



Donato Tedesco¹, Mikhail Makhanov¹, Sylvain Baron¹, Dmitry Suchkov¹, Karim Hyder¹, and Alex Chenchik¹

¹Collecta, Inc., Mountain View, CA USA

Genome-wide loss-of-function screening is a fundamental method to identify genes responsible for driving biological responses, and complex pooled lentiviral-based libraries expressing large numbers of genetic disruptors, such as shRNAs, make large-scale cell screening practical. While RNAi-based approaches have proven to be an effective strategy for these screens, recent work suggests CRISPR technology offers an effective alternative. Although shRNA and sgRNA pooled library screens are similar in concept, the gene interruption with the two techniques occurs by a very different mechanism so some divergence may be expected when comparing results obtained using one method versus the other.

CRISPR technology relies on the stability and specificity of the Cas9/sgrRNA ribonucleoprotein complex. While the variable 5'–20nt 'guide' sequence of the sgrRNA confers target specificity, slight modifications of the 3' constant sequence (tracr) recognized by Cas9 have been shown to affect the detectable amounts of DNA-bound Cas9/sgrRNA complex (Chen, B. et al., Cell 155, 1479–91 2013). To investigate whether Cas9/sgrRNA gene-editing efficiency could be improved by using sgrRNAs with modified tracr structure, and whether such improvement could be beneficial for high-throughput genetic screening applications, we set up to compare the performance of multiple sgrRNA designs in a pooled library HTGS high-throughput genetic screening setting. An optimal sgrRNA tracr structure was identified, and subsequently used for the design of Cellecta's human and mouse genome-wide CRISPR tracrRNA libraries.

To investigate the potential differences in the RNAi and CRISPR HTGS technologies, we then performed parallel shRNA and viability screens using module-1 human genome-wide CRISPR sgRNA library and module-1 human genome-wide RNAi shRNA library, both modules targeting the same set of 6,300 genes with the same number of effectors (sgRNA or shRNA) for each gene. The parallel screens were aimed at the identification of essential genes in a pair of primary isogenic DDX1-pXV-derived CML cell lines. Both approaches showed significant but not complete overlap in the genes identified by each technology in each cell line, suggesting that the results should be interpreted with caution to avoid any major bias in the cell system. However, further analysis of the results showed that the small number of genes could be classified by either CRISPR or RNAi technology. By combining data from the two screening methodologies, a consistent number of viability genes and pathways could be identified and subsequently validated by independent cell based assays at high confirmation rate.

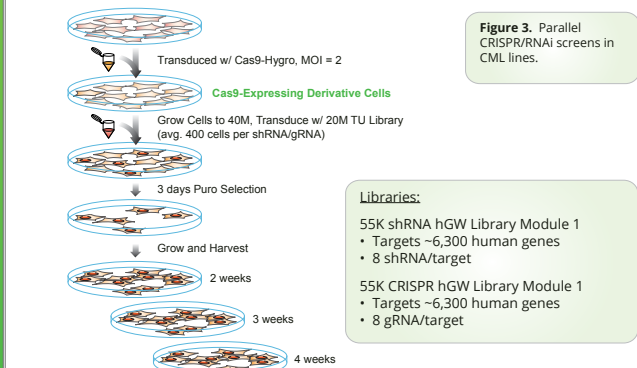


Figure 3. Parallel CRISPR/RNAi screens in CML lines.

Libraries:

- 55K shRNA hGW Library Module 1
- Targets ~6,300 human genes
 - 8 shRNA/target
- 55K CRISPR hGW Library Module 1
- Targets ~6,300 human genes
 - 8 gRNA/target

Figure 1 consists of two line graphs, (A) and (B), showing the performance of shRNA and gRNA over time (0, 2, 4, 6, 8 weeks).

(A) shRNA vs gRNA: depletions

This graph shows the log₂ fold change (y-axis, -1 to 2) for four conditions: gRNA negative (blue), gRNA positive (orange), shRNA negative (green), and shRNA positive (red). The x-axis represents time in weeks. The gRNA positive condition shows a significant decrease in log₂ fold change over time, reaching approximately -1.5 by 8 weeks. The other three conditions remain relatively stable, with log₂ fold change values between 0 and 1.

(B) shRNA vs gRNA: performance

This graph shows the log₂ fold change (y-axis, -1 to 1) for two conditions: gRNA negative (blue) and gRNA positive (orange). The x-axis represents time in weeks. The gRNA positive condition shows a significant decrease in log₂ fold change over time, reaching approximately -1.5 by 8 weeks. The gRNA negative condition remains relatively stable, with log₂ fold change values between 0 and 1.

Figure 4. 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.

Figure 2 consists of four bar charts arranged in a 2x2 grid, showing the distribution of gRNA/shRNA hits across different cell lines. The x-axis for all charts is 'hits/gene' (0 to 8), and the y-axis is 'Gene frequency'.

Top Left: gRNA/shRNA hit distribution (CMC-0 cell line)

This chart shows the distribution of hits for CMC-0 cells. The y-axis ranges from 0% to 100%. A red box labeled 'Off-Target' points to the bar at 0 hits/gene. A blue oval labeled 'True Hits' points to the bar at 4 hits/gene. The legend indicates blue bars for gRNA and orange bars for shRNA.

hits/gene	gRNA (%)	shRNA (%)
0	90	45
1	10	15
2	5	5
3	2	2
4	2	2
5	0	0
6	0	0
7	0	0
8	0	0

Top Right: gRNA/shRNA hit distribution (CMC-P cell line)

This chart shows the distribution of hits for CMC-P cells. The y-axis ranges from 0% to 100%. A red box labeled 'Off-Target' points to the bar at 0 hits/gene. A blue oval labeled 'True Hits' points to the bar at 4 hits/gene. The legend indicates blue bars for gRNA and orange bars for shRNA.

hits/gene	gRNA (%)	shRNA (%)
0	90	45
1	10	15
2	5	5
3	2	2
4	2	2
5	0	0
6	0	0
7	0	0
8	0	0

Bottom Left: gRNA/shRNA hit distribution (CMC-0 cell line)

This chart shows the distribution of hits for CMC-0 cells. The y-axis ranges from 0.0% to 1.6%. A blue oval labeled 'True Hits' points to the bar at 4 hits/gene. The legend indicates blue bars for gRNA and orange bars for shRNA.

hits/gene	gRNA (%)	shRNA (%)
4	1.0	1.4
5	0.6	0.6
6	0.5	0.4
7	0.2	0.2
8	0.1	0.2

Bottom Right: gRNA/shRNA hit distribution (CMC-P cell line)

This chart shows the distribution of hits for CMC-P cells. The y-axis ranges from 0.0% to 1.6%. A blue oval labeled 'True Hits' points to the bar at 4 hits/gene. The legend indicates blue bars for gRNA and orange bars for shRNA.

hits/gene	gRNA (%)	shRNA (%)
4	0.6	0.9
5	0.7	0.4
6	0.4	0.4
7	0.2	0.2
8	0.1	0.1

Figure 5. 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.

Condition	RNAi	CRISPR	RNAi/CRISPR overlap
CME-I	~90	~95	~40
CME-P	~85	~90	~35

Figure 6. 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: exclusive Gene Hits		CRISPR		RNAi	
		CML-R	CML-P	CML-R	CML-P
	ATF6B08	ATF6B08		MTAP	OTODG
	CENH	CHRD0C1		TAF1L	UBC
	DHFR	CHDS5		UBC	
	ERCC2	DHFR			
	FARS5	ERC2			
	FARS8	FARS5			
	NAPL1	NAPL1			
	PGAM1	PGAM1			
	RNMT				
	SCYL1	TRM112			

Figure 7. The number and lists of gene hits exclusively identified by either CRISPR or RNAi technology (no signal with other technology) are displayed.

[illegible]

Figure 8. 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.

Pathway	Tested genes	Protein	Inhibitory Drug	Drug Toxicity
Transferin endocytosis and recycling / iron uptake and transport	ATP6V0B ATP6V0C ATP6V1A ATP6V1D ATP6V1F	V-ATPase	Bafilomycin A1 Bafilomycin B1 Concanamycin A	+ + +
Signaling by Insulin receptor				
Super pathway of geranylgeranyldiphosphate biosynthesis I	FDPS HMGCS3	farnesyl Diphosphate Synthase 3-Hydroxy-3-Methylglutaryl-CoA Synthase 1	Zoledronic acid Aledronic acid Hyemequin	+/- +/- +
Methionine degradation I (to homocysteine)	AHCY AHCYL1	S-adenosylhomocysteine hydrolase Adenosylhomocysteinase 2	Neplanocin	+
Cysteine biosynthesis III (mammalia)	MAT2A	Methionine Adenosyltransferase II, Alpha	FIDAS-5	+

Figure 9. 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.

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

- Specific modifications of the sgRNA invariant tracr portion increase the performance of CRISPR technology in high-throughput genetic screens
- CRISPR technology appears to be less affected by time-dependent genetic drift than RNAi
- CRISPR technology appear to be less affected by off-target gRNA cytotoxicity, compared to off-target shRNA cytotoxicity for RNAi
- Both technologies were able to identify viability gene hits with excellent reproducibility in isogenic cell lines
- The good overlap between viability gene hits identified by the CRISPR and RNAi suggests that the parallel use of both technology could be extremely useful for hit confirmation
- A subset of viability genes identified by CRISPR technology were fully CRISPR-specific, as they showed no signal within the RNAi screens. This observation corroborates the prediction that for some genes, complete knockout is needed to show a viability phenotype.
- Pathway analysis of CRISPR/RNAi combined data was able to identify potential drug targets with high confirmation rate