-ACGCGTGCGTAATGAGAGGATC-3'). Combinatorial sgRNA expression vectors were transduced into the orthogonal clonal line 2 used in the orthogonal screen. For arrayed validation experiments, genetic interaction scores ( $GI_v$ ) were calculated as:

$$GI_v = (\tau_{activation + knockout}) - (\tau_{activation} + \tau_{knockout})$$

$\Psi$  scores from arrayed validation data ( $\Psi_v$ ) were calculated as:

$$\Psi_v = \tau_{activation} \times \tau_{knockout} \times GI_v^2$$

### Cell viability assays

Cells were seeded at 10,000 cells per 96-well in 200  $\mu$ L RPMI-1640 (10% FBS, 1% Anti-Anti) with indicated imatinib and/or R428 concentrations. Viability was determined at indicated time points by mixing 100  $\mu$ L cell suspension with 50  $\mu$ L resazurine medium (50  $\mu$ g/mL, Acros Organics). After 2h incubation, fluorescence was quantified on a plate reader (BMG Labtech) at excitation: 530 nm and emission: 590 nm.

## RNAseq

RNA from K562 CRISPRa cells was extracted using RNeasy Mini kit (Qiagen). Sequencing libraries were prepared using the TruSeq mRNA stranded kit (Illumina) and sequenced via SE 50bp RNAseq on a HiSeq2000 platform. Reads were aligned to *Homo sapiens* Ensembl GRCm38v.78 using STAR\_2.4.2a.

## Quantitative RT-PCR

Total RNA from sgRNA expressing cells was purified using Rneasy Mini columns (Qiagen). Taqman probe assays (Applied Biosystems) were used with FAM labelled probes for target genes and VIC labelled probes for the housekeeping gene HPRT1. Reactions were carried out using the one step qRT-PCR master mix TaqMan RNA-to-C<sub>T</sub> (Applied Biosystems) according to the manufacturer's instructions on the 2900 HT Fast RT-PCR machine (Applied Biosystems).

## Western blot analyses

NF1-Null HEK293T cells were generated using SpCas9 and sgRNA targeting exon 2 with the sequence 5'-AGTCAGTACTGAGCACAACA-3' (Shalem, O., et al., 2013). Following single cell cloning, target sequence amplification by PCR, TOPO cloning, and Sanger sequencing, both NF-1 alleles were confirmed deleted by a 1bp insertion resulting in *NF1*(N39fs) and a 11bp deletion resulting in *NF1*(S35fs). HEK293T cells were transfected with pcDNA3.1 Flag-eGFP (CTRL) and Flag-SPRED2 (SPRED2) using Lipofectamine 2000 (ThermoFisher Scientific, 11668019), serum starved for 24 hours, and stimulated with 20ng/ml recombinant human EGF (Invitrogen, PHG0311). Cells were washed with PBS and lysed in TNM buffer (0.2 M Tris pH 7.5, 1% Triton X-100, 1.5M NaCl, 50 mM MgCl<sub>2</sub>, 1mM DTT, protease and phosphatase inhibitor cocktails). Lysate was cleared and 1,000ug protein was subject GST-Raf1 RBD agarose beads (McCormick lab, in house) for 1.5 hours. Samples were analysed by Western blot using the following antibodies: NF1 (SCBT, sc-67 [D]), Flag (Sigma, F1804), pan-Ras (Cytoskeleton, Inc, AESA02), β-Actin (Sigma, A5441).

K562 orthogonal cells were transduced with lentivirus expressing *S.aureus* sgRNAs: sgNTC 5'-ACGCGTGCCTAATGAGAGGATC-3' (*NF1*-wildtype) or sgNF1 5'-TTGTCTTTGGGTGTATTAGCAA-3' (*NF1*-knockout). At 2 days post transduction, *NF1*-wildtype and -knockout cells were selected with puromycin (2 ug/mL) for 5 days and recovered for 2 additional days before treatment with vehicle (DMSO), 300nM imatinib, 1μM R428, or imatinib + R428 for 48hrs. Whole cell lysates were collected by lysis in RIPA buffer (150 mM NaCl, 1.0% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris pH 8.0) and lysate concentration determined by BCA assay (ThermoFisher). 20 μg of lysate corresponding to each sample was separated by SDS-PAGE and transferred to a PVDF membrane by wet transfer (250mA, 2hrs). Western blot analysis was carried out following standard conditions using p-AXL antibody (R&D Technologies; Y779), followed by re-probing with β-actin antibody (Sigma; AC-74) to confirm equal loading.

## Time-resolved fluorescence resonance energy transfer (TR-FRET) assay

TR-FRET assay utilizing the terbium/Venus as energy donor/acceptor was performed as described previously<sup>35</sup>. Briefly, HEK293T cells were transfected with GST- and Venus-

tagged genes. Cells were lysed in FRET buffer (20 mM Tris, pH 7.0, 0.01% Nonidet-P40, and 50 mM NaCl with proteinase and phosphatase inhibitors) followed by three freeze-and-thaw cycles. Terbium conjugated Anti-GST antibody (Cisbio Bioassays, Codolet, France) was 1:1000 diluted in FRET buffer and dispensed into each well with Multidrop™ Combi Reagent Dispenser (ThermoScientific). The lysate-antibody mixtures were incubated at 4°C before the TR-FRET signal was recorded (EnVision reader setting: Ex 337 nm, Em1: 520 nm, Em2: 486 nm; mirror: D400/D505 dual; time delay: 50 μs). The TR-FRET signal is expressed as the FRET ratio ( $F_{520}/F_{486} \times 10^4$ ).

### Data availability

Sequencing data from the CRISPRa screen and RNAseq are available at Sequence Read Archive accession number SRP127017 under BioProject ID PRJNA422995. All relevant additional data has been published with the manuscript, either as part of the main text or in the supplement. Plasmids and their sequences are deposited at Addgene.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

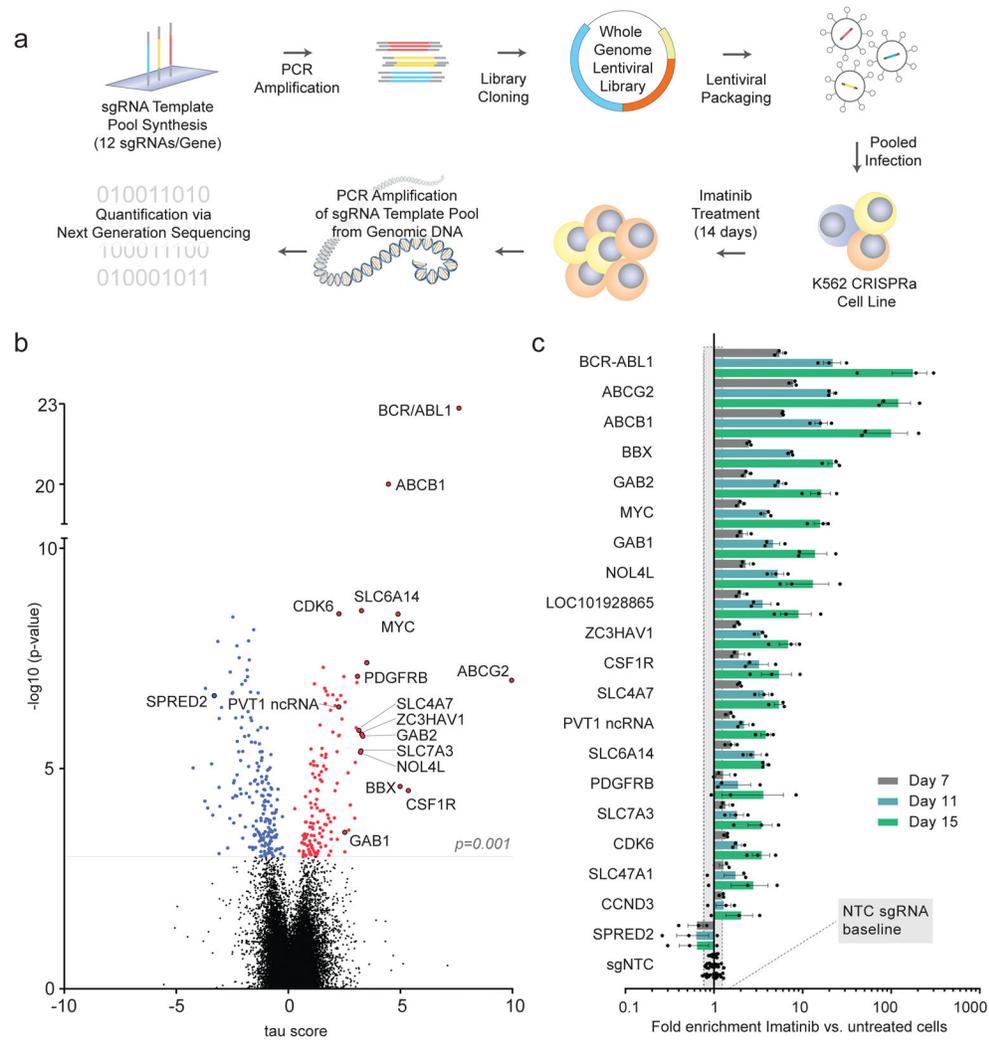
Special thanks go to members of the McManus Lab who provided critical feedback during the course of this project and Elizabeth Cahill for excellent technical support. We also thank Luke A. Gilbert and Marvin E. Tanenbaum for sharing the CRISPRa cell line ahead of its publication. M.T.M. was supported by NIH/CTD<sup>2</sup> (U01CA168370) and IDG (1U01MH105028). M.K. was supported by NIH/NIGMS New Innovator Award DP2 GM119139, NIH/NCI K99/R00 CA181494, a Stand Up to Cancer Innovative Research Grant and the Chan Zuckerberg Biohub. J.B. was supported by NIH Training grant T32 GM00715 and an AFPE Predoctoral Fellowship. H.F. was supported by NIH/CTD<sup>2</sup> (U01CA168449).

### References

1. Beltrao P, Cagney G, Krogan NJ. Quantitative genetic interactions reveal biological modularity. *Cell*. 2010; 141:739–745. DOI: 10.1016/j.cell.2010.05.019 [PubMed: 20510918]
2. Fischer B, et al. A map of directional genetic interactions in a metazoan cell. *eLife*. 2015; 4
3. Drees BL, et al. Derivation of genetic interaction networks from quantitative phenotype data. *Genome biology*. 2005; 6:R38. [PubMed: 15833125]
4. St Onge RP, et al. Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nature genetics*. 2007; 39:199–206. DOI: 10.1038/ng1948 [PubMed: 17206143]
5. Avery L, Wasserman S. Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *Trends in genetics : TIG*. 1992; 8:312–316. [PubMed: 1365397]
6. Costanzo M, et al. A global genetic interaction network maps a wiring diagram of cellular function. *Science*. 2016:353.
7. Wright AV, Nunez JK, Doudna JA. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell*. 2016; 164:29–44. DOI: 10.1016/j.cell.2015.12.035 [PubMed: 26771484]
8. Shalem O, Sanjana NE, Zhang F. High-throughput functional genomics using CRISPR-Cas9. *Nature reviews. Genetics*. 2015; 16:299–311. DOI: 10.1038/nrg3899
9. Boettcher M, McManus MT. Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. *Molecular cell*. 2015; 58:575–585. DOI: 10.1016/j.molcel.2015.04.028 [PubMed: 26000843]

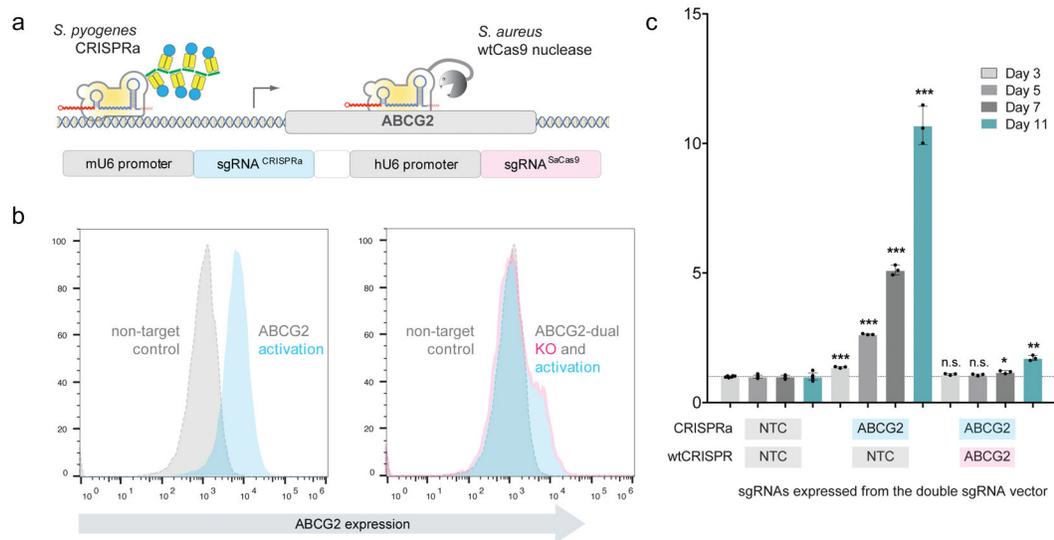
10. O'Hare T, Deininger MW, Eide CA, Clackson T, Druker BJ. Targeting the BCR-ABL signaling pathway in therapy-resistant Philadelphia chromosome-positive leukemia. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2011; 17:212–221. DOI: 10.1158/1078-0432.CCR-09-3314 [PubMed: 21098337]
11. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell*. 2014
12. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*. 2009; 4:44–57. DOI: 10.1038/nprot.2008.211 [PubMed: 19131956]
13. Dhillon AS, Hagan S, Rath O, Kolch W. MAP kinase signalling pathways in cancer. *Oncogene*. 2007; 26:3279–3290. DOI: 10.1038/sj.onc.1210421 [PubMed: 17496922]
14. Milojkovic D, Apperley J. Mechanisms of Resistance to Imatinib and Second-Generation Tyrosine Inhibitors in Chronic Myeloid Leukemia. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2009; 15:7519–7527. DOI: 10.1158/1078-0432.CCR-09-1068 [PubMed: 20008852]
15. de Jong R, ten Hoeve J, Heisterkamp N, Groffen J. Crkl is complexed with tyrosine-phosphorylated Cbl in Ph-positive leukemia. *The Journal of biological chemistry*. 1995; 270:21468–21471. [PubMed: 7545163]
16. Cilloni D, Saglio G. Molecular pathways: BCR-ABL. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2012; 18:930–937. DOI: 10.1158/1078-0432.CCR-10-1613 [PubMed: 22156549]
17. Cong F, et al. Cytoskeletal protein PSTPIP1 directs the PEST-type protein tyrosine phosphatase to the c-Abl kinase to mediate Abl dephosphorylation. *Molecular cell*. 2000; 6:1413–1423. [PubMed: 11163214]
18. Calvisi DF, et al. Inactivation of Ras GTPase-activating proteins promotes unrestrained activity of wild-type Ras in human liver cancer. *Journal of hepatology*. 2011; 54:311–319. DOI: 10.1016/j.jhep.2010.06.036 [PubMed: 21067840]
19. Kuo TC, Chavarria-Smith JE, Huang D, Schlissel MS. Forced expression of cyclin-dependent kinase 6 confers resistance of pro-B acute lymphocytic leukemia to Gleevec treatment. *Molecular and cellular biology*. 2011; 31:2566–2576. DOI: 10.1128/MCB.01349-10 [PubMed: 21536647]
20. Sherr CJ, Beach D, Shapiro GI. Targeting CDK4 and CDK6: From Discovery to Therapy. *Cancer discovery*. 2016; 6:353–367. DOI: 10.1158/2159-8290.CD-15-0894 [PubMed: 26658964]
21. Cheah CY, et al. Patients with myeloid malignancies bearing PDGFRB fusion genes achieve durable long-term remissions with imatinib. *Blood*. 2014; 123:3574–3577. DOI: 10.1182/blood-2014-02-555607 [PubMed: 24687085]
22. Li F, et al. FGFR-Mediated Reactivation of MAPK Signaling Attenuates Antitumor Effects of Imatinib in Gastrointestinal Stromal Tumors. *Cancer discovery*. 2015; 5:438–451. DOI: 10.1158/2159-8290.CD-14-0763 [PubMed: 25673643]
23. Chase A, et al. Imatinib sensitivity as a consequence of a CSF1R-Y571D mutation and CSF1/CSF1R signaling abnormalities in the cell line GDM1. *Leukemia*. 2009; 23:358–364. DOI: 10.1038/leu.2008.295 [PubMed: 18971950]
24. Dufies M, et al. Mechanisms of AXL overexpression and function in Imatinib-resistant chronic myeloid leukemia cells. *Oncotarget*. 2011; 2:874–885. DOI: 10.18632/oncotarget.360 [PubMed: 22141136]
25. Ran FA, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*. 2015; 520:186–191. DOI: 10.1038/nature14299 [PubMed: 25830891]
26. Konermann S, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2015; 517:583–588. DOI: 10.1038/nature14136 [PubMed: 25494202]
27. Qi LS, et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*. 2013; 152:1173–1183. DOI: 10.1016/j.cell.2013.02.022 [PubMed: 23452860]
28. Vojta A, et al. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic acids research*. 2016; 44:5615–5628. DOI: 10.1093/nar/gkw159 [PubMed: 26969735]

29. Dahlman JE, et al. Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nature biotechnology*. 2015; 33:1159–1161. DOI: 10.1038/nbt.3390
30. Stowe IB, et al. A shared molecular mechanism underlies the human rasopathies Legius syndrome and Neurofibromatosis-1. *Genes & development*. 2012; 26:1421–1426. DOI: 10.1101/gad.190876.112 [PubMed: 22751498]
31. Graham DK, DeRyckere D, Davies KD, Earp HS. The TAM family: phosphatidyserine sensing receptor tyrosine kinases gone awry in cancer. *Nature reviews. Cancer*. 2014; 14:769–785. DOI: 10.1038/nrc3847 [PubMed: 25568918]
32. Gay CM, Balaji K, Byers LA. Giving AXL the axe: targeting AXL in human malignancy. *British journal of cancer*. 2017; 116:415–423. DOI: 10.1038/bjc.2016.428 [PubMed: 28072762]
33. Postel-Vinay S, Ashworth A. AXL and acquired resistance to EGFR inhibitors. *Nat Genet*. 2012; 44:835–836. DOI: 10.1038/ng.2362 [PubMed: 22836088]
34. Sheridan C. First Axl inhibitor enters clinical trials. *Nature biotechnology*. 2013; 31:775–776. DOI: 10.1038/nbt0913-775a
35. Li Z, et al. The OncoPPI network of cancer-focused protein-protein interactions to inform biological insights and therapeutic strategies. *Nature communications*. 2017; 8:14356.
36. Laufer C, Fischer B, Billmann M, Huber W, Boutros M. Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. *Nat Methods*. 2013; 10:427–431. DOI: 10.1038/nmeth.2436 [PubMed: 23563794]
37. Roguev A, et al. Quantitative genetic-interaction mapping in mammalian cells. *Nat Methods*. 2013; 10:432–437. DOI: 10.1038/nmeth.2398 [PubMed: 23407553]
38. Bassik MC, et al. A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. *Cell*. 2013; 152:909–922. DOI: 10.1016/j.cell.2013.01.030 [PubMed: 23394947]
39. Vizeacoumar FJ, et al. A negative genetic interaction map in isogenic cancer cell lines reveals cancer cell vulnerabilities. *Molecular systems biology*. 2013; 9:696. [PubMed: 24104479]
40. Shen JP, et al. Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic interactions. *Nature methods*. 2017
41. Han K, et al. Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nature biotechnology*. 2017
42. Blomen VA, et al. Gene essentiality and synthetic lethality in haploid human cells. *Science*. 2015; 350:1092–1096. DOI: 10.1126/science.aac7557 [PubMed: 26472760]
43. Du D, et al. Genetic interaction mapping in mammalian cells using CRISPR interference. *Nature methods*. 2017
44. Kampmann M, Bassik MC, Weissman JS. Integrated platform for genome-wide screening and construction of high-density genetic interaction maps in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*. 2013; 110:E2317–2326. DOI: 10.1073/pnas.1307002110 [PubMed: 23739767]
45. Quintas-Cardama A, Cortes J. Molecular biology of bcr-abl1-positive chronic myeloid leukemia. *Blood*. 2009; 113:1619–1630. DOI: 10.1182/blood-2008-03-144790 [PubMed: 18827185]
46. Gilbert LA, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*. 2014
47. LaMontagne KR Jr, Flint AJ, Franza BR Jr, Pandergast AM, Tonks NK. Protein tyrosine phosphatase 1B antagonizes signalling by oncoprotein tyrosine kinase p210 bcr-abl in vivo. *Molecular and cellular biology*. 1998; 18:2965–2975. [PubMed: 9566916]
48. Bae S, Park J, Kim JS. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 2014; 30:1473–1475. DOI: 10.1093/bioinformatics/btu048 [PubMed: 24463181]
49. Burk O, Klempnauer KH. Myb and Ets transcription factors cooperate at the myb-inducible promoter of the tom-1 gene. *Biochimica et biophysica acta*. 1999; 1446:243–252. [PubMed: 10524199]
50. Shannon P, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 2003; 13:2498–2504. DOI: 10.1101/gr.1239303 [PubMed: 14597658]



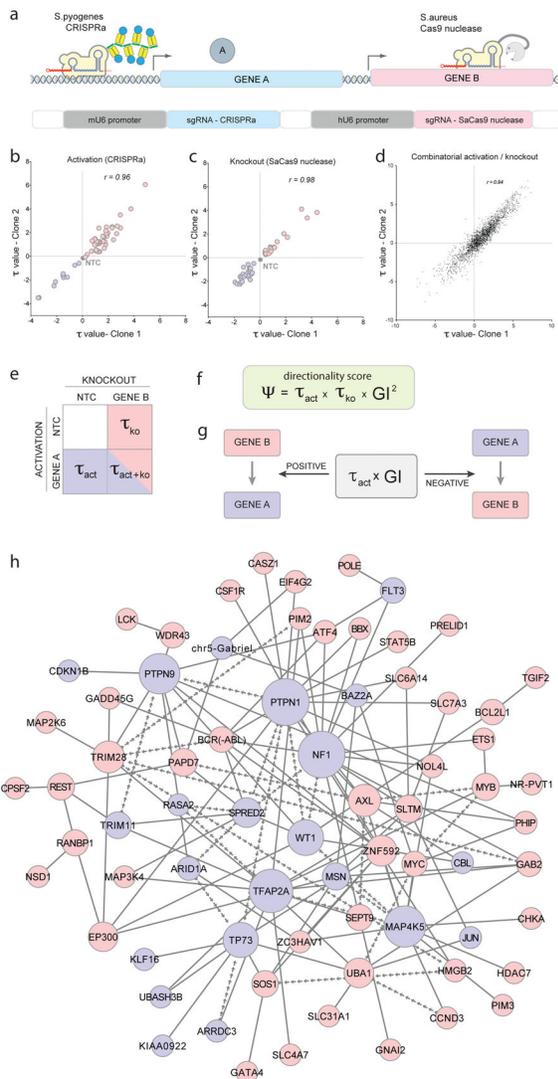
**Figure 1. Ultra-complex CRISPRa screen identifies hundreds of genes involved in cancer signalling pathways**

**a**, Schematic of genome-scale CRISPRa screening approach (see text for details). **b**, Overview of CRISPRa screen results. Negative  $\tau$  values indicate depletion and positive values enrichment of cells following imatinib selection. Significant candidate genes (FDR < 0.05,  $p < 0.001$ ) are in colour (blue = depleted, red = enriched). Validated candidate genes are labelled in black. Mann-Whitney U test was used to calculate p-values as described previously<sup>44</sup>. To correct for multiple hypothesis testing, we first performed random sampling with replacement among the set of  $\tau$  values for non-targeting control sgRNAs and calculated p-values for each sampling. Then, we calculated the false discover rate (FDR) based on the distribution of p-values for all genes in the library and for non-targeting controls generated above. **c**, Candidate gene validation. Enrichment of candidate sgRNA expressing cells was measured over time. Values represent the mean of three different sgRNAs targeting each gene with s.e.m. Grey shading = two standard deviations of sgNTCs at day 15. All values from separate sgRNAs on days 7, 11 and 15 normalised to baseline or untreated cells are shown in Supplemental Table 3.



**Figure 2. The orthogonal CRISPR system**

**a**, Schematic of the orthogonal system on the example of imatinib efflux transporter *ABCG2*. Combination of CRISPR systems from *S.pyogenes* (CRISPRa) and *S.aureus* (Cas9 nuclease) allows the simultaneous activation and knockout of genes in the same cell simply by expressing two appropriate sgRNAs. **b**, Orthogonal system is able to modulate *ABCG2* protein levels. Flow cytometry analysis of *ABCG2* levels following CRISPRa mediated activation of *ABCG2* without (left) or with (right) SaCas9 nuclease mediated knockout of *ABCG2* (grey histogram = sgNTC for both CRISPR systems). A representative result from  $n>10$  independent experiments with similar results is shown. **c**, Orthogonal system can control imatinib response. Enrichment of imatinib treated cells with activated *ABCG2* with/out SaCas9 nuclease mediated knockout of *ABCG2*. Values represent the mean of independent experiments ( $n=3$ ) with s.e.m. and statistical significance was determined via two-tailed, homoscedastic t-test with \* =  $p<0.05$ , \*\* =  $p<0.01$  and \*\*\* =  $p<0.001$ .



**Figure 3. Orthogonal CRISPR screens can quantify directional genetic interactions**  
**a**, Concept of the application of the orthogonal system for directional gene interaction studies. In the same cell, one gene is activated (CRISPRa) while another gene is knocked out (SaCas9 nuclease). **b–d**, Correlation of  $\tau$  values from two clonal cell line replicates is shown for **b**, gene activation, **c**, gene knockout and **d**, all possible combinations thereof. Correlation values ( $r$ ) are Pearson product-moment correlation coefficients. **e**, Schematic of perturbation data set from each gene pair (blue = depleted, red = enriched, NTC = non-target control sgRNA) **f**, Formula for calculating  $\Psi$  scores. Negative  $\Psi$  scores define interactions in which directionality could be inferred. **g**, To determine which of both interaction partners acts up- or downstream,  $\tau_{\text{activation}}$  values were multiplied with genetic interaction scores. Positive values indicate a downstream function, negative values an upstream function of the activated gene. **h**, Based on GI and  $\Psi$  scores determined by the full orthogonal interaction screen, a genetic interaction model was constructed. For positive regulators of cell fitness, nodes are shown in red and negative regulators in blue. Arrow-shaped edges indicate inferred directional interactions between nodes. Line-shaped edges symbolise genetic

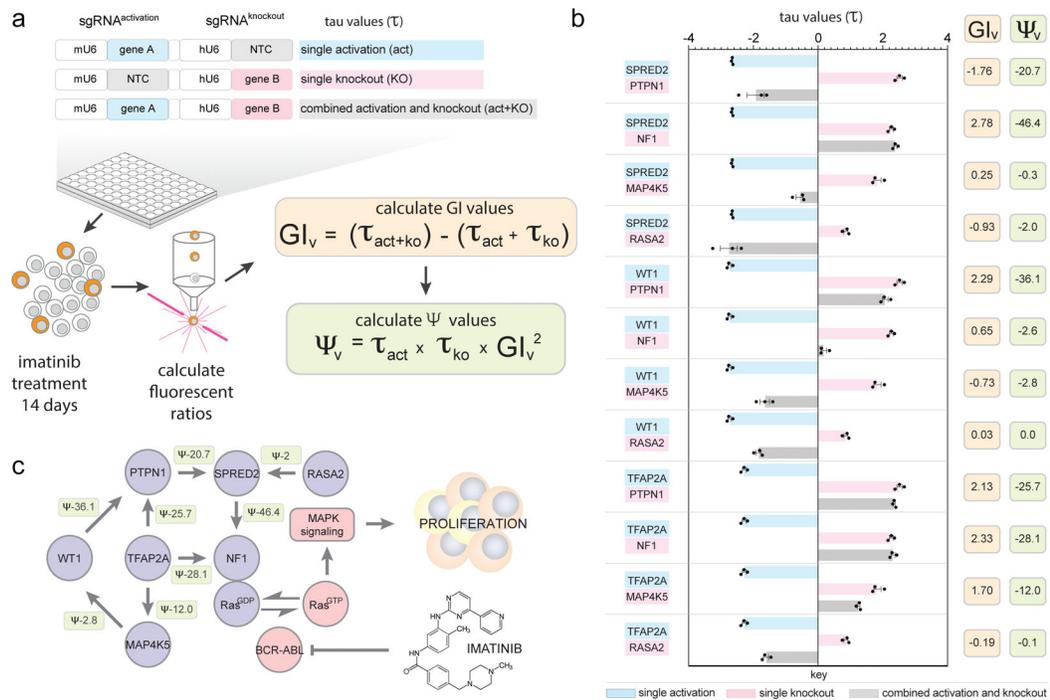
interactions where directionality could not be inferred. Node sizes are proportional to the degree of connectivity. In total, 2258 gene:gene combinations that passed the filter criteria were considered for the construction of the directional genetic interaction network.

Author Manuscript

Author Manuscript

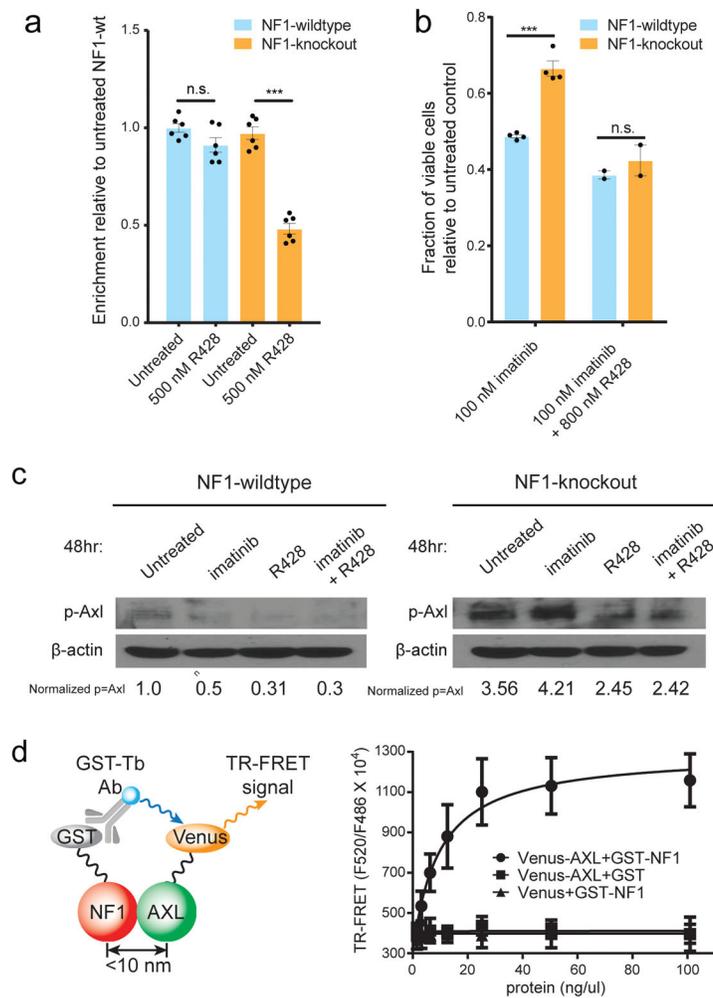
Author Manuscript

Author Manuscript



**Figure 4. Validation of a directional Ras-centric genetic sub-network**

**a**, Relative fitness ( $\tau$ ) was measured over 14 days following gene activation, knockout or the combination of both. From those values, genetic interaction ( $GI_v$ ) as well as directionality ( $\Psi_v$ ) scores were calculated. NTC = non-target control sgRNA. **b**, Twelve activation/knockout combinations were re-tested in an arrayed format from which ten were predicted by the orthogonal screen to show a directional genetic interaction. Eight combinations displayed the same trend of directional interactions predicted by the orthogonal screen data while two interactions did not reproduce (see also Supplemental Table 9). Single perturbation, and combinatorial  $\tau$  values are shown following 14 days of imatinib selection (mean with s.e.m. from technical replicates ( $n=3$ )) along with calculated  $GI_v$  and  $\Psi_v$  scores for each gene:gene combination. **c**, A directional genetic interaction model was assembled based on validated interactions from **b**. Arrows indicate the direction of the functional dependencies as explained in the text but do not suggest direct physical interactions. Values represent  $\Psi_v$  scores calculated from  $\tau$  values in **b**. Each directional interaction was reproduced three times independently.



### Figure 5. Exploiting genetic dependencies for cancer therapy

**a**, *NF1*-knockout K562 cells are significantly more sensitive to the AXL kinase inhibitor R428 than *NF1*-wildtype cells. Cells were treated for 8 days with 500 nM R428 on day 0 and day 4. (mean with s.e.m. from technical replicates (n=6)). **b**, *NF1*-knockout K562 cells are significantly more resistant to imatinib but can be re-sensitised by R428 treatment (mean with s.e.m. from technical replicates (n=4 for imatinib and n=2 for imatinib + R428 treated cells)). In panel **a** and **b**, statistical significance was determined via two-tailed, homoscedastic t-test with \* =  $p < 0.05$ , \*\* =  $p < 0.01$  and \*\*\* =  $p < 0.001$ . **c**, *NF1*-knockout cells accumulate elevated levels of phosphorylated AXL kinase (p-AXL) which can be reduced by treatment with the AXL kinase inhibitor R428. Quantification of the ratio of band intensity from p-AXL/ $\beta$ -actin, normalised to p-AXL levels in *NF1*-wt untreated cells is shown. The experiment was performed once. **d**, TR-FRET assay shows direct interaction between NF1 and AXL in HEK293T cells. Shown is the mean with s.d. from three independent experiments.