HHS Public Access

Author manuscript

Cancer Discov. Author manuscript; available in PMC 2017 February 01.

Published in final edited form as:

Cancer Discov. 2016 August ; 6(8): 824–826. doi:10.1158/2159-8290.CD-16-0665.

Genomic amplifications cause false positives in CRISPR screens

Ankur Sheel[1] and **Wen Xue**[1,2,#]

[1] RNA Therapeutics Institute, University of Massachusetts Medical School, 368 Plantation Street, AS4.2043, Worcester, MA 01605

[2] Program in Molecular Medicine and Department of Molecular, Cell and Cancer Biology, University of Massachusetts Medical School, 368 Plantation Street, AS4.2043, Worcester, MA 01605. Phone: 774-455-3783 Fax: 508-856-6696

Summary

In CRISPR-based screens for essential genes, Aguirre et al. and Munoz et al. show that gene-independent targeting of genomic amplifications in human cancer cell lines reduces proliferation or survival. The correlation between CRISPR target site copy number and lethality demonstrates the need for scrutiny and complementary approaches to rule out off-target effects and false positives in CRISPR screens.

Reverse-genetic screens are powerful tools for decoding how genotype translates to phenotype (1). RNA interference (RNAi) and CRISPR are two major tools for loss-of-function, reverse-genetic studies in mammalian cells (2). High-throughput RNAi- and CRISPR-based screens are typically performed by transducing cells with a lentivirus library of short hairpin RNAs (shRNAs) or single-guide RNAs (sgRNAs), and the shRNAs or sgRNAs that produce a desired phenotype are identified in a population as enriched (e.g., inactivation promotes cell growth) or depleted (e.g., inactivation reduces viability). Recent studies reveal profound differences between genes identified by RNAi- and CRISPR-based screens even in the same genetic context (3). The disparate cellular responses to RNAi-based knockdown and CRISPR-based genome editing indicate the need for more scrutiny when interpreting the results of genetics screens using these technologies.

RNAi-based genetic screens have been widely used in mammalian cells, and false-positive and false-negative results of RNAi screens have been well characterized. Partial silencing of a target gene by RNAi may be insufficient to produce a phenotype, leading to false-negative results for some genes. Unintended or “off-target” silencing of mRNAs homologous to a target mRNA or quenching of the microRNA pathway by overproduction of an RNAi construct can lead to false-positive results.

The commonly used *S. pyogenes* CRISPR system targets the DNA rather than the transcript of a gene and introduces a double-strand DNA break. Imprecise repair of the break by non-homologous end joining results in a small insertion or deletion that disrupts the reading

[#] Correspondence: Wen.Xue@umassmed.edu.

The authors disclose no potential conflicts of interest.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

frame. Thus CRISPR inactivates target genes more readily than RNAi does. Nevertheless, potential false-positive and false-negative results of CRISPR screens should be expected. Indeed, complete gene knockout can trigger compensatory transcriptional activation of functionally redundant genes (4), causing false-negative results. As with RNAi, false-positives may result if CRISPR modifies off-target genomic sites with sequence similarity to the intended CRISPR target site (5). But the caveats of CRISPR-based screens have not been systematically documented. In this issue of *Cancer Discovery,* Aguirre et al. (6) and Munoz et al. (7) identify target site copy number as an unexpected trigger of false-positives in CRISPR screens for essential genes.

The technical features of the two CRISPR screens are summarized in **Fig. 1A**. Aguirre et al. used a genome-wide library at 6 sgRNA per gene. This study was performed in a large collection of 33 cell lines representing a variety of cancer types and genetic contexts. Munoz et al. used a focused CRISPR library targeting ~2700 genes in five cell lines. While not at genome-wide scale, focused CRISPR library can achieve more sgRNA per gene (20 vs. 6) and higher representation (1000 cell per sgRNA vs. 500) compared to a genome-wide library.

Aguirre et al. show that their CRISPR screening approach identifies essential genes in cancer cells, including both oncogene drivers and non-oncogene dependencies. Their study parallels previous studies in which they performed genome-wide RNAi screens using the same cell lines, allowing them to make some comparisons between the CRISPR and RNAi screens (8). Because copy number amplification frequently leads to oncogene overexpression, they sought to identify cancer drivers associated with genomic amplification. They unexpectedly found that CRISPR guides targeting genes within genomic amplifications reduce proliferation or survival as compared to guides that target genes outside of the amplifications (**Fig. 1B**).

Seeking to compare RNAi and CRISPR screening technologies, Munoz et al. constructed complementary shRNA and sgRNA libraries (**Fig. 1A**). Screening in five cancer cell lines—three diploid and two aneuploid—they found that CRISPR screens identify 2- to 5-times more essential genes than RNAi screens, possibly due to more complete inactivation by CRISPR. To rule out the possibility that the CRISPR screens had lower false-negative rates or higher false-positive rates, the authors examined the lethality scores of non-expressed genes. Notably, the CRISPR screens in the three diploid cell lines produced virtually no false positives. But the screens in the two aneuploid cell lines did produce false positives, and the false positives mapped to genomic amplifications.

Both studies proceed to more carefully examine the effects of CRISPR target site copy number on cell viability. Consistent with the finding that CRISPR-mediated lethality is independent of transcriptional status, CRISPR guides targeting intergenic sequences in genomic amplifications are as lethal as those targeting essential genes. Thus reduced proliferation or survival by sgRNAs targeting amplified genes is not due to gene inactivation. Comparing aggregate analysis of apparent essentiality due to amplified genes, Aguirre et al. find that increasingly essential genes (based on level of CRISPR guide depletion) were more likely to reside in genomic amplifications in CRISPR screens than

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

they were in RNAi screens—also consistent with the finding that lethality does not result from gene inactivation.

Both studies find that the anti-proliferative effect of CRISPR targeting genomic amplifications positively correlates with target site copy number. They show that CRISPR guides that target multiple sites throughout the genome induce lethality, and that the best predictor of off-target lethality is the number of genomic sites with perfect complementarity to an intended CRISPR target site. The authors of both studies therefore propose that excessive DNA damage due to CRISPR cutting underlies the lethality, and they show that CRISPR guides targeting genomic amplifications or multiple sites throughout the genome activates the DNA damage response (i.e., gamma-H2AX phosphorylation and foci) and increases G2/M cell cycle arrest. These findings help to explain observations of CRISPR-induced off-target lethality as well as previous work showing that sgRNAs targeting a non-genic region of the BCR-ABL amplification decrease cell viability (9).

Aguirre et al. suggest two cellular responses to CRISPR genome editing in cancer cells: an early anti-proliferative DNA damage response and a later target gene inactivation. Overall, the anti-proliferative effect is independent of the target gene or chromosome structure, and increases with the number of cuts conferred by individual sgRNAs. The authors propose that the early anti-proliferative cell response induced by CRISPR represents a critical vulnerability of cancer cells with genome amplifications that might enable cancer-specific therapy. In the second response, sgRNAs targeting essential genes are depleted following loss of protein expression, representing a true positive in a CRISPR screen. They note the caveat, however, that copy number amplifications may protect some essential genes from complete knockout by CRISPR.

The study by Munoz et al. also provides insight into the design and functionality of sgRNAs used in CRISPR screens. Screening three different cell lines with a CRISPR tiling array against 139 essential genes and ~364 sgRNAs per gene, they determined that the best predictor of sgRNA performance is targeting a conserved Pfam protein domain—consistent with a recent study by Shi et al. (10)—followed by sequence conservation across vertebrate species. In addition to these features, the authors advise the following criteria to avoid excessive double-strand DNA breaks: design sgRNAs with minimal matches across the genome, transduce lentiviral sgRNA libraries at low multiplicity of infection, and use of sgRNA targeting non-expressing or known non-essential regions as controls compared to scrambled sgRNA.

In summary, these papers report comprehensive loss-of-function CRISPR screens across a panel of human cancer cell lines. Both papers show that cancer driver genes can be readily identified, demonstrating the feasibility of CRISPR-based functional genomic screens. Both studies also highlight potential false-positive results in CRISPR screens in cancer cell lines harboring copy number variations, adding a layer of complexity to interpreting CRISPR screens. The findings call for improved CRISPR libraries, use of diverse cell lines, and scrutiny of CRISPR screen data. Alternative approaches—including loss-of-function RNAi knockdown, gain-of-function cDNA rescue, and CRISPR-based transcriptional inactivation

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

—are needed to screen for cancer drivers in genomic amplifications and should be used to complement or validate CRISPR screens.

Acknowledgments

This work was supported by NCI (5R00CA169512), American Cancer Society (129056-RSG-16-093) and the Lung Cancer Research Foundation to W.X.

References

1. Shalem O, Sanjana NE, Zhang F. High-throughput functional genomics using CRISPR-Cas9. Nat Rev Genet. 2015; 16(5):299–311. [PubMed: 25854182]
2. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014; 346(6213):1258096. [PubMed: 25430774]
3. Housden BE, Perrimon N. Comparing CRISPR and RNAi-based screening technologies. Nat Biotechnol. 2016; 34(6):621–3. [PubMed: 27281421]
4. Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, et al. Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature. 2015; 524(7564):230–3. [PubMed: 26168398]
5. Tsai SQ, Joung JK. Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. Nat Rev Genet. 2016; 17(5):300–12. [PubMed: 27087594]
6. Aguirre AJ, Meyers RM, Weir BA, Vazquez F, Zhang C-Z, Ben-David U, et al. Genomic copy number dictates a gene-independent cell response to CRISPR-Cas9 targeting. Cancer Discov. 2016
7. Munoz DM, Cassiani PJ, Li L, Billy E, Korn JM, Jones MD, et al. CRISPR screens provide a comprehensive assessment of cancer vulnerabilities but generate false-positive hits for highly amplified genomic regions. Cancer Discov. 2016
8. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483(7391):603–7. [PubMed: 22460905]
9. Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, et al. Identification and characterization of essential genes in the human genome. Science. 2015; 350(6264):1096–101. [PubMed: 26472758]
10. Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB, Vakoc CR. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nat Biotechnol. 2015; 33:661–7. [PubMed: 25961408]

Author Manuscript

A.

	Aguirre et al.	Munoz et al.
No. cell lines	33 (K562, PANC1, T47D, etc)	5 (DLD1, RKO, HT-1080, etc)
No. genes	19,050	2,722
sgRNA per gene	6	20
Cell per sgRNA	~500	≥1000

B.

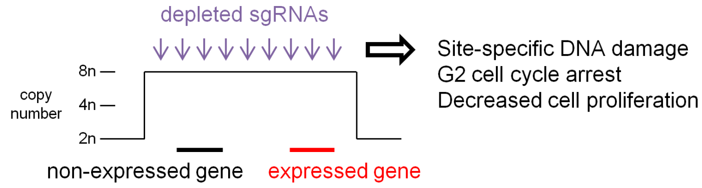

Fig. 1. Genomic copy number causes false positive hits in CRISPR screens

(A) Technical features of the two CRISPR-based screens for essential genes in human cancer cell lines. "Cell per sgRNA" denotes the number of cells per sgRNA, or the representation of library maintained at each cell passage. (B) Both papers find that sgRNAs are depleted if they target highly amplified genomic regions, independent of gene expression status.