bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Transcriptome-wide Cas13 guide RNA design for model organisms and viral RNA pathogens

Xinyi Guo[1,2], Hans-Hermann Wessels[1,2], Alejandro Méndez-Mancilla[1,2], Daniel Haro[1,2], Neville E. Sanjana[1,2,*]

1. New York Genome Center, New York, NY, USA.
2. Department of Biology, New York University, New York, NY, USA.

* E-mail: neville@sanjanalab.org

Keywords: Cas13, CRISPR, model organisms, RNA viruses, on-target efficiency prediction

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

CRISPR-Cas13 mediates robust transcript knockdown in human cells through direct RNA targeting. Compared to DNA-targeting CRISPR enzymes like Cas9, RNA targeting by Cas13 is transcript- and strand-specific: It can distinguish and specifically knock-down processed transcripts, alternatively spliced isoforms and overlapping genes, all of which frequently serve different functions. Previously, we identified optimal design rules for *Rfx*Cas13d guide RNAs (gRNAs), and developed a computational model to predict gRNA efficacy for all human protein-coding genes. However, there is a growing interest to target other types of transcripts, such as noncoding RNAs (ncRNAs) or viral RNAs, and to target transcripts in other commonly-used organisms. Here, we predicted relative Cas13-driven knock-down for gRNAs targeting messenger RNAs and ncRNAs in six model organisms (human, mouse, zebrafish, fly, nematode and flowering plants) and four abundant RNA virus families (SARS-CoV-2, HIV-1, H1N1 influenza and MERS). To allow for more flexible gRNA efficacy prediction, we also developed a web-based application to predict optimal gRNAs for any RNA target entered by the user. Given the lack of Cas13 guide design tools, we anticipate this resource will facilitate CRISPR-Cas13 RNA targeting in common model organisms, emerging viral threats to human health, and novel RNA targets.

* * *

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

CRISPR-Cas13 mediates robust transcript knockdown in human cells through direct RNA targeting[1-4]. Compared to DNA-targeting CRISPR enzymes like Cas9, RNA targeting by Cas13 is transcript- and strand-specific: It can distinguish and specifically knock-down processed transcripts, alternatively spliced isoforms and overlapping genes, all of which frequently serve different functions. Previously, we have described a set of optimal design rules for *Rfx*Cas13d guide RNAs (gRNAs), and developed a computational model to predict gRNA efficacy for all human protein-coding genes[5]. However, there is a growing interest to target other types of transcripts, such as noncoding RNAs (ncRNAs)[6,7] or viral RNAs[8,9], and to target transcripts in other commonly-used organisms[10-13]. Here, we predicted relative Cas13-driven knock-down for gRNAs targeting messenger RNAs and ncRNAs in six model organisms (human, mouse, zebrafish, fly, nematode and flowering plants) and four abundant RNA virus families (SARS-CoV-2, HIV-1, H1N1 influenza and MERS). To allow for more flexible gRNA efficacy prediction, we also developed a web-based application to predict optimal Cas13d guide RNAs for any RNA target entered by the user.

To select optimal gRNAs for transcripts produced from the reference genomes of human, mouse, zebrafish, fly, nematode and flowering plants, we created a user-friendly Cas13 online platform (https://cas13design.nygenome.org/) (**Fig. 1a**). We previously found that optimal Cas13 gRNAs depend on specific sequence and structural features, including position-based nucleotide preferences in the gRNA and the predicted folding energy (secondary structure) of the combined direct repeat plus gRNA[5]. Using this algorithm, we pre-computed gRNA efficacies, where possible, for all mRNAs and ncRNAs with varying transcript length for the 6 model organisms (**Fig. 1b, Supplementary Fig. 1**).

For the scored gRNAs for each organism, we found that approximately 20% are ranked in the top quartile (Q4 guides) for both mRNAs and ncRNAs (**Fig. 1b**). Remarkably, even though the nucleotide composition can very between RNAs from different species[14-16], we find a similar proportion of optimal *Rfx*Cas13d gRNAs across all six species.

Next, we examined how many predicted high efficacy gRNAs are present, on average, in different transcripts. To do this, we determined what fraction of the transcripts in each organism include *n* top-scoring

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Q4) gRNAs for values of *n* between 1 and 25. We found that coding sequences contained a higher number of top-scoring gRNA per transcript across all organisms, whereas targeting the noncoding transcriptome is more challenging and varies across different organisms (**Fig. 1c**). On average, we were able to find at least 25 Q4 gRNAs for >99% of coding exons in mRNAs but only 80% of ncRNAs. Beyond targeting transcripts from the reference genomes of these model organisms, there are also many other applications of Cas13, such as targeting transcripts from non-model organisms, cleavage of synthetic RNAs, and targeting of transcripts carrying genetic variants not found in the reference genome. Therefore, in addition to these pre-scored gRNAs, we have also developed a graphical interface that allows the user to input a custom RNA sequence for scoring and selection of optimal gRNAs.

Recently, several groups have proposed using CRISPR-Cas13 nucleases to directly target viral RNAs[8,17], which has become an area of rapid technology development due to the recent COVID-19 pandemic[18]. However, these approaches do not use optimized Cas13 guide RNAs. Previously, we showed that optimal guide RNAs targeting an EGFP transgene can result in a ~10-fold increase in knock-down efficacy compared to other gRNAs[5]. Therefore, to speed development of effective CRISPR-based antiviral therapeutics, we applied our design algorithm to target SARS-CoV-2 and other serious viral threats using Cas13d.

To ensure coverage of diverse patient isolates, we collected 7,630 sequenced SARS-CoV-2 genomes submitted to the Global Initiative on Sharing All Influenza Data (GISAID) database from 58 countries/regions[19] (**Fig. 2a**). Using the first sequenced SARS-CoV-2 isolate from New York City (USA/NY1-PV08001/2020) as a reference[20], we evaluated how many individual SARS-CoV-2 genomes each reference gRNA can target (**Fig. 2b**). Guide RNAs targeting protein-coding regions are mostly well-conserved across all genomes, with lower conservation in more variable regions such as Non-Structural-Protein 14 (NSP14) and Spike (S) protein. We found that gRNAs targeting in the 5' and 3' untranslated regions tended to be poorly conserved, as might be expected given the lack of coding function of these regions (**Supplementary Fig. 2**). Upon examination of each of the 26 SARS-CoV-2 genes, we found that all gene transcripts could be targeted with Q4 gRNAs.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Similarly, we designed and scored all gRNAs for the coronavirus MERS and two other RNA viruses, HIV-1 which drives Acquired Immunodeficiency Syndrome (AIDS) and H1N1 pandemic influenza. Unlike SARS-CoV-2, where a single high-efficacy (Q4) gRNA can target all genomes analyzed, we found that at least two gRNAs are needed to target nearly all available genomes. For the highly mutagenic virus HIV-1[21], we found that nine gRNAs are needed to target all available genomes (**Fig. 2c**). Given the tremendous interest in viral RNA targeting using Cas13 enzymes, this dataset of optimized gRNAs provides a platform for the development of CRISPR therapeutics for broad targeting of viral populations from diverse patient isolates. All designed gRNAs for model organism and viral transcripts can be interactively browsed or downloaded in bulk on the design tool website.

RNA-targeting CRISPR-Cas13 has great potential for transcriptome perturbation and antiviral therapeutics. In this study, we have designed and scored Cas13d gRNAs for both mRNAs and ncRNAs in six common model organisms and identified optimized gRNAs to target virtually all sequenced viral RNAs for SARS-CoV-2, HIV-1, H1N1 influenza and MERS. We further expanded our web-based platform to make the Cas13 gRNA design readily accessible for model organisms and created a new application to enable gRNA prediction for user-provided target RNA sequences. Given the current lack of Cas13 guide design tools, we anticipate this resource will greatly facilitate CRISPR-Cas13 RNA targeting in model organisms, emerging viral threats to human health and novel RNA targets.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., Verdine, V., Cox, D. B. T., Kellner, M. J., Regev, A., Lander, E. S., Voytas, D. F., Ting, A. Y. & Zhang, F. RNA targeting with CRISPR-Cas13. *Nature* 550, 280-284, (2017).

2. Smargon, A. A., Cox, D. B. T., Pyzocha, N. K., Zheng, K., Slaymaker, I. M., Gootenberg, J. S., Abudayyeh, O. A., Essletzbichler, P., Shmakov, S., Makarova, K. S., Koonin, E. V. & Zhang, F. Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol Cell* 65, 618-630 e617, (2017).

3. Konermann, S., Lotfy, P., Brideau, N. J., Oki, J., Shokhirev, M. N. & Hsu, P. D. Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell* 173, 665-676 e614, (2018).

4. Yan, W. X., Chong, S., Zhang, H., Makarova, K. S., Koonin, E. V., Cheng, D. R. & Scott, D. A. Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein. *Mol Cell* 70, 327-339 e325, (2018).

5. Wessels, H.-H., Méndez-Mancilla, A., Guo, X., Legut, M., Daniloski, Z. & Sanjana, N. E. Massively parallel Cas13 screens reveal principles for guide RNA design. *Nature Biotechnology* 38, 722-727, (2020).

6. Li, S., Li, X., Xue, W., Zhang, L., Cao, S.-M., Lei, Y.-N., Yang, L.-Z., Guo, S.-K., Zhang, J.-L., Gao, X., Wei, J., Li, J., Yang, L. & Chen, L.-L., (2020).

7. Xu, D., Cai, Y., Tang, L., Han, X., Gao, F., Cao, H., Qi, F. & Kapranov, P. A CRISPR/Cas13-based approach demonstrates biological relevance of vlinc class of long non-coding RNAs in anticancer drug response. *Sci Rep* 10, 1794, (2020).

8. Abbott, T. R., Dhamdhere, G., Liu, Y., Lin, X., Goudy, L., Zeng, L., Chemparathy, A., Chmura, S., Heaton, N. S., Debs, R., Pande, T., Endy, D., La Russa, M. F., Lewis, D. B. & Qi, L. S. Development of CRISPR as an Antiviral Strategy to Combat SARS-CoV-2 and Influenza. *Cell* 181, 865-876 e812, (2020).

9. Cui, J., Techakriengkrai, N., Nedumpun, T. & Suradhat, S. Abrogation of PRRSV infectivity by CRISPR-Cas13b-mediated viral RNA cleavage in mammalian cells. *Sci Rep* 10, 9617, (2020).

10. Kushawah, G., Abugattas-Nuñez del Prado, J., Martinez-Morales, J. R., DeVore, M., Guelfo, J. R., Brannan, E. O., Wang, W., Corbin, T. J., Moran, A. M., Alvarado, A. S., Málaga-Trillo, E., Takacs, C. M., Bazzini, A. A. & Moreno-Mateos, M. A., (2020).

11. Buchman, A. B., Brogan, D. J., Sun, R., Yang, T., Hsu, P. D. & Akbari, O. S. Programmable RNA Targeting Using CasRx in Flies. *CRISPR J* 3, 164-176, (2020).

12. Mahas, A., Aman, R. & Mahfouz, M. CRISPR-Cas13d mediates robust RNA virus interference in plants. *Genome Biol* 20, 263, (2019).

13. Zhou, H., Su, J., Hu, X., Zhou, C., Li, H., Chen, Z., Xiao, Q., Wang, B., Wu, W., Sun, Y., Zhou, Y., Tang, C., Liu, F., Wang, L., Feng, C., Liu, M., Li, S., Zhang, Y., Xu, H., Yao, H., Shi, L. & Yang, H. Glia-to-Neuron Conversion by CRISPR-CasRx Alleviates Symptoms of Neurological Disease in Mice. *Cell* 181, 590-603.e516, (2020).

14. Boyle, A. P., Araya, C. L., Brdlik, C., Cayting, P., Cheng, C., Cheng, Y., Gardner, K., Hillier, L. W., Janette, J., Jiang, L., Kasper, D., Kawli, T., Kheradpour, P., Kundaje, A., Li, J. J., Ma, L., Niu, W., Rehm, E. J., Rozowsky, J., Slattery, M., Spokony, R., Terrell, R., Vafeados, D., Wang, D., Weisdepp, P., Wu, Y. C., Xie, D., Yan, K. K., Feingold, E. A., Good, P. J., Pazin, M. J., Huang, H., Bickel, P. J., Brenner, S. E., Reinke, V., Waterston, R. H., Gerstein, M., White, K. P., Kellis, M. & Snyder, M. Comparative analysis of regulatory information and circuits across distant species. *Nature* 512, 453-456, (2014).

15. Gerstein, M. B., Rozowsky, J., Yan, K. K., Wang, D., Cheng, C., Brown, J. B., Davis, C. A., Hillier, L., Sisu, C., Li, J. J., Pei, B., Harmanci, A. O., Duff, M. O., Djebali, S., Alexander, R. P., Alver, B. H., Auerbach, R., Bell, K., Bickel, P. J., Boeck, M. E., Boley, N. P., Booth, B. W., Cherbas, L., Cherbas, P., Di, C., Dobin, A., Drenkow, J.,

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Ewing, B., Fang, G., Fastuca, M., Feingold, E. A., Frankish, A., Gao, G., Good, P. J., Guigo, R., Hammonds, A., Harrow, J., Hoskins, R. A., Howald, C., Hu, L., Huang, H., Hubbard, T. J., Huynh, C., Jha, S., Kasper, D., Kato, M., Kaufman, T. C., Kitchen, R. R., Ladewig, E., Lagarde, J., Lai, E., Leng, J., Lu, Z., MacCoss, M., May, G., McWhirter, R., Merrihew, G., Miller, D. M., Mortazavi, A., Murad, R., Oliver, B., Olson, S., Park, P. J., Pazin, M. J., Perrimon, N., Pervouchine, D., Reinke, V., Reymond, A., Robinson, G., Samsonova, A., Saunders, G. I., Schlesinger, F., Sethi, A., Slack, F. J., Spencer, W. C., Stoiber, M. H., Strasbourger, P., Tanzer, A., Thompson, O. A., Wan, K. H., Wang, G., Wang, H., Watkins, K. L., Wen, J., Wen, K., Xue, C., Yang, L., Yip, K., Zaleski, C., Zhang, Y., Zheng, H., Brenner, S. E., Graveley, B. R., Celniker, S. E., Gingeras, T. R. & Waterston, R. Comparative analysis of the transcriptome across distant species. *Nature* 512, 445-448, (2014).

16. Long, H., Sung, W., Kucukyildirim, S., Williams, E., Miller, S. F., Guo, W., Patterson, C., Gregory, C., Strauss, C., Stone, C., Berne, C., Kysela, D., Shoemaker, W. R., Muscarella, M. E., Luo, H., Lennon, J. T., Brun, Y. V. & Lynch, M. Evolutionary determinants of genome-wide nucleotide composition. *Nat Ecol Evol* 2, 237-240, (2018).

17. Blanchard, E. L., Vanover, D., Bawage, S. S., Tiwari, P. M., Rotolo, L., Beyersdorf, J., Peck, H. E., Bruno, N. C., Hincapie, R., Finn, M. G., Michel, F., Lafontaine, E. R., Hogan, R. J., Zurla, C. & Santangelo, P. J., (2020).

18. World Health Organization. *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*, < https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>, (2020).

19. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill* 22, (2017).

20. Gonzalez-Reiche, A. S., Hernandez, M. M., Sullivan, M. J., Ciferri, B., Alshammary, H., Obla, A., Fabre, S., Kleiner, G., Polanco, J., Khan, Z., Alburquerque, B., van de Guchte, A., Dutta, J., Francoeur, N., Melo, B. S., Oussenko, I., Deikus, G., Soto, J., Sridhar, S. H., Wang, Y. C., Twyman, K., Kasarskis, A., Altman, D. R., Smith, M., Sebra, R., Aberg, J., Krammer, F., Garcia-Sastre, A., Luksza, M., Patel, G., Paniz-Mondolfi, A., Gitman, M., Sordillo, E. M., Simon, V. & van Bakel, H. Introductions and early spread of SARS-CoV-2 in the New York City area. *Science*, (2020).

21. Cuevas, J. M., Geller, R., Garijo, R., Lopez-Aldeguer, J. & Sanjuan, R. Extremely High Mutation Rate of HIV-1 In Vivo. *PLoS Biol* 13, e1002251, (2015).

22. Kuhn, R. M., Haussler, D. & Kent, W. J. The UCSC genome browser and associated tools. *Brief Bioinform* 14, 144-161, (2013).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Acknowledgements

We thank the entire Sanjana laboratory for support and advice. We thank M. Zaran and S. Brock for assistance with the web-tool server. N.E.S. is supported by New York University and New York Genome Center startup funds, National Institutes of Health (NIH)/National Human Genome Research Institute (grant nos. DP2HG010099), NIH/National Cancer Institute (grant no. R01CA218668), Defense Advanced Research Projects Agency (grant no. D18AP00053), the Sidney Kimmel Foundation, the Melanoma Research Alliance, and the Brain and Behavior Foundation.

Author contributions

N.E.S. and H.H.W. conceived the project. N.E.S., H.H.W., X.G. and A.M.-M. designed the study. A.M.-M. and D.H. designed Cas13d gRNAs for all model organisms. H.H.W built the web tool and performed analyses for model organisms. X.G. designed Cas13d gRNAs and performed analyses for viruses. X.G., H.H.W and D.H. produced the figures. N.E.S. supervised the work. All authors contributed to drafting and reviewing the manuscript, provided feedback and approved the manuscript in its final form.

Competing interests

The New York Genome Center and New York University have applied for patents relating to the work in this article. N.E.S. is an adviser to Vertex.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Methods

gRNA design for model organisms

Reference transcriptomes and corresponding annotations were obtained for each model organism: *H. sapiens* (GENCODE v19, GRCh37), *M. musculus* (GENCODE M24, mm10), *D. rerio* (Ensembl v99, GRCz11), *D. melanogaster* (Ensembl v99, BDGP6), *C. elegans* (Ensembl v99, WBcel235) and *A. thaliana* (Ensembl Plants v46, TAIR10). For each organism, we performed the on-target efficiency predictions for both mRNAs and ncRNAs using command-line *Rfx*Cas13d designer version 0.2 as previously described[5]. We scored gRNAs for all RNA targets with a length of at least 80 nucleotides.

RNA virus genome collection

All full-length RNA virus genomes were downloaded on April 17th, 2020. We downloaded 7,630 complete SARS-CoV-2 viral genomes classified as high coverage and 4,237 Influenza A H1N1 viral genomes with a complete set of eight genomic segments. SARS-CoV-2 and H1N1 genomes were obtained from GISAID (https://www.gisaid.org/). We also analyzed 522 MERS-CoV and 5,557 full length HIV-1 viral genomes, which were downloaded from NCBI Virus (https://www.ncbi.nlm.nih.gov/labs/virus/).

gRNA design to target SARS-CoV-2

We split multi-FASTA files into single-entry FASTA files using the UCSC tool faSplit[22]. All possible 23-mer gRNAs targeting individual genomes were scored with the *Rfx*Cas13 on-target model described previously[5]. All scored guides were classified into four quartiles. Quartile 4 guides (or Q4) are designated to be the predicted best-performing guides. We used USA/NY1-PV08001/2020 (refer to as NY1 isolate) for the SARS-CoV-2 reference gRNA design. Compared to the original (Wuhan) isolate, NY1 contains 3 nucleotide substitutions (G3243A, C25214T, G29027T) resulting in two amino acid mutations (N: A252S, ORF1a: G993S). The SARS-CoV-2 transcript annotation was obtained from NCBI (GenBank: NC_045512.2).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Prediction of minimal numbers of gRNAs to target RNA viruses

For each RNA virus, we identified a minimal set of high-scoring Q4 gRNAs that could target all genomes collected. We used a greedy algorithm as described previously[8]: For each iteration, the gRNAs with the highest number of targeting genomes are added to the set. During each iteration, if multiple gRNAs target the same highest number of genomes, we will pick one for the minimal set and start the next iteration.

Code availability

All designed Cas13 guide RNAs (for model organisms and RNA viruses) and the interactive design tool are available here: https://cas13design.nygenome.org/. For additional reproducibility, we provide shell scripts and R code to reproduce the figures here: https://gitlab.com/sanjanalab/cas13_webtool. The Cas13 guide design algorithm is available here: https://gitlab.com/sanjanalab/cas13.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

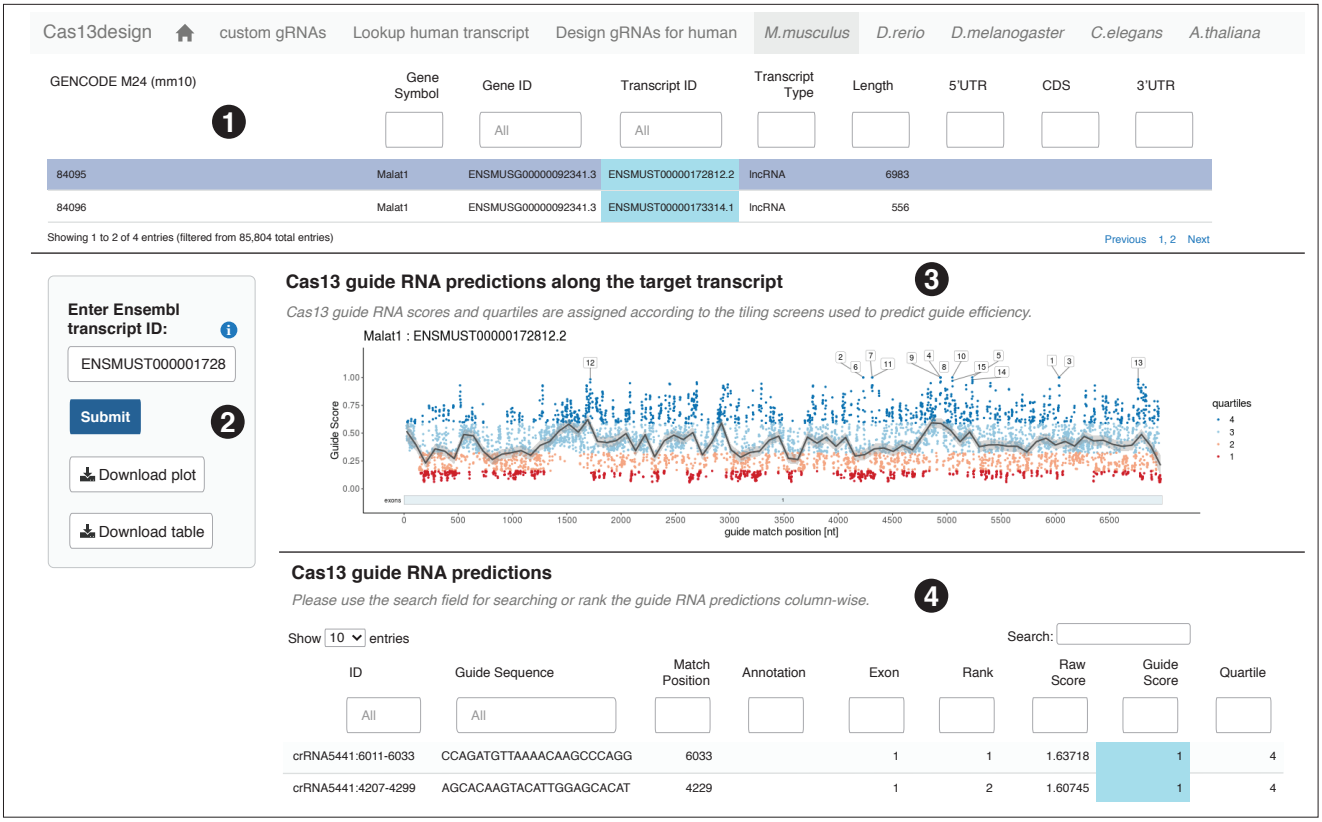
Figure legends

Figure 1. A graphical interface for optimized CRISPR-Cas13d gRNA design for messenger RNAs (mRNAs) and noncoding RNAs (ncRNAs) from six common model organisms. (**a**) Example output of Cas13 Design webtool. Circle (1) allows searches by gene symbol or transcript ID for gRNA design (2), with options to download generated plots and data tables. Circle (3) displays a distribution of gRNAs along the target transcript color-coded by the predicted targeting efficacy scores separated into four quartiles. Q4 gRNAs correspond to those with the highest predicted efficacy and Q1 correspond to those with the lowest predicted efficacy. Circle (4) displays sgRNA options with on-target score predictions. (**b**) The predicted guide efficacy quartiles for mRNAs and ncRNAs across six model organisms. Numbers above bars indicate the percent of transcripts scored that meet the minimal length requirement for target RNAs (80 nt). (**c**) The fraction of processed transcripts that contain at least 1 (up to 25) Q4 gRNAs (predicted high-scoring gRNAs).

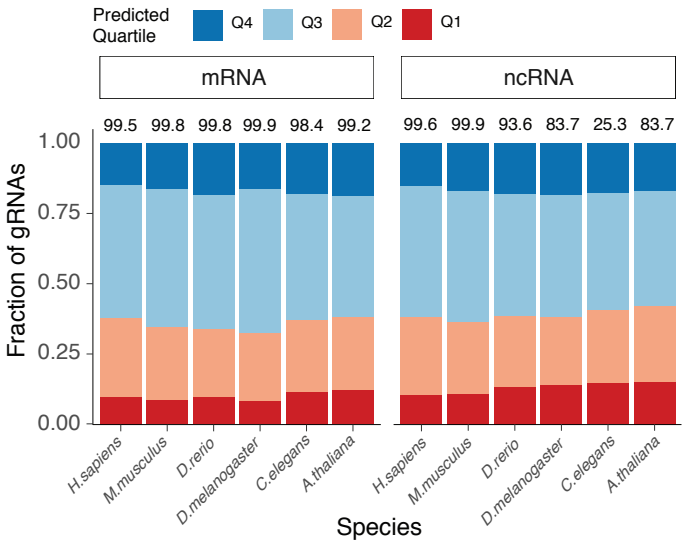
Figure 2. Optimal CRISPR-Cas13d gRNAs to target common human pathogenic RNA viruses. (a) World map of analyzed SARS-CoV-2 isolates (data from GISAID, April 17th, 2020). **(b)** Guide RNA design for each SARS-CoV-2 gene. *Top panel:* SARS-CoV-2 gene annotations. *Middle panel:* Percent of SARS-CoV-2 genomes targeted by each NY1 reference gRNA. *Bottom panel*: Fraction of gRNAs in Q4 per gene (pie) and total number of Q4 gRNAs per gene that targets at least 99% of the total genomes (bar). **(c)** Predicted minimum number of Q4 gRNAs to target all SARS-CoV-2, MERS-CoV, H1N1, and HIV-1 genomes analyzed (n = 7630, 522, 4237 and 5557 viral genomes, respectively).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

a



b



c

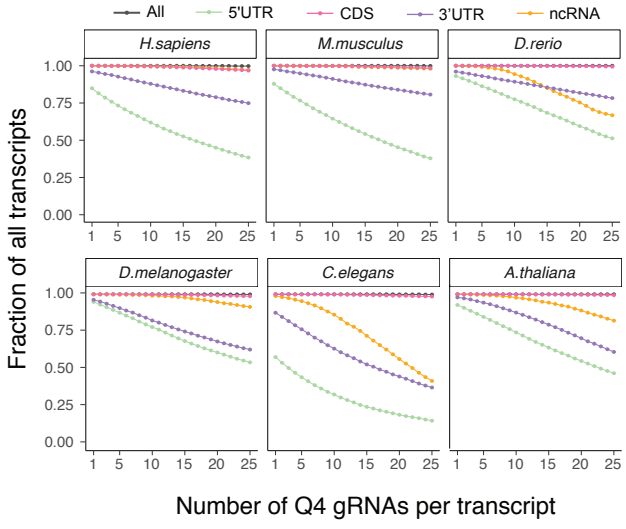


Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.20.259762; this version posted August 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

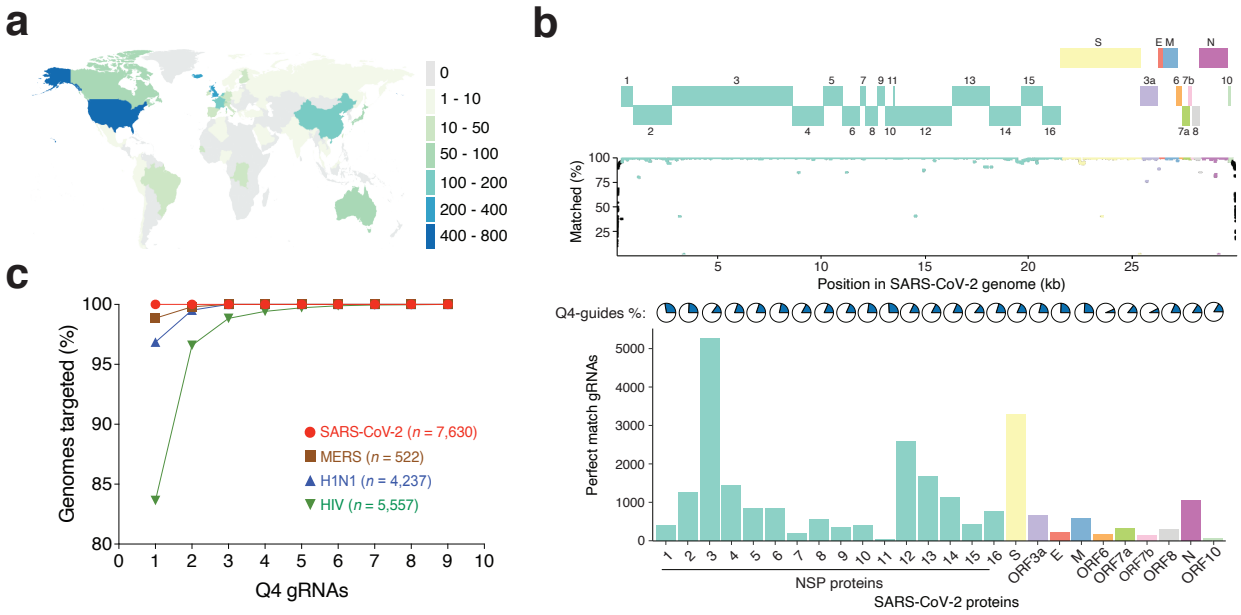


Figure 2