

# CRISP-view: a database of functional genetic screens spanning multiple phenotypes

Yingbo Cui<sup>1,†</sup>, Xiaolong Cheng<sup>2,†</sup>, Qing Chen<sup>2</sup>, Bicna Song<sup>2</sup>, Anthony Chiu<sup>2,3</sup>, Yuan Gao<sup>2,4</sup>, Tyson Dawson<sup>2,5,6</sup>, Lumen Chao<sup>2</sup>, Wubing Zhang<sup>7</sup>, Dian Li<sup>7</sup>, Zexiang Zeng<sup>7</sup>, Jijun Yu<sup>8</sup>, Zexu Li<sup>9</sup>, Teng Fei<sup>9</sup>, Shaoliang Peng<sup>10</sup> and Wei Li<sup>2,11,\*</sup>

<sup>1</sup>Sanyi Road, Changsha, Hunan Province, People's Republic of China, <sup>2</sup>Center for Genetic Medicine Research, Children's National Hospital. 111 Michigan Ave NW, Washington, DC 20010, USA, <sup>3</sup>School of Medicine and Health Sciences, George Washington University, 2300 I Street NW, Washington, DC 20037, USA, <sup>4</sup>Department of Biochemistry and Molecular Biology, George Washington University, 2300 I Street NW, Washington, DC 20037, USA, <sup>5</sup>Institute for Biomedical Sciences, George Washington University, 2300 I Street NW, Washington, DC 20037, USA, <sup>6</sup>Computational Biology Institute, Milken Institute School of Public Health, George Washington University, 45085 University Drive, Ashburn, VA 20148, USA, <sup>7</sup>Department of Data Sciences, Dana-Farber Cancer Institute and Harvard T.H. Chan School of Public Health. 450 Brookline Ave., Boston MA 02215, USA, <sup>8</sup>Beijing Key Laboratory of Therapeutic Gene Engineering Antibody. Beijing, People's Republic of China, <sup>9</sup>College of Life and Health Sciences, Northeastern University. 110819 Shenyang, People's Republic of China, <sup>10</sup>Lushan South Road, Changsha, Hunan Province, People's Republic of China and <sup>11</sup>Department of Genomics and Precision Medicine, George Washington University. 111 Michigan Ave NW, Washington, DC 20010, USA

Received August 12, 2020; Revised September 12, 2020; Editorial Decision September 15, 2020; Accepted September 30, 2020

## ABSTRACT

**High-throughput genetic screening based on CRISPR/Cas9 or RNA-interference (RNAi) enables the exploration of genes associated with the phenotype of interest on a large scale. The rapid accumulation of public available genetic screening data provides a wealth of knowledge about genotype-to-phenotype relationships and a valuable resource for the systematic analysis of gene functions. Here we present CRISP-view, a comprehensive database of CRISPR/Cas9 and RNAi screening datasets that span multiple phenotypes, including *in vitro* and *in vivo* cell proliferation and viability, response to cancer immunotherapy, virus response, protein expression, etc. By 22 September 2020, CRISP-view has collected 10 321 human samples and 825 mouse samples from 167 papers. All the datasets have been curated, annotated, and processed by a standard MAGeCK-VISPR analysis pipeline with quality control (QC) metrics. We also developed a user-friendly webserver to visualize, explore, and search these datasets. The webserver is freely available at <http://crispview.weillab.org>.**

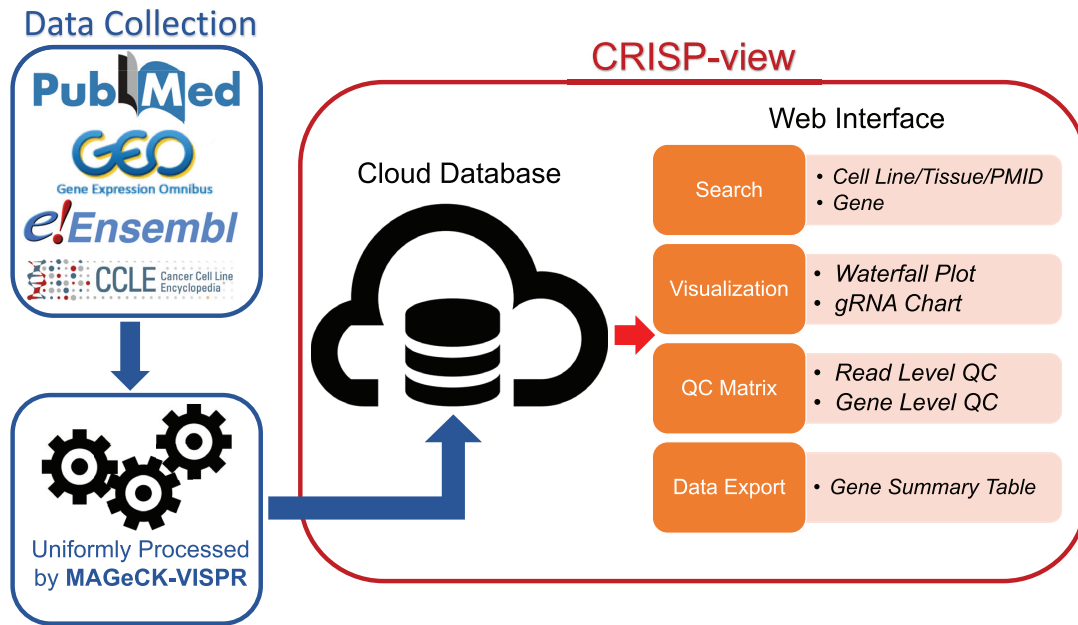
## INTRODUCTION

Functional genetic screening is a high-throughput, cost-effective technology to identify genes or genomic elements that are pertinent to a phenotype-of-interest (1–7). Based on RNA-interference (RNAi) or CRISPR/Cas9 (8–11), screening explores the functions of genes (or non-coding elements) in various contexts including cancer progression, interaction with immune system, response to drug treatment, virus infection etc. (12–17). The broad spectrum of possible phenotypes that can be covered by genetic screening provides a wealth of information on our understanding of the genes, non-coding elements, and their associated pathways in different aspects (18–25).

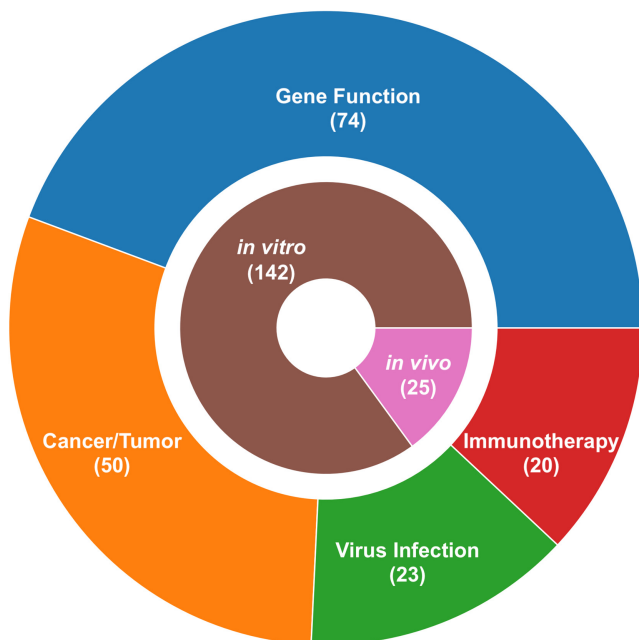
With the rapid accumulation of genetic screens in recent years, several databases are developed to collect, visualize and compare these datasets. GenomeCRISPR (26) collects human cell-line CRISPR/Cas9 screening datasets and enables users to explore the behavior of genes and sgRNAs. PICKLES (27) enables pooled *in vitro* knock-out CRISPR screening data to be visualized together with other types of data including copy number variation, expression and mutation. BioGRID ORCS (28) reports scores of CRISPR/Cas9 screening, generated by different analysis algorithms, from the original publication. In addition, large-scale genome-wide screening projects, including Project Drive (29), DepMap (30) and Sanger DepMap (or Project Score) (31), provided a centralized web interface for

\*To whom correspondence should be addressed. Tel: +1 202 476 4986; Email: wli2@childrensnational.org

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.



**Figure 1.** A Schematic view of CRISP-view. The CRISP-view database consists of three parts: a curated metadata collection, a unified screening data processing and a web interface.



**Figure 2.** Statistics of datasets collected in CRISP-view so far. The dataset includes *in vitro* and *in vivo* experiments and covers a variety of research areas, including gene function, cancer/tumor, Immunotherapy and virus infection.

users to browse *in vitro* screening data as well as the associated genomic profiles from hundreds of cell lines. However, several limitations restrict the wide application of these datasets. For example, all these databases collected screens that mainly measure the *in vitro* proliferation (or viability) of cancer cell lines, while screens of other phenotypes are lacking. Among those, DepMap and Sanger DepMap

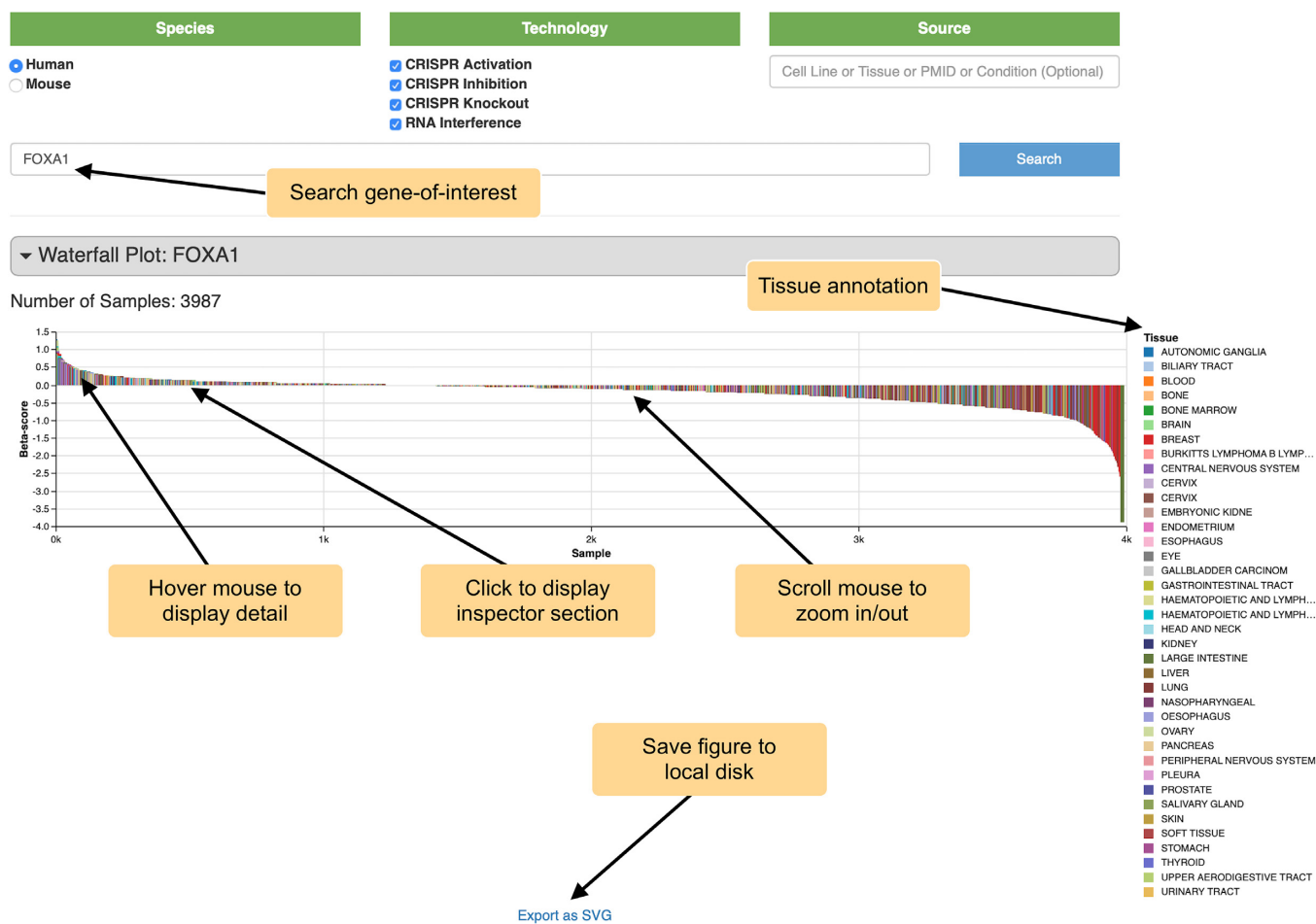
only included in-house screening datasets on cancer cell lines. Other datasets only stored published analysis results of screens reported in the paper (like BioGRID ORCS, GenomeCRISPR), or do not include up-to-date screening datasets. As the analysis results are generated by different methods from different studies, and the quality of the datasets may vary, it is difficult to systematically compare screens across different datasets. Therefore, a central challenge is to store, process, and evaluate genetic screening data spanning multiple phenotypes and in a standardized way, where users are able to gain new biological insights through data mining.

Here we present CRISP-view, a database of CRISPR- or RNAi-based genetic screening spanning various phenotypes, including *in vitro* and *in vivo* cell proliferation or viability, immune or immunotherapy response, virus infection, protein expression (by GFP sorting), etc. As of 22 September 2020, CRISP-view collected 11,146 samples with unified metadata annotation, curation and standardized quality control metrics. CRISP-view also provides a web interface for users to search and browse all the datasets (and their associated metadata). CRISP-view represents the most comprehensive collection of screening datasets up to date, and is constantly being updated as screening datasets accumulates in the public domain.

## DATABASE CONTENT

CRISP-view is a comprehensive annotated resource of public genetic screening data in human and mouse. CRISP-view contains >11 000 genetic *in vitro* and *in vivo* screening samples from 167 (and growing) publications. A variety of different screening technologies are covered, including CRISPR activation (CRISPRa), CRISPR inhibition (CRISPRi), CRISPR knockout and RNA interference

## CRISP-view: Database of CRISPR Screen Data



**Figure 3.** Searching genes of interest in CRISP-view. FOXA1 is shown as an example and a waterfall plot is shown below. Different colours in the waterfall plot represent different tissues of origin of cell line.

(RNAi). The CRISP-view database consists of three parts: a curated metadata collection, unified screening data processing and a web interface. All data sources and web interface features are summarized in Figure 1.

### DATA SOURCES AND METADATA ANNOTATION

CRISP-view collects publicly available CRISPR/Cas9 and RNAi screening data from gene expression omnibus (GEO), the supplemental materials of papers or from the author directly. We systematically annotated the metadata of these samples manually, including species, screening type, cell line, library, screening conditions, associated PubMed ID and citation.

In total, our database contains 11 146 samples from 167 high-throughput genetic experiments, including 10 321 (and 825) human (and mouse) samples, respectively. Included in the database are 1271 human and 29 mouse cell line screens, respectively. There are 81 different screening libraries used in these high-throughput experiments. The database also covers a very wide range of research subjects, including proliferation and viability in cancer cells

and mouse models, immune and immunotherapy related screens, virus related screens, and the expression of certain gene marker (Figure 2 and Supplementary Figure S1). We list all the publications incorporated in CRISP-view database in the Statistics Page (<http://crispview.weillab.org/statistics>).

### DATA PROCESSING AND QUALITY CONTROL METRICS

To ensure the consistency between different datasets, we limit our screening datasets to those where raw sequence data (fastq format) or raw count tables are available. The data is uniformly processed by the MAGeCK-VISPR pipeline (32,33) that generates QC measurements and beta scores for all the perturbed genes. Beta-score is a measurement to reflect the functions of genes in the screen, similar to the term 'log fold change' in differential expression analysis: a positive (or negative) beta score indicates a positive (or negative) selection of the corresponding gene in the screen, respectively.

## CRISP-view: Database of CRISPR Screen Data

**Species**

Human  
 Mouse

**Technology**

CRISPR Activation  
 CRISPR Inhibition  
 CRISPR Knockout  
 RNA Interference

**Source**

K-562

Gene Symbol (Optional)

Search cell line, tissue, PMID or treatment condition

▼ Results

Click search result to display inspector section

Show:  Search:

Sample ▲	Species	Cell Line	Type	Publication	QC
Anti_A	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●
Anti_B	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●
Anti_C	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●
AO_10_10_A	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●
AO_10_10_B	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●
AO_10_10_C	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●
Chlor_A	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●
Chlor_B	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●
Chlor_C	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●
DO_A	human	K-562	CRISPR ko	To T, et al. Cell 2019	● ● ● ● ● ● ● ●

Showing 1 to 10 of 97 records

Pages: Previous **1** 2 3 ... 10 Next

**Figure 4.** Searching in CRISP-view. ‘K-562’ is used as an example of search term and CRISP-view returns all matched samples. The corresponding sample information and QC metrics are listed in the table.

▼ Inspector

Metadata annotation   Positively selected genes

Metadata   QC   Positive Selection   Negative Selection

QC report   Negatively selected genes

Sample	Anti_A
Species	Human
Technology	CRISPR ko
Cell Line	K-562
Tissue	Haematopoietic And Lymphoid Tissue
Source	Count table
Treatment	100 nM Antimycin
Other Condition	
Culture Days	15
Normalized Method	Brunello Controls
Library	<a href="#">Brunello</a>
Citation	To T, et al. A Compendium of Genetic Modifiers of Mitochondrial Dysfunction Reveals Intra-organelle Bufferin. Cell 2019
PMID	<a href="#">31730859</a>

**Figure 5.** Metadata Inspector view of individual sample. The inspector section displays the detailed data annotations, quality control report and positively/negatively selected genes of individual sample in separated tabs.



Figure 6. Positive selection Inspector view of individual sample.

The QC measurement, generated by MAGeCK-VISPR, includes the number of reads and the percentage of mapped reads, the number of sgRNAs with zero read count, and the Gini index of read count distribution. In addition, we evaluate the degree of negative selection on ribosomal genes using GSEA (34) as a measurement of quality in proliferation-based dropout screens, because the knockout of ribosomal genes is expected to have a strong negative selection phenotype (1,35). We also define a threshold for each QC metric for pass or fail, based on the distribution of that metric in all samples. A QC metric is considered as pass if it is better than 2/3 of the samples in the database.

## DATABASE INTERFACE AND TUTORIAL

The CRISP-view website is available at <http://crispview.weililab.org>. The main page provides options for users to select species and screening technologies, and to search datasets by gene symbol, publication ID (PMID), treatment conditions or biological source (cell line name or tissue type).

User can explore the beta scores of interesting gene by searching gene symbols, and view detailed data annotations, quality control metric and positively/negatively selected genes of individual sample. After entering a gene symbol, a waterfall plot will display the ranked beta-score of that gene in all samples (Figure 3). Different colors in waterfall plot represent different tissues of origin for cell lines. The waterfall plot allows users to zoom-in, zoom-out or move within the image for exploration. When mouse hovers over the waterfall plot, a smaller view window will show the basic information of the selected sample (e.g. cell line name, tissue of origin, screening library, treatment condition, the beta-score, and source). After clicking on any data point in waterfall plot, users can navigate the inspector section (behind the waterfall plot) to inspect the corresponding sample in more details. The waterfall plot can be saved locally in SVG format.

User can also search interesting cell line, tissue, publication or treatment condition, where matched samples will be shown in a table (Figure 4). User can view the detailed information of each sample in the inspector section (Figure 5). In the inspector section, CRISP-view provides three layers

of content for each sample in four tabs: a manually curated metadata annotation (first tab), QC results (second tab) and a list of positively and negatively selected genes and associated sgRNAs (third and fourth tabs). Metadata annotation includes sample name, species, screening technology, cell line, tissue of origin, source data format, normalized method, screening library and citation. In the QC results, the quality control report is shown briefly by colored circles: green (and red) indicates the metric passed (or failed) the threshold, respectively. A list of positively (and negatively) selected genes and associated sgRNAs are shown in the third (and fourth) tabs, respectively, where genes are ranked according to their beta-scores (Figure 6). User can view normalized read counts of gRNAs in selected sample and its corresponding initial condition(s), and download gene selection information as tab-separated text file.

## IMPLEMENTATION

The client of the database is implemented in jQuery and Twitter Bootstraps, and the server is implemented with Apache in AWS EC2. The communication between servers and clients uses JSON format (by Ajax), which efficiently reduces response time and improves user experience. In order to build a robust and scalable database, all the datasets are stored in MySQL database with AWS Rational Database Service (RDS) that contains thousands of samples and millions of data points (gene beta scores and sgRNA normalized counts). We used several strategies to speed up the database. First, we use the Ajax web model that allows the server to send or receive data asynchronously in the background, without interfering with the display of the client. Second, the client only loads the top 100 positively or negatively selected genes by default, and will load additional genes per user's request. Third, we constructed a multi-column index for the MySQL database to enable quick search over values of different columns. This strategy efficiently reduces the amount of data transferring to the client by default and improves the speed of inquiry.

## DISCUSSION AND FURTHER DIRECTIONS

We present CRISP-view, a comprehensive annotated database of RNAi and CRISPR/Cas9 genetic screens in human and mouse samples. All the collected raw data is processed by a uniform MAGeCK-VISPR pipeline with complete metadata annotation and QC measurements. CRISP-view provides a web interface to visualize the datasets and associated metadata, and enables researchers to explore interesting genes, cell lines, tissues, studies or conditions. The CRISP-view database is updated on a regular basis to incorporate newly published genetic screening data, providing a powerful resource for researchers to explore gene functions associated with different phenotypes.

Future works of CRISP-view include data mining and machine learning approaches to better understand the gene functions in different datasets or conditions, and more visualization and analysis tools to associate screening data with other types of data (like expression, mutation, copy number variation, and epigenetic profiles).

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors would like to thank Jingyu Peng, Shengqing Gu, Feizheng Wu for the helpful discussion about the project, and Johannes Köster for the help of developing the visualization module and database.

## FUNDING

W.L., X.C., B.S., Q.C. and L.C are supported by the research starter award from Pharmaceutical Research and Manufacturers of American Foundation (PhRMA) and the startup fund from the Center of Genetic Medicine Research at Children's National Hospital; A.C. is supported by the W.T. Gill Fellowship program at George Washington University. Funding for open access charge: Wei Li's startup fund at Children's National Hospital; PhRMA fund from Wei Li.

*Conflict of interest statement.* None declared.

## REFERENCES

1. Wang, T., Wei, J.J., Sabatini, D.M. and Lander, E.S. (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, **343**, 80–84.
2. Chen, S., Sanjana, N.E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D.A., Song, J., Pan, J.Q. and Weissleder, R. (2015) Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell*, **160**, 1246–1260.
3. Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.J. and Joung, J.K. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.*, **31**, 227–229.
4. Zhou, Y., Zhu, S., Cai, C., Yuan, P., Li, C., Huang, Y. and Wei, W. (2014) High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature*, **509**, 487–491.
5. Koike-Yusa, H., Li, Y., Tan, E.P., Velasco-Herrera Mdel, C. and Yusa, K. (2014) Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.*, **32**, 267–273.
6. Wang, T., Wei, J.J., Sabatini, D.M. and Lander, E.S. (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, **343**, 80–84.
7. Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G. *et al.* (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, **343**, 84–87.
8. Carpenter, A.E. and Sabatini, D.M. (2004) Systematic genome-wide screens of gene function. *Nat. Rev. Genet.*, **5**, 11–22.
9. Hannon, G.J. (2002) RNA interference. *Nature*, **418**, 244–251.
10. Echeverri, C.J., Beachy, P.A., Baum, B., Boutros, M., Buchholz, F., Chanda, S.K., Downward, J., Ellenberg, J., Fraser, A.G. and Hacohen, N. (2006) Minimizing the risk of reporting false positives in large-scale RNAi screens. *Nat. Methods*, **3**, 777–779.
11. Echeverri, C.J. and Perrimon, N. (2006) High-throughput RNAi screening in cultured cells: a user's guide. *Nat. Rev. Genet.*, **7**, 373–384.
12. Han, J., Perez, J.T., Chen, C., Li, Y., Benitez, A., Kandasamy, M., Lee, Y., Andrade, J. and Manicassamy, B. (2018) Genome-wide CRISPR/Cas9 screen identifies host factors essential for influenza virus replication. *Cell Rep.*, **23**, 596–607.
13. Shi, J., Wang, E., Milazzo, J.P., Wang, Z., Kinney, J.B. and Vakoc, C.R. (2015) Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.*, **33**, 661–667.
14. Manguso, R.T., Pope, H.W., Zimmer, M.D., Brown, F.D., Yates, K.B., Miller, B.C., Collins, N.B., Bi, K., LaFleur, M.W. and Juneja, V.R.

- (2017) In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. *Nature*, **547**, 413–418.
15. Chen, S., Sanjana, N.E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D.A., Song, J., Pan, J.Q., Weissleder, R. *et al.* (2015) Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell*, **160**, 1246–1260.
  16. Xiao, T., Li, W., Wang, X., Xu, H., Yang, J., Wu, Q., Huang, Y., Geradts, J., Jiang, P., Fei, T. *et al.* (2018) Estrogen-regulated feedback loop limits the efficacy of estrogen receptor-targeted breast cancer therapy. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 7869–7878.
  17. Parnas, O., Jovanovic, M., Eisenhaure, T.M., Herbst, R.H., Dixit, A., Ye, C.J., Przybylski, D., Platt, R.J., Tirosh, I., Sanjana, N.E. *et al.* (2015) A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. *Cell*, **162**, 675–686.
  18. Rauscher, B., Valentini, E., Hardeland, U. and Boutros, M. (2017) Phenotype databases for genetic screens in human cells. *J. Biotechnol.*, **261**, 63–69.
  19. Haribowo, A.G., Hannich, J.T., Michel, A.H., Megyeri, M., Schuldiner, M., Kornmann, B. and Riezman, H. (2019) Cytotoxicity of 1-deoxysphingolipid unraveled by genome-wide genetic screens and lipidomics in *Saccharomyces cerevisiae*. *Mol. Biol. Cell*, **30**, 2814–2826.
  20. Loregger, A., Raaben, M., Nieuwenhuis, J., Tan, J.M., Jae, L.T., van den Hengel, L.G., Hendrix, S., van den Berg, M., Scheij, S. and Song, J.-Y. (2020) Haploid genetic screens identify SPRING/C12ORF49 as a determinant of SREBP signaling and cholesterol metabolism. *Nat. Commun.*, **11**, 1128.
  21. Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P. *et al.* (2015) BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature*, **527**, 192–197.
  22. Fei, T., Li, W., Peng, J., Xiao, T., Chen, C.H., Wu, A., Huang, J., Zang, C., Liu, X.S. and Brown, M. (2019) Deciphering essential cistromes using genome-wide CRISPR screens. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 25186–25195.
  23. Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M., Cleary, B., Lander, E.S. and Engreitz, J.M. (2016) Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science*, **354**, 769–773.
  24. Sanjana, N.E., Wright, J., Zheng, K., Shalem, O., Fontanillas, P., Joung, J., Cheng, C., Regev, A. and Zhang, F. (2016) High-resolution interrogation of functional elements in the noncoding genome. *Science*, **353**, 1545–1549.
  25. Zhu, S., Li, W., Liu, J., Chen, C.H., Liao, Q., Xu, P., Xu, H., Xiao, T., Cao, Z., Peng, J. *et al.* (2016) Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.*, **34**, 1279–1286.
  26. Rauscher, B., Heigwer, F., Breinig, M., Winter, J. and Boutros, M. (2016) GenomeCRISPR—a database for high-throughput CRISPR/Cas9 screens. *Nucleic Acids Res.*, **45**, D679–D686.
  27. Lenoir, W.F., Lim, T.L. and Hart, T. (2018) PICKLES: the database of pooled in-vitro CRISPR knockout library essentiality screens. *Nucleic Acids Res.*, **46**, D776–D780.
  28. Oughtred, R., Stark, C., Breitkreutz, B.-J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G. and McAdam, R. (2019) The BioGRID interaction database: 2019 update. *Nucleic Acids Res.*, **47**, D529–D541.
  29. McDonald, E.R. 3rd, de Weck, A., Schlabach, M.R., Billy, E., Mavrakis, K.J., Hoffman, G.R., Belur, D., Castelletti, D., Frias, E., Gampa, K. *et al.* (2017) Project DRIVE: a compendium of cancer dependencies and synthetic lethal relationships uncovered by large-scale, deep RNAi screening. *Cell*, **170**, 577–592.
  30. Meyers, R.M., Bryan, J.G., McFarland, J.M., Weir, B.A., Sizemore, A.E., Xu, H., Dharia, N.V., Montgomery, P.G., Cowley, G.S., Pantel, S. *et al.* (2017) Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.*, **49**, 1779–1784.
  31. Behan, F.M., Iorio, F., Picco, G., Goncalves, E., Beaver, C.M., Migliardi, G., Santos, R., Rao, Y., Sassi, F., Pinnelli, M. *et al.* (2019) Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature*, **568**, 511–516.
  32. Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M. and Liu, X.S. (2014) MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.*, **15**, 554.
  33. Li, W., Köster, J., Xu, H., Chen, C.-H., Xiao, T., Liu, J.S., Brown, M. and Liu, X.S. (2015) Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol.*, **16**, 281.
  34. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R. and Lander, E.S. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 15545–15550.
  35. Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E. and Doench, J.G. (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, **343**, 84–87.