CRISPR/Cas9 genome-wide sgRNA libraries for loss-of-function and gain-of-function genetic screens

CELLECTA
www.cellecta.com

Donato Tedesco[1], Mikhail Makhanov[1], Debbie Deng[1],Sylvain Baron[1], Paul Diehl[1], Dmitry Suchkov[1], Costa Frangou[1], and Alex Chenchik[1].
[1]Cellecta, Inc., Mountain View, CA USA

Abstract

Genome-wide loss-of-function screening is a fundamental method to identify genes responsible for driving biological processes. Complex pooled lentiviral-based libraries expressing large numbers of genetic disruptors, such as shRNAs (RNAi) or sgRNAs (CRISPR), make large-scale cell screening practical.

Although loss-of-function shRNA and sgRNA pooled library screens are similar in concept, the gene loss of function is achieved by different mechanisms (mRNA degradation with RNAi, full gene disruption with CRISPR-KO, transcriptional inhibition with CRISPRi), so some divergences are expected and indeed observed when comparing results obtained using one method versus the other. Furthermore, contrary to RNAi, CRISPR technology can be modified to activate gene expression (CRISPRa), thus enabling the use of genome-wide gain-of-function screening in gene function studies. Here we present and compare results from genetic screens in human cancer cells using CRISPR-KO, CRISPRi, CRISPRa and RNAi technologies. To investigate differences in the RNAi and CRISPR HTGS technologies, we performed parallel viability screens using sgRNA and shRNA libraries targeting the same set of ~ 6,300 genes with the same number of effectors for each gene. The screens were aimed at the identification of essential genes in a pair of primary isogenic PDX-derived CML cell lines. The screen results show significant but not complete overlap in the genes identified by each technology in each cell line, indicating that both approaches are effective at identifying the majority of essential genes in a cell system. However, deeper analysis revealed that a small number of genes could only be identified by either CRISPR or RNAi technology. Importantly, by combining data from the two screening methodologies, a higher number of viability genes and pathways could be identified and subsequently validated by independent cell-based assays with a high confirmation rate. To assess the performance of CRISPRa and CRISPRi HTGS technologies, multiple parallel screens were performed in MDA-MB-231 cells with different sgRNA design libraries targeting the same set of ~2000 genes. Gene target identification rates achieved by the different libraries were then used to identify the best sgRNA designs for CRISPRa and CRISPRi applications.

sgRNA Scaffold Improvement

Natural CRISPR Target Complex in bacteria

crRNA: 5' - NNNNNNNNNNNNNNNNNNN GUUUUAGAGCUAUGCUGUUUG 3'
tracrRNA: 5' - UCGUGGCUGAGCCACGGUGAAAAAGUUCAACUAUUGCCUGAUCGGAAUAAAAUU CGAUACGACAAACUUACCAAGG - 5'
GAA

Standard CRISPR single-guide RNA (sgRNA)

sgRNA (wt): 5' - NNNNNNNNNNNNNNNNNNN GUUUUAGAGCUA G A
3' - CGUGGCUGAGCCACGGUGAAAAAGUUCAACUAUUGCCUGAUCGGAAUAAAAUU CGAU A A
GAA

Improved CRISPR single-guide RNA (sgRNA)

sgRNA (HEAT): 5' - NNNNNNNNNNNNNNNNNNN GAUUUAGAGCUA UGCUGU G A
3' - CGUGGCUGAGCCACGGUGAAAAAGUUCAACUAUUGCCUGAUCGGAAUUAAAAUU CGAU ACGACA A A
GAA

Modified sgRNA structure increases knockout effectiveness

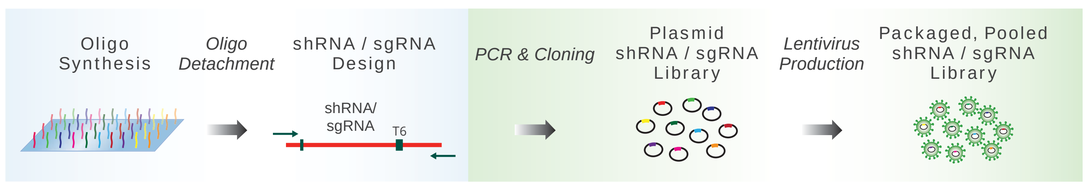
- Improves robustness of screen by increasing signal (sgRNA depletion or enrichment) over random genetic variation (noise).

A-T inversion removes transcriptions stop signal of TTTTT

Extension of first loop to increase Cas9 binding

Combined modifications produce **H**elix **E**xtended **A-T** inversion ("**HEAT**") sgRNA scaffold

Library Construction and Screening



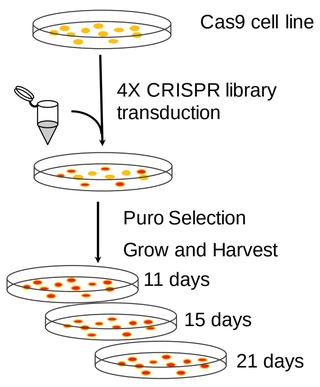
sgRNA Structure Improvement

4X CRISPR LIBRARY:

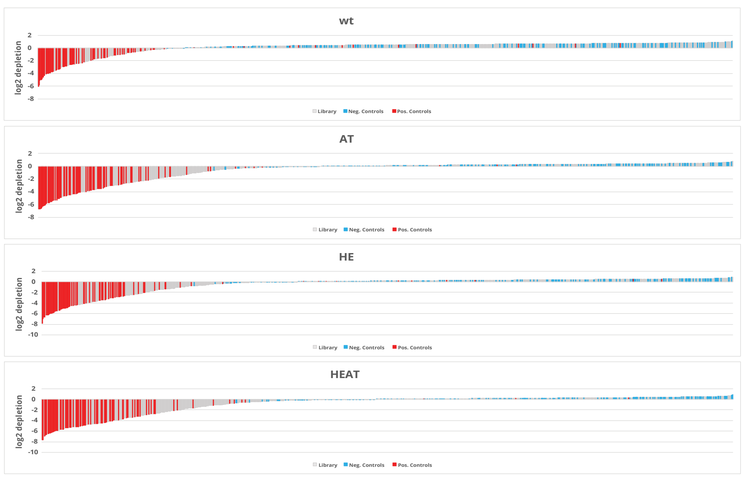
- ~100 genes, 8 sgRNAs/gene (including known essential genes)
- ~100 non-targeting sgRNAs (negative controls)
- ~100 intron-targeting sgRNAs (negative controls)

Each sgRNA repeated 4X:

- sgRNA-wt structure
- sgRNA-AT structure
- sgRNA-HE structure
- sgRNA-HEAT structure



sgRNA structure optimization screen

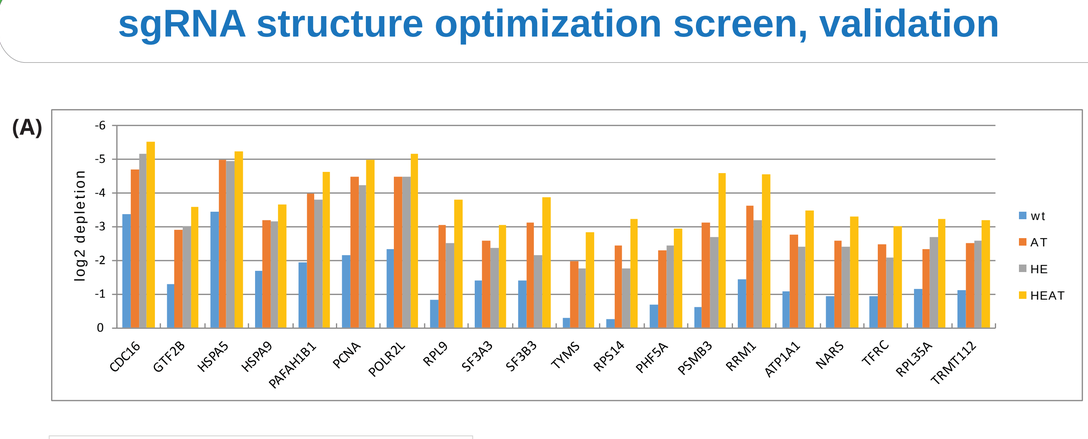


sgRNA handle designs:

- **wt:** wt Tracr sequence
- **AT:** TTTT terminator elimination
- **HE:** stem-loop stabilization
- **HEAT:** HE + AT modifications

Figure 1. To identify the best performing sgRNA design, we built a sgRNA design optimization library of different handle structures, all targeting the same ~900 human genomic sites. Targeted sites included essential genes, non-essential genes, introns, and non-human control sequences. In order to identify the best performing sgRNA design, the optimization library was transduced into CAS9 expressing cell lines validated with CRISPRTest Cas9 Assay Kit [2]. Samples were then collected at multiple time points post transduction. On the left, the waterfall plots display the depletion fold of positive and negative controls within the library for each design.

sgRNA structure optimization screen, validation



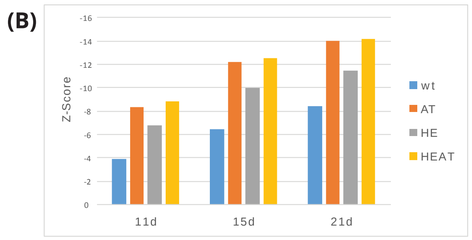
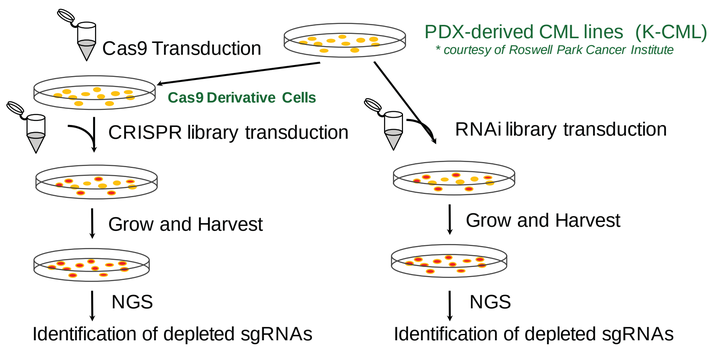


Figure 2. Based on the calculated depletion folds of positive controls (A) in the design optimization screen, Z-scores were calculated for each sgRNA handle design (B). Following analysis of the depletion values and Z-scores across all the tested cell lines at multiple time points, the HEAT design was identified as the best performing and selected for the construction of Cellecta's human and mouse genome-wide CRISPR libraries.

CRISPR vs RNAi Screens



CRISPR and RNAi screen comparison with 50k libraries targeting the same 6,200 genes each with 8 sgRNA or shRNA per gene.

- Results largely overlap with a few more hits in the CRISPR screen
- CRISPR screen has much less backround. Off-target noise very smalled compared with shRNA

Analysis of Controls

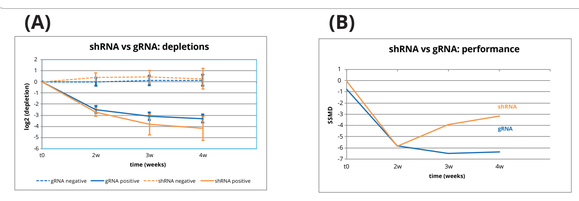


Figure 3. Parallel CRISPR/RNAi Screens in CML-R cell line: Analysis of Controls. (a) Changes in representation (depletion) of negative and positive (known cytotoxic) controls in the shRNA and gRNA libraries was monitored over a 4 week period. (b) The standardized mean depletion difference between positive and negative controls was calculated, as a means to compare RNAi vs. CRISPR screening platform performance.

shRNA / gRNA Hit Distribution

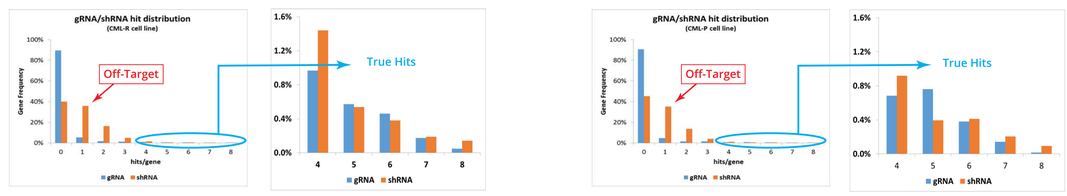


Figure 4. Measured distribution of gRNA and shRNA hits within target genes in CML-R and CML-P cell lines. In both cell lines, about 35% of target genes in the shRNA library were identified by a single shRNA hit (for a total of ~2,300 shRNAs in the shRNA library) compared to only about 5% genes (~320 gRNAs) in the gRNA library. Single shRNA/gRNA hits are most likely off-target hits. Target genes identified by multiple hits (4 or more) were found at comparable frequencies in CRISPR and RNAi screens. Multiple shRNA/gRNA hits are most likely on-target hits.

RNAi / CRISPR Gene Hit Overlap

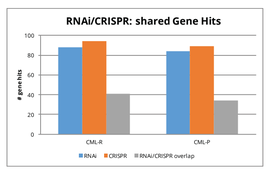
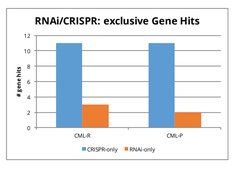


Figure 5. Gene hit analysis: Overlap and Reproducibility. RNAi/CRISPR overlap: the number of RNAi and CRISPR gene hits are displayed for the CML-R and CML-P cell lines. The overlap between RNAi and CRISPR hits is reported in the grey bar.

RNAi / CRISPR Unique Gene Hits



CRISPR		RNAi	
CML-R	CML-P	CML-R	CML-P
ATP6V0B	ATP6V0B	MTAP	OTOGL
CCNH	CHORDC1	TAF1L	UBC
DHFR	DHDDS		
ERCC2	DHFR		
FARSA	ERCC2		
FARSB	FARSA		
NARFL	NARFL		
PGAM1	PGAM1		
RNMT	SCYL1		
SCYL1	TRMT112		
TRMT112	YARS2		

Figure 6. The number and lists of gene hits exclusively identified by either CRISPR or RNAi technology (no signal with other technology) are displayed.

Pathway Analysis

Pathway	# genes	Z Score	AdjustedP	Pathway	# genes	Z Score	AdjustedP
Cytosolic tRNA aminoacylation	16	14.2	0.005	purine nucleotides de novo biosynthesis II	5	7.4	0.005
Activation of APC/C and APC/C:Cdc20 mediated degradation of mitotic proteins	12	12.2	0.005	Cap-dependent Translation Initiation	7	7.3	0.005
Regulation of APC/C activators between G1/S and early anaphase	12	11.8	0.005	**Transferrin endocytosis and recycling**	4	**6.6**	**0.005**
CDT1 association with the CDC6:ORC:origin complex	10	11.1	0.005	Pyrimidine biosynthesis	2	6.4	0.005
Vif-mediated degradation of APOBEC3G	9	11.0	0.005	Polo-like kinase mediated events	2	6.4	0.005
SCF-beta-TrCP mediated degradation of Emi1	9	11.0	0.005	Formation of RNA Pol II elongation complex	6	6.2	0.005
mRNA Splicing - Major Pathway	7	10.9	0.005	**Signaling by Insulin receptor**	7	**5.9**	**0.005**
Cell Cycle Checkpoints	15	10.7	0.005	Nucleotide Excision Repair	5	5.8	0.005
Synthesis of DNA	10	10.7	0.005	E2F transcription factor network	7	4.6	0.005
Stabilization of p53	10	10.7	0.005	**superpathway of geranylgeranyldiphosphate biosynthesis I (via mevalonate)**	3	**4.4**	**0.005**
Orc1 removal from chromatin	10	10.7	0.005	pyrimidine deoxyribonucleotides de novo biosynthesis	3	4.4	0.005
Autodegradation of the E3 ubiquitin ligase COP1	10	10.7	0.005	Insulin Synthesis and Processing	3	4.4	0.005
Vpu mediated degradation of CD4	9	10.5	0.005	ATM pathway	18	4.3	0.005
Destabilization of mRNA by AUF1 (hnRNP D0)	9	10.5	0.005	Validated targets of C-MYC transcriptional activation	7	3.8	0.005
Cyclin A:Cdk2-associated events at S phase entry	13	10.4	0.005	**Iron uptake and transport**	4	**4.4**	**0.009**
Ubiquitin Mediated Degradation of Phosphorylated Cdc25A	10	10.3	0.005	Mitochondrial tRNA aminoacylation	3	3.2	0.009
Ubiquitin-dependent degradation of Cyclin D1	9	10.1	0.005	**methionine degradation I (to homocysteine)**	1	**4.1**	**0.012**
Signaling by Wnt	9	10.1	0.005	Unfolded Protein Response	2	3.4	0.012
SCF(Skp2)-mediated degradation of p27/p21	9	10.1	0.005	Regulation of nuclear beta catenin signaling and target gene transcription	5	2.3	0.012
Regulation of activated PAK-2p34 by proteasome mediated degradation	9	10.1	0.005	Golgi Cisternae Pericentriolar Stack Reorganization	2	3.4	0.016
Degradation of beta-catenin by the destruction complex	9	10.1	0.005	Recruitment of mitotic centrosome proteins and complexes	2	3.2	0.016
Cyclin E associated events during G1/S transition	13	9.9	0.005	Centrosome maturation	2	3.2	0.016
Regulation of ornithine decarboxylase (ODC)	9	9.7	0.005	PI3K Cascade	1	3.4	0.018
ER-Phagosome pathway	9	9.7	0.005	superpathway of cholesterol biosynthesis	3	2.8	0.018
Mitotic G1-G1/S phases	15	9.7	0.005	GPCR downstream signaling	1	3.4	0.021
RNA Polymerase II Transcription	11	8.6	0.005	p73 transcription factor network	4	2.5	0.021
mRNA Processing	10	8.5	0.005	Insulin receptor signalling cascade	2	2.5	0.031
Regulation of Apoptosis	9	8.4	0.005	**cysteine biosynthesis III (mammalia)**	1	**3.4**	**0.034**
Pausing and recovery of elongation	5	8.3	0.005	Activation of ATR in response to replication stress	1	3.1	0.040

Figure 7. Data from CRISPR and RNAi screens was combined for pathway analysis of gene hits. Identified pathways were filtered for redundancy and sorted by significance base on Z-scores and adjusted p-values (0.005<p<0.05). A subset of lower ranking pathways was then chosen for confirmation studies.

Confirmation Studies

Pathway	Tested genes	Protein	Inhibitory Drug	Drug Toxicity
Transferrin endocytosis and recycling / Iron uptake and transport Signaling by Insulin receptor	ATP6V0B ATP6V0C ATP6V1A ATP6V1D ATP6V1F	V-ATPase	Bafilomycin A1 Bafilomycin B1 Concanamycin A	+ + +
Superpathway of geranylgeranyldiphosphate biosynthesis I	FDPS	Farnesyl Diphosphate Synthase	Zoledronic acid Aledronate	+/- +/-
	HMGCS1	3-Hydroxy-3-Methylglutaryl-CoA Synthase 1	Hymeglusin	+
Methionine degradation I (to homocysteine)	AHCY	S-adenosylhomocysteine hydrolase	Neplanocin	+
	AHCYL1	Adenosylhomocysteinase 2		
Cysteine biosynthesis III (mammalia)	MAT2A	Methionine Adenosyltransferase II, Alpha	FIDAS-5	+

Figure 8. Confirmation studies on a subset of pathways identified by combined CRISPR/RNAi data analysis were performed. The cytotoxic effect of inhibitory compounds specific to the protein products of the identified target genes was assayed in CML-R cells.

CRISPRa and CRISPRi systems

Fig. 9

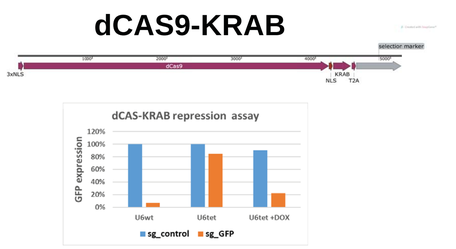


Figure 9. Cells expressing dCas9-KRAB repressor and CMV-GFP were transduced with targeting (sg_GFP) or non-targeting (sg_control) sgRNAs, expressed from either wt (U6wt) or tet-inducible (U6tet) U6 promoter. CMV repression was assayed by loss of GFP fluorescence intensity.

Figure 10. Cells expressing dCas9-VPH transactivator were transduced with sgRNAs targeting the promoters of the indicated genes. Promoter transactivation was measured by qRT-PCR

Fig. 10

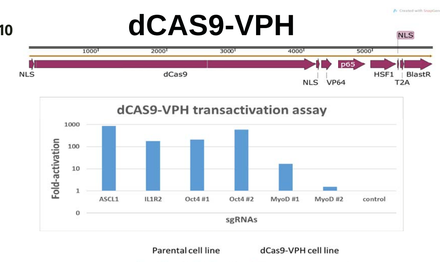


Fig. 11

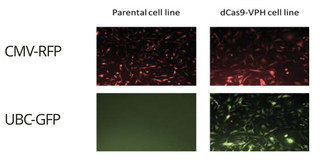


Figure 11. Parental cells and cells expressing dCas9-VPH transactivator (both expressing GFP and RFP from CMV and UBC promoters, respectively) were transduced with UBC-targeting sgRNA. UBC promoter transactivation was assayed by fluorescence intensity.

CRISPRa/i screens

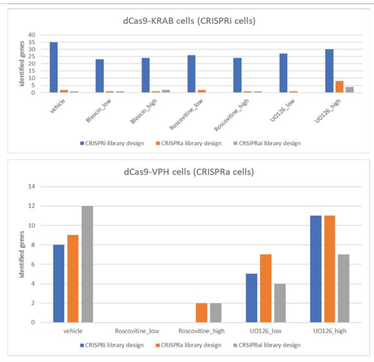


Figure 12. 3 different sgRNA design algorithms were tested by means of genetic screens in CRISPRa and CRISPRi cells. CRISPRa cells are MDA-MB-231 cells expressing the dCas9-VPH transcriptional activator, CRISPRi cells are MDA-MB-231 cells expressing the dCas9-KRAB transcriptional repressor. 3 separate libraries were built, each targeting the promoters of the same ~2,000 genes with a redundancy of 3 sgRNA/promoters, each library design based on a different algorithm for sgRNA target selection. The CRISPRa library uses an algorithm specifically developed for CRISPRa applications, the CRISPRi library uses an algorithm specifically developed for CRISPRi applications, the CRISPRai library uses an algorithm which does not differentiate between CRISPRa and CRISPRi applications for sgRNA target site selection. The 3 libraries were cotransduced into both CRISPRa and CRISPRi cells, and genetic screens were performed aimed at identifying genes which would promote or hinder viability under standard growth conditions, or incubation at high or low doses of protein kinase inhibitors roscovitine and UO126, and DNA damage causing agent Bleocin. The summarized results show that CRISPRi sgRNA library design works the best in CRISPRi screens, while both CRISPRa and CRISPRai designs seem to perform better than CRISPRi design in CRISPRa screens.

Conclusions

In the presented studies, the CRISPR/Cas9 screening platform showed comparable or better performance than RNAi.

- Modifications in the sgRNA structure enhance performance in CRISPR genetic screens
- Fewer off-target hits are seen in CRISPR vs RNAi screens
- More "pathway coverage" is seen in CRISPR vs RNAi screens
- In some instances, RNAi still performs better than CRISPR
- Combined CRISPR/RNAi screens outperform replicate CRISPR-only and RNAi-only screens
- CRISPRi-specific sgRNA design improves performance of loss-of-function CRISPRi screens
- CRISPRa-specific sgRNA design improves performance of gain-of-function CRISPRa screens