UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Investigation and engineering of RNA-targeting CRISPR-Cas systems for practical biotechnology applications

Permalink

https://escholarship.org/uc/item/0r01m0rx

Author

Brogan, Daniel J.

Publication Date

2023

Supplemental Material

https://escholarship.org/uc/item/0r01m0rx#supplemental

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Investigation and engineering of RNA-targeting CRISPR-Cas systems for practical biotechnology applications

A dissertation submitted
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in

Biochemistry & Molecular Biophysics

by

Daniel J. Brogan

Committee in charge:

Professor Omar S. Akbari, Chair
Professor Alexis C. Komor, Co-Chair
Professor Kevin D. Corbett
Professor Brian S. Leigh
Professor Elizabeth Villa

2024

Copyright
Daniel J. Brogan, 2024
All rights reserved

The dissertation of Daniel J. Brogan is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

I dedicate this dissertation to those who set me on my path toward science, as well as those who have guided me along my journey.

My parents, Joe and Kathy, always allowed me to challenge myself and choose my own direction in life. And my sister, Amanda, has been there to support me throughout my life. My friends and family have always provided cherished memories and laughter. The love and support I have had since childhood has allowed me to persevere and never quit.

My scientific journey started with Ryan McQuillen. Ryan gave me the opportunity to work and learn at Johns Hopkins, where I met my first mentor, Bin Wu. Bin took me on as a research technician and the first member of his lab. He spent a lot of time mentoring me and allowing me to develop as a researcher. And Yang Liu, a postdoc from Bin's lab, trained me on scientific technique and experimentation. Without these 3 people providing me opportunity and training, I would never have found my passion for research.

Research is invigorating and exciting, but it is also quite consuming. It draws you in and begs you to continue through the continuous blows of failure. After all, full commitment is required to obtain knowledge. The effort in research can leave you exhausted and ready to give in. But the people who live by us, support us, and embrace us, allow researchers to continue forward. For me, the person who has stood as my rock and pushed me to be my best self is my wife, Katarina. I love you, Katarina. Thank you for everything.

EPIGRAPH

“Never put a limit on the high stakes”
-Mac Miller
Jet Fuel

TABLE OF CONTENTS

Dissertation Approval Page…….. *iii*

Dedication……. *iv*

Epigraph…… *v*

Table of Contents…….. *vi*

List of Abbreviations…… ix

List of Figures…….. *xi*

List of Tables…… *xiv*

List of Supplemental Files……... *xv*

Acknowledgements…….. *xvi*

Vita……... *xix*

Abstract of Dissertation……..... *xx*

Chapter 1: Introduction……. 1
1.1 The diversity of RNA-targeting CRISPR-Cas systems…….. 1
1.2 The target-dependent collateral cleavage activity of CRISPR-Cas13.... 2
1.3 Type III-E effectors are promising for therapeutic applications……. 3
1.4 Dissertation overview……. 4
1.5 References…….. 6

Chapter 2: Investigation and characterization of a CRISPR-Cas13d *in vivo* RNA-targeting platform in *Drosophila melanogaster*……. 10
2.1 Abstract…….. 10
2.2 Introduction…… 10
2.3 Results…… 13
2.3.1 Genetically encoding CasRx in flies……... 13
2.3.2 Programmable RNA targeting of endogenous target genes… 15
2.3.3 Tissue-Specific RNA Targeting by CasRx……. 16
2.3.4 Criteria for CasRx-mediated phenotypes……... 18

2.3.5 Quantification of CasRx-mediated on/off-target activity…… 20
2.4 Discussion…….. 22
2.5 Methods……. 25
2.5.1 Design and assembly of constructs……. 25
2.5.2 Fly genetics and imaging…… 27
2.5.3 Illumina RNA sequencing…….. 29
2.5.4 Bioinformatics: Quantification and differential expression analysis…… 30
2.6 Acknowledgements……... 32
2.7 Figures……... 33
2.8 Supplemental figures……. 39
2.9 Supplemental tables……... 45
2.10 References……... 48

Chapter 3: Manipulation of Cas13d collateral cleavage activity for the development of a rapid and sensitive nucleic acid diagnostic assay……. 53
3.1 Abstract…….. 53
3.2 Introduction…… 53
3.3 Results…… 56
3.3.1 SENSR workflow……... 56
3.3.2 Target selection and validation……... 57
3.3.3 Optimization of SENSR……. 58
3.3.4 Fluorescence limit of detection…….. 60
3.3.4 Lateral flow assay development……. 61
3.3.5 SENSR detection of SARS-CoV-2 from patient isolates…... 62
3.4 Discussion…….. 63
3.5 Methods……. 65
3.5.1 *In vitro* gRNA cleavage assays…….. 65
3.5.2 RT-RPA amplification of viral genomic sequences……... 66
3.5.3 Fluorescence-based detection of SARS-CoV-2……. 66
3.5.3 Lateral flow detection of SARS-CoV-2……. 67
3.5.4 LoD analysis…….. 68
3.5.5 SENSR detection of patient samples……. 69
3.6 Acknowledgements…… 70
3.7 Figures…… 71
3.8 Supplemental figures……. 78
3.9 Supplemental tables…… 95
3.10 References…… 98
Chapter 4: Advancement of CRISPR diagnostics toward the point of care…… 103
4.1. Abstract……. 103

4.2 Introduction…… 103
4.3 Single reaction CRISPR-Dx…….. 105
4.4 Conclusions…… 108
4.5 Acknowledgments…….. 109
4.6 Figures……..... 110
4.7 Supplementary tables…….. 111
4.8 References…….. 112

Chapter 5: Engineering type III-E CRISPR-Cas systems to develop highly effective and specific RNA-therapeutic technologies……. 116
5.1 Abstract…….. 116
5.2 Introduction…… 116
5.3 Results…… 118
5.3.1 Identification of novel type III-E-like architecture containing a Cas1 domain…….. 118
5.3.2 Engineering synthetic Cas effectors with varying architectures……. 118
5.3.3 Programmable *in vitro* cleavage of ssRNA with SynCas effectors……... 120
5.3.4 Generation of compact SynCas effectors for RNA-targeting……. 122
5.4 Discussion…….. 123
5.5 Methods……. 125
5.5.1 Design and cloning of constructs……... 125
5.5.2 Mammalian cell culture……. 125
5.5.3 RNA-knockdown Assay…… 126
5.5.3 Total RNA collection and qPCR…… 126
5.5.4 Flow Cytometry……. 127
5.5.5 Recombinant Protein Expression and Purification…… 127
5.5.6 Nucleic acid target and crRNA preparation…….. 129
5.5.7 Nuclease assays……. 129
5.6 Acknowledgements……... 130
5.7 Figures……... 131
5.8 Supplemental figures……. 137
5.9 Supplemental tables…….. 145
5.10 References…… 146

LIST OF ABBREVIATIONS

CRISPR	Clustered Regularly Interspaced Short Palindromic Repeat
Cas	CRISPR associated protein
crRNA	CRISPR RNA
gRNA	guide RNA
ssRNA	Single-stranded RNA
HEPN	Higher eukaryotes and prokaryotes nucleotide-binding
qPCR	Quantitative Polymerase Chain Reaction
RT-qPCR	Reverse Transcription quantitative Polymerase Chain Reaction
RNP	Ribonucleoprotein
INS	Insertion
RNAi	RNA interference
CasRx	RfxCas13d
$gRNA^{array}$	Array of multiple gRNAs
RNAseq	RNA sequencing
dCas	Catalytically inactive Cas enzyme
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SENSR	Sensitive Enzymatic Nucleic acid Sequence Reporter
CRISPR-Dx	CRISPR Diagnostics
vRNA	viral RNA
PFS	Protospacer Flanking Sequence

RT-RPA	Reverse transcription Recombinase Polymerase Amplification
CCR	Cas Cleavage Reaction
LF	Lateral Flow
S	SARS-CoV-2 Spike protein
N	SARS-CoV-2 Nucleocapsid protein
HMF	Half-Maximum Fluorescence
FQ	Fluorescence Quencher
NTC	No-template Control
LoD	Limit of Detection
POC	Point-of-Care
SynCas	Synthetic Cas
ODS	Orthologous Domain Substitution
WT	Wild-Type
AA	Amino Acid
DR	Direct Repeat

LIST OF FIGURES

Figure 2.1: Genetic assessment of programmable CasRx-mediated transcript reduction in flies……... 33

Figure 2.2: CasRx-mediated target transcript reduction in restricted tissue types using the binary Gal4/UAS system…… 34

Figure 2.3: Robust CasRx-mediated reduction of GFP…….. 35

Figure 2.4: Quantification of potential CasRx-mediated on/off target activity…….. 37

Figure S2.1: Schematic representation of constructs generated for this study……... 39

Figure S2.2: Complete inheritance data for bidirectional crosses…….. 40

Figure S2.3: Development-related inheritance and lethality of Ubiq-CasRx and Ubiq-dCasRx transheterozygotes……. 41

Figure S2.4: Complete inheritance data for binary Gal4/UAS crosses…….. 42

Figure S2.5: modENCODE transcript expression relative to *Drosophila melanogaster* development…….. 43

Figure S2.6: CasRx-gRNA^{array} transcript target selection and construct generation…….. 44

Figure 3.1: Harnessing SENSR to detect evidence of SARS-CoV-2 viral transcripts…… 71

Figure 3.2: SARS-CoV-2 detection by SENSR via Fluorescence and Lateral Flow assay. 73

Figure 3.3: SENSR detection of positive SARS-CoV-2 validated patient samples…… 75

Figure S3.1: Summary of current CRISPR-based anti-Covid technologies organized by Cas enzyme used and by role as diagnostic or detection tool, or as a putative prophylactic…….. 76

Figure S3.2: CasRx protein purification workflow and quality control…… 78

Figure S3.3: Molecular scale overview of CasRx detection protocol workflow…….. 79

Figure S3.4: Depiction of unique gRNA target sequences across SAR-CoV-2 genome.. 81

Figure S3.5: Preliminary assessment of candidate guides assayed in this study…….. 82

Figure S3.6: Probe di-nucleotide preference selection for SENSR fluorescence detection assay…… 83

Figure S3.7: Optimization of SENSR volume input parameters…….. 84

Figure S3.8: Assessment of optimal amplification time for detection of *Spike* and *Nucleocapsid* targets……. 85

Figure S3.9: Optimization of amplification and cleavage time for lateral flow readout of SENSR…… 86

Figure S3.10: SENSR detection of positive SARS-CoV-2 validated patient samples with unoptimized protocol…… 87

Figure S3.11: SENSR detection of positive SARS-CoV-2 validated patient samples….. 88

Figure S3.12: SENSR lateral flow analysis of RT-qPCR validated SARS-CoV-2 negative patient samples……... 89

Figure S3.13: Viral copy number per μL for RT-qPCR verified positive samples…….. 90

Figure S3.14: Schematic representation of SENSR application for detection of other pathogens……..... 91

Figure S3.15: CasRx direct detection of synthetic Spike ssRNA…… 92

Figure 4.1: CRISPR-Dx formats for detection of nucleic acids……. 110

Figure 5.1: Cas7-1 sequence and predictive structure in relation to type III-E effectors… 131

Figure 5.2: SynCas construction and analysis of initial variants…… 132

Figure 5.3: Biochemical assessment of SynCas cleavage activity……. 133

Figure 5.4: Generation and validation of compact SynCas effectors……. 134

Figure S5.1: Sequence and structure prediction of Cas7-1…….. 135

Figure S5.2: Domain breakdown of initial SynCas variants…… 136

Figure S5.3: Design and assessment of SynCas effectors…… 137

Figure S5.4: Alignment of 29AA N-terminus of Cas1 domain to RT-Cas1 fusion proteins……... 138

Figure S5.5: Biochemical assessment of SynCas.v10……. 139

Figure S5.6: Domain breakdown of compact SynCas variants……... 140

Figure S5.7: Representative gating method for flow cytometry analysis of EGFP knockdown…… 142

LIST OF TABLES

Table S2.2: Complete data set for the Ubiq-CasRx and Ubiq-dCasRx bidirectional crosses…… 45

Table S2.3: Complete data set for the Gal4/UASt-CasRx or Gal4/ UASt-dCasRx crosses…… 46

Table S2.4: Illumina RNA sequencing whole-transcriptome analysis samples……... 47

Table S2.13: Primers used for vector construction……... 47

Table S3.1: Summary Table of current CRISPR-based anti-Covid technologies…… 93

Table S3.2: Analysis identifying putative 30 nt CasRx gRNA target sites conserved across, and specific to, the SARS-CoV-2 genome…….. 94

Table S3.5: RT-RPA protocols for sample volume input variation experiments……. 95

Table S3.6: CCR protocols for RT-RPA volume variation experiments……. 96

Table S3.8: Time-to-detection from fluorescence analysis of all positive patient samples……. 97

Table S4.1: Summary of CRISPR-Dx Technologies Using Collateral Cleavage…… 111

Table S5.1: Breakdown of domains for type III-E effectors based on clustal omega sequence alignment…… 143

Table S5.3: List of crRNAs used in this study……. 144

Table S5.4: List of select plasmids used in this study with respective Addgene IDs….. 145

LIST OF SUPPLEMENTAL FILES

Brogan_Table_S2.1.xlsx

Brogan_Table_S2.5.xlsx

Brogan_Table_S2.6.xlsx

Brogan_Table_S2.7.xlsx

Brogan_Table_S2.8.xlsx

Brogan_Table_S2.9.xlsx

Brogan_Table_S2.10.xlsx

Brogan_Table_S2.11.xlsx

Brogan_Table_S2.12.xlsx

Brogan_File_S2.1.pl

Brogan_File_S2.2.pl

Brogan_Table_S3.3.xlsx

Brogan_Table_S3.4.xlsx

Brogan_Table_S3.7.xlsx

Brogan_Table_S5.2.xlsx

Brogan_File_S5.1.clustal

Brogan_File_S5.2.clustal

Brogan_File_S5.3-5.25.pdb

Brogan_File_S5.26.clustal

Brogan_File_S5.27-5.62.pdb

ACKNOWLEDGMENTS

I would like to acknowledge the following people:

Dr. Omar S. Akbari for supporting and guiding me throughout my graduate studies. Your mentorship provided me a freedom to explore and innovate that has allowed me to be the scientist I am today.

Dr. Alexis C. Komor for being there for me since the beginning of my time at UCSD and for constantly setting aside time to help me with anything from science to life.

Dr. Kevin D. Corbett for taking the time to listen to my project ideas and goals and providing incredibly constructive feedback.

Dr. Brian S. Leigh for showing me how to engage students so that they learn while being challenged and engaged.

Dr. Elizabeth Villa for taking the time to provide me with advice and the many fun conversations in Tata Hall.

Dr. Elizabeth A. Komives for supporting my research through fantastic collaborations and guiding me on biochemical and biophysical approaches to my work.

Dr. Robyn Raban for all of the support and conversations about research, science in general, the biotech industry, and life.

Dr. Elena Dalla Benetta for being the best person to work on a project with and helping me develop as both a scientist and a mentor.

Calvin P. Lin for being an incredible friend and collaborator. Without you my life in grad school would have never been the same and I will never forget that.

Tianqi (Nancy) Wang for allowing me to grow as a mentor and for all your help with experiments.

Dr. Igor Antoshechkin for the help in all things bioinformatics.

Dr. Stephanie Gamez, Dr. Michelle Bui, and Junru (aka the grad students of the Akbari Lab), for all of the advice, discussions, and fun conversations throughout grad school.

My mentees, Claire Lin and Nathan Pham, who were incredible undergraduate researchers and patient while I learned to mentor.

The members of the Komor lab for providing materials and advice and helping with any questions I had about research or experimental design.

My friends, Kai Schwennicke, Quinn T. Cowan, and Dr. Ximena Garcia Arceo, for all of the laughs and calmness during the chaos of graduate school.

Finally, I would like to thank Dr. Nikolay P. Kandul, Dr. Ming Li, Dr. Hsing-Han Li, Dr. Shih-Che Weng, and Dr. Fangying Chen for help or advice on projects during my graduate studies and all current and former members of the Akbari Lab.

The material in this dissertation is based on the following papers which are either published, submitted for publication, or preparing for submission.

Chapter 2, in full, is a reprint of the material as it appears in The CRISPR Journal. Anna B. Buchman[*], Daniel J Brogan[*], Ruichen Sun[*], Ting Yang, Igor Antoshechkin, Patrick D. Hsu, Omar S. Akbari. “Programmable RNA-targeting using CasRx in flies”. The CRISPR Journal. (2020). [* = Equal contribution]. The dissertation/thesis author was the co-first-author and primary investigator of this paper.

Chapter 3, in full, is a reprint of the material as it appears in ACS Sensors. Daniel J Brogan[*], Duverney Chaverra-Rodriguez[*], Calvin P Lin[*], Andrea L Smidler, Ting Yang, Lenissa M. Alcantara, Igor Antoshechkin, Junru Liu, Robyn R Raban, Pedro Belda-Ferre, Rob Knight,

Elizabeth A Komives, Omar S. Akbari. “Development of a Rapid and Sensitive CasRx-Based Diagnostic Assay for SARS-CoV-2”. ACS Sensors. (2021). [* = Equal contribution]. The dissertation/thesis author was the first-author and primary investigator of this paper.

Chapter 4, in full, is a copy of the material currently in preparation for publication and submitted to Nature Biotechnology. Daniel J Brogan, Elena Dalla Benetta, Tianqi Wang, Calvin P. Lin, Fangying Chen, Harry Li, Claire Lin, Elizabeth A. Komives, Omar S. Akbari. “Engineered type III-E CRISPR-Cas effectors”. (2023). The dissertation/thesis author is the first-author and primary investigator of this paper.

VITA

2016	Bachelor of Science, Ursinus College
2016-2018	Research Technician, Johns Hopkins University
2018-2019	Instructional Assistant, University of California San Diego
2020	Master of Science, University of California San Diego
2018-2023	Graduate Research Assistant, University of California San Diego
2024	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

"Engineered type III-E CRISPR-Cas effectors" Submitted to Nature Communications. (2023).

"Development and evaluation of a CasRx-based SENSR assay for detection of feline coronavirus nucleic acid" Submitted to GEN Biotechnology. (2023).

"Targeting Sex Determination to Suppress Mosquito Populations" eLife. (2023).

"Engineered Antiviral Sensor Targets Infected Mosquitoes" Submitted to The CRISPR Journal. (2023).

"CRISPR Diagnostics: Advances toward the Point of Care" Biochemistry. (2022)

"Development of a Rapid and Sensitive CasRx-Based Diagnostic Assay for SARS-CoV-2" ACS Sensors. (2021).

"Ubiquitous and Tissue-specific RNA Targeting in Drosophila Melanogaster using CRISPR/CasRx" Journal of Visualized Experiments. (2021).

"Programmable RNA Targeting Using CasRx in Flies" The CRISPR Journal. (2020).

ABSTRACT OF DISSERTATION

Investigation and engineering of RNA-targeting CRISPR-Cas systems for practical biotechnology applications

by

Daniel J. Brogan

Doctor of Philosophy in Biochemistry & Molecular Biophysics

University of California San Diego, 2024

Professor Omar Sultan Akbari, Chair

The discovery of CRISPR-Cas prokaryotic and archaeal immune systems has spurred the development of revolutionary precision molecular tools. While CRISPR DNA endonucleases have been rigorously characterized, novel RNA-targeting CRISPR systems are being explored as new discoveries rapidly emerge. Since their discovery, RNA-targeting CRISPR systems have quickly been applied for programmable transcript targeting, RNA therapeutics, and pathogen detection. Upon discovery, the Cas13d subtype was the smallest and most efficient RNA-targeting CRISPR system discovered. However, like all discovered Cas13 enzymes, Cas13d ribonucleases exhibit target-dependent collateral cleavage activity that limits the applicability of this technology. More recently, type III-E CRISPR-Cas effectors were

characterized and exhibit significant RNA knockdown without collateral cleavage activity. Unlike Cas13 systems, type III-E effectors are very large and efficiencies can vary dramatically. In this body of work, we characterize and harness the collateral cleavage activity of Cas13d and subsequently establish a method to engineer and develop practical type III-E effectors for therapeutic applications.

Chapter 1: Introduction

1.1 The diversity of RNA-targeting CRISPR-Cas systems

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and their associated proteins (Cas) encompass a rapidly expanding set of programmable endonucleases that have evolved as prokaryotic and archaeal regulatory immune system (Makarova et al. 2011). CRISPR-Cas systems are initially divided into two classes: class 1 and class 2. Class 1 CRISPR systems typically accomplish nuclease activity through coordination of multiple subunits into a large complex, whereas class 2 systems operate using a single, large endonuclease (Makarova et al. 2015). Since the first demonstration of programmable DNA-editing by the class 2 effector Cas9 in 2012 (Jinek et al. 2012), there has been an explosion of CRISPR discovery and engineering pipelines all working to develop innovative genome engineering technologies. Significant effort has been focused on DNA-targeting CRISPR systems (Cong et al. 2013; Mali et al. 2013; Zetsche et al. 2015; Komor et al. 2016), but the recent emergence of RNA-targeting CRISPR-Cas systems provides a unique pathway to produce precise transcriptome engineering technologies for transient control gene expression.

RNA-targeting CRISPR-Cas systems encompass a diverse array of ribonuclease effectors capable of programmably cleaving single-stranded RNA (ssRNA) molecules. Two main types of effectors exist: Class 1, type III and Class 2, type VI. Type III CRISPR-Cas systems are composed of 6 subtypes (A-F) and typically utilize multisubunit complex formation to accomplish RNA (or DNA) cleavage (van Beljouw et al. 2023). Type VI CRISPR-Cas systems (Cas13), however, are composed of four subtypes (A-D) and operate as a single-effector protein (O'Connell 2019). Cas13d effectors were discovered in 2018 and were found to be highly efficient and also the most compact Cas13 effector subtype (Yan et al. 2018; Konermann et al.

2018). The type III-E (Cas7-11) subtype was discovered and characterized in 2021 and, unlike other class 1 CRISPR systems, accomplishes RNA-targeting with the largest known single-effector CRISPR-Cas proteins (van Beljouw et al. 2021; Özcan et al. 2021). Both Cas13d and type III-E effectors have advantages and disadvantages, and easily enable development of extraordinary biotechnology platforms.

1.2 The target-dependent collateral cleavage activity of CRISPR-Cas13

Since the discovery of Cas13 in 2016 (East-Seletsky et al. 2016; Abudayyeh et al. 2016), these enzymes have been thoroughly characterized and engineered (Abudayyeh et al. 2017; Cox et al. 2017; Meeske, Nakandakari-Higa, and Marraffini 2019). As tools, Cas13, and more specifically Cas13d, had great potential as a platform technology for therapeutics due to the robust knockdown activity and small size. However, all Cas13 enzymes exhibit a target-dependent collateral cleavage activity *in vitro* (East-Seletsky et al. 2016; Abudayyeh et al. 2016; Smargon et al. 2017; Yan et al. 2018; Konermann et al. 2018), where, upon target cleavage, Cas13 cleaves bystander ssRNA molecules. Both the on-target and collateral cleavage activity are driven by two HEPN domains that are conserved among all Cas13 effectors (O'Connell 2019). Inactivation of these domains through targeted mutagenesis generates a catalytically inactive Cas13 (dCas13) mutant that is capable of programmable RNA-binding. Efforts to engineer (Tong et al. 2022) or discover (Wei et al. 2022) Cas13 enzymes with lower or no collateral activity are ongoing, but it is also possible to take advantage of this activity to build nucleic acid diagnostic systems.

CRISPR-based nucleic acid diagnostics offer a promising avenue for the development of practical nucleic acid diagnostics (Gootenberg et al. 2017). Cas13 collateral cleavage is target

dependent and therefore can serve as a reporter for the presence of specific nucleic acid sequences. This is accomplished by incubating the target RNA with a collateral probe often modified with a 5' fluorescence tag and a 3' fluorescence quencher (East-Seletsky et al. 2016). Other readout modifications, such as for lateral flow, have also been demonstrated (Gootenberg et al. 2018). Cas13 diagnostics can be combined with isothermal amplification systems and this simplicity offers the development of portable diagnostic tools with numerous applications (Myhrvold et al. 2018), compared to complex, yet more sensitive diagnostics such as qPCR. Although Cas13 diagnostic technologies already present promise for application, these tools are still in their nascency and will require future innovations to advance toward the point-of-care.

1.3 Type III-E offer a viable alternative to Cas13 for therapeutic applications

Collateral cleavage activity, as well as complications with immunogenicity (Tang et al. 2022), currently limit the development of Cas13-based RNA-targeting therapeutic platforms. Cas7-11 effectors programmably target ssRNA molecules, but completely lack collateral cleavage activity (Özcan et al. 2021). Cas7-11 effectors are composed of four Cas7 domains, a Cas11 domain, and a large insertion (INS) domain, where the Cas7 and Cas11 domains coordinate target cleavage (Kato, Zhou, et al. 2022; Goswami et al. 2022). Type III-E CRISPR systems evolved with a protease-activated caspase immune system. When the Cas7-11 ribonucleoprotein (RNP) complex binds with the target molecule, the protein undergoes a conformational change that permits binding with an associated protease (Hu et al. 2022; Ekundayo et al. 2023). Interestingly, catalytic inactivation of Cas7-11 resulted in a more robust immune response (Kato, Okazaki, et al. 2022), suggesting target binding is the key determinant for sufficient immunity against invading nucleic acids.

Due to the evolution of type III-E systems, these enzymes accomplish catalysis slowly and therefore are typically less effective than Cas13 enzymes. Cas7-11 effectors are also extremely large proteins, which complicates delivery of these systems for therapeutic applications. Efforts have quickly been made to reduce the size of Cas7-11 proteins. One study produced a more compact Cas7-11 effector through deletion of the INS domain, which minimally impacted RNA-targeting activity compared to wild-type (Kato, Zhou, et al. 2022). This initial engineering effort presents promise for future iterations of engineered Cas7-11 variants. These engineering efforts could improve catalytic activity of Cas7-11 proteins, while also permitting development of practical Cas7-11-based therapeutic technologies. Since Cas7-11 proteins were discovered so recently, there is promise that evolved forms of these technologies will outperform their native ancestors.

1.4 Dissertation Overview

In this body of work, we will explore CRISPR-based RNA-targeting applications and demonstrate their versatility as platform technologies. We start with the investigation of a Cas13 enzyme, RfxCas13d or CasRx, as an *in vivo* RNA-targeting platform in *Drosophila melanogaster*, where we uncover, at the genetic and molecular levels, the toxic and lethal effects of collateral cleavage activity (Buchman et al. 2020). We then manipulate the collateral cleavage activity of CasRx to generate a sensitive and potentially portable (Brogan and Akbari 2022) diagnostic technology capable of detecting infectious diseases in patient isolates (Brogan et al. 2021). Furthermore, this diagnostic system is applicable beyond infectious disease and can be used to quantify expected collateral activity for *in vivo* expression (Benetta et al. 2023) or detection of transgenic elements in engineered vectors (Li et al. 2023). Finally, we establish an

engineering method that reveals a unique modularity of Cas7-11 proteins and apply this method to generate synthetic CRISPR proteins with comparable activity to wild-type. We also find an unusual crRNA processing activity of Cas7-11 proteins that enables RNA engineering such as developing aptamer-based RNA-targeting technologies. Together, this will present the understanding, utilization, and engineering of RNA-targeting CRISPR-Cas systems for a variety of applications and demonstrating the importance of these systems as biotechnology platforms.

1.5 References

Abudayyeh, Omar O., Jonathan S. Gootenberg, Patrick Essletzbichler, Shuo Han, Julia Joung, Joseph J. Belanto, Vanessa Verdine, et al. 2017. “RNA Targeting with CRISPR-Cas13.” *Nature* 550 (7675): 280–84. https://doi.org/10.1038/nature24049.

Abudayyeh, Omar O., Jonathan S. Gootenberg, Silvana Konermann, Julia Joung, Ian M. Slaymaker, David B. T. Cox, Sergey Shmakov, et al. 2016. “C2c2 Is a Single-Component Programmable RNA-Guided RNA-Targeting CRISPR Effector.” *Science* 353 (6299): aaf5573. https://doi.org/10.1126/science.aaf5573.

Beljouw, Sam P. B. van, Anna C. Haagsma, Alicia Rodríguez-Molina, Daan F. van den Berg, Jochem N. A. Vink, and Stan J. J. Brouns. 2021. “The gRAMP CRISPR-Cas Effector Is an RNA Endonuclease Complexed with a Caspase-like Peptidase.” *Science* 373 (6561): 1349–53. https://doi.org/10.1126/science.abk2718.

Beljouw, Sam P. B. van, Jasper Sanders, Alicia Rodríguez-Molina, and Stan J. J. Brouns. 2023. “RNA-Targeting CRISPR-Cas Systems.” *Nature Reviews. Microbiology* 21 (1): 21–34. https://doi.org/10.1038/s41579-022-00793-y.

Benetta, Elena Dalla, Adam J. López-Denman, Hsing-Han Li, Reem A. Masri, Daniel J. Brogan, Michelle Bui, Ting Yang, et al. 2023. “Engineered Antiviral Sensor Targets Infected Mosquitoes.” *bioRxiv*. https://doi.org/10.1101/2023.01.27.525922.

Brogan, Daniel J., and Omar S. Akbari. 2022. “CRISPR Diagnostics: Advances toward the Point of Care.” *Biochemistry*, March. https://doi.org/10.1021/acs.biochem.2c00051.

Brogan, Daniel J., Duverney Chaverra-Rodriguez, Calvin P. Lin, Andrea L. Smidler, Ting Yang, Lenissa M. Alcantara, Igor Antoshechkin, et al. 2021. “Development of a Rapid and Sensitive CasRx-Based Diagnostic Assay for SARS-CoV-2.” *ACS Sensors*, October. https://doi.org/10.1021/acssensors.1c01088.

Buchman, Anna B., Dan J. Brogan, Ruichen Sun, Ting Yang, Patrick D. Hsu, and Omar S. Akbari. 2020. “Programmable RNA Targeting Using CasRx in Flies.” *The CRISPR Journal* 3 (3): 164–76. https://doi.org/10.1089/crispr.2020.0018.

Cong, Le, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu, et al. 2013. “Multiplex Genome Engineering Using CRISPR/Cas Systems.” *Science* 339 (6121): 819–23. https://doi.org/10.1126/science.1231143.

Cox, David B. T., Jonathan S. Gootenberg, Omar O. Abudayyeh, Brian Franklin, Max J. Kellner, Julia Joung, and Feng Zhang. 2017. “RNA Editing with CRISPR-Cas13.” *Science* 358 (6366): 1019–27. https://doi.org/10.1126/science.aaq0180.

East-Seletsky, Alexandra, Mitchell R. O’Connell, Spencer C. Knight, David Burstein, Jamie H. D. Cate, Robert Tjian, and Jennifer A. Doudna. 2016. “Two Distinct RNase Activities of

CRISPR-C2c2 Enable Guide-RNA Processing and RNA Detection." *Nature* 538 (7624): 270–73. https://doi.org/10.1038/nature19802.

Ekundayo, Babatunde, Davide Torre, Bertrand Beckert, Sergey Nazarov, Alexander Myasnikov, Henning Stahlberg, and Dongchun Ni. 2023. "Structural Insights into the Regulation of Cas7-11 by TPR-CHAT." *Nature Structural & Molecular Biology* 30 (2): 135–39. https://doi.org/10.1038/s41594-022-00894-5.

Gootenberg, Jonathan S., Omar O. Abudayyeh, Max J. Kellner, Julia Joung, James J. Collins, and Feng Zhang. 2018. "Multiplexed and Portable Nucleic Acid Detection Platform with Cas13, Cas12a, and Csm6." *Science* 360 (6387): 439–44. https://doi.org/10.1126/science.aaq0179.

Gootenberg, Jonathan S., Omar O. Abudayyeh, Jeong Wook Lee, Patrick Essletzbichler, Aaron J. Dy, Julia Joung, Vanessa Verdine, et al. 2017. "Nucleic Acid Detection with CRISPR-Cas13a/C2c2." *Science*. https://doi.org/10.1126/science.aam9321.

Goswami, Hemant N., Jay Rai, Anuska Das, and Hong Li. 2022. "Molecular Mechanism of Active Cas7-11 in Processing CRISPR RNA and Interfering Target RNA." *eLife* 11 (October). https://doi.org/10.7554/eLife.81678.

Hu, Chunyi, Sam P. B. van Beljouw, Ki Hyun Nam, Gabriel Schuler, Fran Ding, Yanru Cui, Alicia Rodríguez-Molina, et al. 2022. "Craspase Is a CRISPR RNA-Guided, RNA-Activated Protease." *Science* 377 (6612): 1278–85. https://doi.org/10.1126/science.add5064.

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012. "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity." *Science* 337 (6096): 816–21. https://doi.org/10.1126/science.1225829.

Kato, Kazuki, Sae Okazaki, Cian Schmitt-Ulms, Kaiyi Jiang, Wenyuan Zhou, Junichiro Ishikawa, Yukari Isayama, et al. 2022. "RNA-Triggered Protein Cleavage and Cell Growth Arrest by the Type III-E CRISPR Nuclease-Protease." *Science* 378 (6622): 882–89. https://doi.org/10.1126/science.add7347.

Kato, Kazuki, Wenyuan Zhou, Sae Okazaki, Yukari Isayama, Tomohiro Nishizawa, Jonathan S. Gootenberg, Omar O. Abudayyeh, and Hiroshi Nishimasu. 2022. "Structure and Engineering of the Type III-E CRISPR-Cas7-11 Effector Complex." *Cell* 0 (0). https://doi.org/10.1016/j.cell.2022.05.003.

Komor, Alexis C., Yongjoo B. Kim, Michael S. Packer, John A. Zuris, and David R. Liu. 2016. "Programmable Editing of a Target Base in Genomic DNA without Double-Stranded DNA Cleavage." *Nature* 533 (7603): 420–24. https://doi.org/10.1038/nature17946.

Konermann, Silvana, Peter Lotfy, Nicholas J. Brideau, Jennifer Oki, Maxim N. Shokhirev, and Patrick D. Hsu. 2018. "Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors." *Cell* 173 (3): 665–76.e14. https://doi.org/10.1016/j.cell.2018.02.033.

Li, Ming, Nikolay P. Kandul, Ruichen Sun, Ting Yang, Elena D. Benetta, Daniel J. Brogan, Igor Antoshechkin, et al. 2023. “Targeting Sex Determination to Suppress Mosquito Populations.” *bioRxiv*. https://doi.org/10.1101/2023.04.18.537404.

Makarova, Kira S., Daniel H. Haft, Rodolphe Barrangou, Stan J. J. Brouns, Emmanuelle Charpentier, Philippe Horvath, Sylvain Moineau, et al. 2011. “Evolution and Classification of the CRISPR-Cas Systems.” *Nature Reviews. Microbiology* 9 (6): 467–77. https://doi.org/10.1038/nrmicro2577.

Makarova, Kira S., Yuri I. Wolf, Omer S. Alkhnbashi, Fabrizio Costa, Shiraz A. Shah, Sita J. Saunders, Rodolphe Barrangou, et al. 2015. “An Updated Evolutionary Classification of CRISPR-Cas Systems.” *Nature Reviews. Microbiology* 13 (11): 722–36. https://doi.org/10.1038/nrmicro3569.

Mali, Prashant, Luhan Yang, Kevin M. Esvelt, John Aach, Marc Guell, James E. DiCarlo, Julie E. Norville, and George M. Church. 2013. “RNA-Guided Human Genome Engineering via Cas9.” *Science* 339 (6121): 823–26. https://doi.org/10.1126/science.1232033.

Meeske, Alexander J., Sandra Nakandakari-Higa, and Luciano A. Marraffini. 2019. “Cas13-Induced Cellular Dormancy Prevents the Rise of CRISPR-Resistant Bacteriophage.” *Nature* 570 (7760): 241–45. https://doi.org/10.1038/s41586-019-1257-5.

Myhrvold, Cameron, Catherine A. Freije, Jonathan S. Gootenberg, Omar O. Abudayyeh, Hayden C. Metsky, Ann F. Durbin, Max J. Kellner, et al. 2018. “Field-Deployable Viral Diagnostics Using CRISPR-Cas13.” *Science* 360 (6387): 444–48. https://doi.org/10.1126/science.aas8836.

O’Connell, Mitchell R. 2019. “Molecular Mechanisms of RNA Targeting by Cas13-Containing Type VI CRISPR–Cas Systems.” *Journal of Molecular Biology* 431 (1): 66–87. https://doi.org/10.1016/j.jmb.2018.06.029.

Özcan, Ahsen, Rohan Krajeski, Eleonora Ioannidi, Brennan Lee, Apolonia Gardner, Kira S. Makarova, Eugene V. Koonin, Omar O. Abudayyeh, and Jonathan S. Gootenberg. 2021. “Programmable RNA Targeting with the Single-Protein CRISPR Effector Cas7-11.” *Nature*, September, 1–6. https://doi.org/10.1038/s41586-021-03886-5.

Smargon, Aaron A., David B. T. Cox, Neena K. Pyzocha, Kaijie Zheng, Ian M. Slaymaker, Jonathan S. Gootenberg, Omar A. Abudayyeh, et al. 2017. “Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28.” *Molecular Cell*. https://doi.org/10.1016/j.molcel.2016.12.023.

Tang, Xin-Zi Emily, Shu Xuan Tan, Shawn Hoon, and Gene W. Yeo. 2022. “Pre-Existing Adaptive Immunity to the RNA-Editing Enzyme Cas13d in Humans.” *Nature Medicine* 28 (7): 1372–76. https://doi.org/10.1038/s41591-022-01848-6.

Tong, Huawei, Jia Huang, Qingquan Xiao, Bingbing He, Xue Dong, Yuanhua Liu, Xiali Yang, et

al. 2022. “High-Fidelity Cas13 Variants for Targeted RNA Degradation with Minimal Collateral Effects.” *Nature Biotechnology*, August. https://doi.org/10.1038/s41587-022-01419-7.

Wei, Jingyi, Peter Lotfy, Kian Faizi, Eleanor Wang, Hannah Slabodkin, Emily Kinnaman, Sita Chandrasekaran, et al. 2022. “Deep Learning and CRISPR-Cas13d Ortholog Discovery for Optimized RNA Targeting.” *bioRxiv*. https://doi.org/10.1101/2021.09.14.460134.

Yan, Winston X., Shaorong Chong, Huaibin Zhang, Kira S. Makarova, Eugene V. Koonin, David R. Cheng, and David A. Scott. 2018. “Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein.” *Molecular Cell* 70 (2): 327–39.e5. https://doi.org/10.1016/j.molcel.2018.02.028.

Zetsche, Bernd, Jonathan S. Gootenberg, Omar O. Abudayyeh, Ian M. Slaymaker, Kira S. Makarova, Patrick Essletzbichler, Sara E. Volz, et al. 2015. “Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System.” *Cell* 163 (3): 759–71. https://doi.org/10.1016/j.cell.2015.09.038.

Chapter 2: Investigation and characterization of a CRISPR-Cas13d *in vivo* RNA-targeting platform in *Drosophila melanogaster*

2.1 Abstract

CRISPR-Cas genome editing technologies have revolutionized the fields of functional genetics and genome engineering, but with the recent discovery and optimization of RNA-targeting Cas ribonucleases, we may soon see a similar revolution in the study of RNA function and transcriptome engineering. However, to date, successful proof of principle for Cas ribonuclease RNA targeting in eukaryotic systems has been limited. Only recently has successful modification of RNA expression by a Cas ribonuclease been demonstrated in animal embryos. This previous work, however, did not evaluate endogenous expression of Cas ribonucleases and only focused on function in early developmental stages. A more comprehensive evaluation of this technology is needed to assess its potential impact. Here we report on our efforts to develop a programmable platform for RNA targeting using a Cas ribonuclease, CasRx, in the model organism *Drosophila melanogaster*. By genetically encoding CasRx in flies, we demonstrate moderate transcript targeting of known phenotypic genes in addition to unexpected toxicity and lethality. We also report on the off-target effects following on-target transcript cleavage by CasRx. Taken together, our results present the current state and limitations of a genetically encoded programmable RNA-targeting Cas system in *Drosophila melanogaster*, paving the way for future optimization of the system.

2.2 Introduction

The development of CRISPR as a programmable genome engineering tool has revolutionized the life sciences by providing transformative applications for both medicine and

biotechnology (Adli 2018). While much of the recent focus has been on exploiting CRISPR technologies to target DNA, recent findings that certain CRISPR systems can also be programmed to target RNA have suggested new possibilities for CRISPR technologies in transcriptome engineering (Abudayyeh et al. 2016; Konermann et al. 2018; East-Seletsky et al. 2017). For example, one recent advancement was the engineering and biochemical characterization of Cas ribonuclease (CasRx) as a compact single-effector Cas enzyme that exclusively targets RNA with superior efficiency and specificity as compared to RNA interference (RNAi) (Konermann et al. 2018). In human cells, CasRx demonstrated highly efficient on-target gene reduction with limited off-target activity, making it a potential tool for gene reduction. However, this technology has yet to be comprehensively adapted for facile use in other systems (although see (Kushawah et al. 2020)), such as *Drosophila melanogaster* (flies), which are a common tool for exploring new biological questions and developing novel bioengineering tools *in vivo.* Non-RNAi-based techniques for reducing gene expression (without permanently altering the genome) in animals would provide for a more flexible technique to modulate gene expression in a biologically relevant way.

CasRx belongs to the Cas13 family of RNA-targeting Cas enzymes, a group of highly specific ribonucleases (Konermann et al. 2018; Abudayyeh et al. 2017). Though these enzymes possess promiscuous RNase activity resulting in cleavage of non-target RNA (Konermann et al. 2018; East-Seletsky et al. 2016; Smargon et al. 2017; Yan et al. 2018), a possible drawback for applying Cas13 ribonuclease-based transcriptome engineering technologies, they may serve as a starting point for optimizing these RNA-targeting platforms for *in vivo* applications. For example, RNA-targeting CRISPR technologies could enable the development of robust gene silencing techniques in animals in which RNAi poorly functions (Konermann et al. 2018;

Perrimon, Ni, and Perkins 2010). Another potential application may involve using RNA-targeting CRISPR technologies to engineer mosquito populations resistant to infection with RNA viruses. Numerous RNA viruses of global medical importance, such as dengue, Zika, and chikungunya virus, are transmitted primarily by one species of *Aedes* mosquito. Engineering this mosquito population with virus resistance may be a tool to reduce disease transmission (Champer, Buchman, and Akbari 2016); however, no current technologies have successfully targeted all of these viruses simultaneously (Buchman et al. 2020; Mathur et al. 2010; Franz et al. 2006; Yen et al. 2018; Buchman et al. 2019). RNA-targeting CRISPR systems may provide a platform to reduce the spread of mosquito-borne viruses by targeting viral RNA genomes in a programmable manner. Therefore, it is of high priority to further understand the utility of RNA-targeting CRISPR systems in relevant model organisms.

RNAi-based approaches are the current standard for transcriptome modification in *Drosophila.* This technology has increased our understanding of the function and regulation of many genes (Perrimon, Ni, and Perkins 2010; Dietzl et al. 2007; Ni et al. 2009, 2008, 2011), yet RNAi was reported to show occasional high false negative rates, particularly in highly expressed genes due to insufficient small RNA expression (Perrimon, Ni, and Perkins 2010; Dietzl et al. 2007; Heigwer, Port, and Boutros 2018), and at some other times high false positive rates due to positional or off-target effects (Kulkarni et al. 2006; Ma et al. 2006; Perrimon and Mathey-Prevot 2007; Markstein et al. 2008). Co-expression of Dicer2 can reduce false negative rates, but would in turn increase the prevalence of false positives (Perrimon, Ni, and Perkins 2010; Dietzl et al. 2007) and render the entire process not as clean. Ideally, an RNA-targeting system should be easily programmable, not require expression of multiple factors, and should work in a simplified manner. CasRx, like other CRISPR systems, is easily programmable (Jinek et al. 2012;

Chakraborty, Teoh, and Das 2017) and is capable of targeting nearly any coding gene, but unlike other Cas13 enzymes, it lacks a protospacer flanking sequence requirement (Konermann et al. 2018), making it more versatile for programmable targeting. Additionally, CasRx is a simplified RNA-targeting system as it requires no additional helper enzymes to efficiently target and degrade RNA (Konermann et al. 2018). For these reasons, the CasRx ribonuclease is a practical starting point for establishing a single-effector RNA-targeting platform for *in vivo* gene reduction studies. Here we report the first use of a CasRx-mediated RNA-targeting system in flies. We show that separately encoding CasRx and guide RNA arrays (gRNAarray) in the genome promotes robust expression of these elements throughout development. Furthermore, we demonstrate that binary genetic crosses with ubiquitous and tissue-specific CasRx- and gRNAarray-expressing fly lines can produce clear, highly penetrant phenotypes and by using RNA sequencing (RNAseq) we demonstrate that CasRx is capable of moderate targeted transcript reduction at various stages of fly development, albeit with various degrees of off-target activity. Moreover, we also found that CasRx expression and targeting was often toxic and resulted in unexpected lethality indicating further optimization will be necessary for this to be a versatile tool for *Drosophila* genetics.

2.3 Results

2.3.1 Genetically encoding CasRx in flies

To genetically determine the RNA-targeting capabilities of CasRx, *in vivo*, we engineered flies encoding two versions of the CasRx ribonuclease: the active enzyme and a catalytically inactive negative control (dCasRx). We did this by generating transgenes that use a broadly expressing ubiquitin (Ubiq) promoter (Akbari et al. 2009) to drive expression of either CasRx

(Ubiq-CasRx) or dCasRx (Ubiq-dCasRx) (Supplementary Fig. S2). CasRx exhibits two distinct RNase activities for processing its cognate gRNAarray and cleaving target RNA. Because we wanted our negative control to still bind target RNA and efficiently process the gRNAarray, we eliminated programmable RNA cleavage in dCasRx by inactivating the positively charged catalytic residues of the HEPN motifs (Konermann et al. 2018). We established these transgenic lines by integrating each transgene site-specifically using an available φC31 docking site located on the second chromosome (attp40w) (Supplementary Fig. S2; Supplementary Table S2). These flies were viable, though we were unable to generate homozygotes for either CasRx or dCasRx, presumably due to high levels of ubiquitous ribonuclease expression. While homozygotes are desirable because, when outcrossed, all progeny would receive a copy of the transgene, we were still able to assess CasRx activity by maintaining these stocks as heterozygotes balanced over the chromosome Curly-of-Oster (CyO), which ensures a non-lethal expression level of CasRx while retaining the transgene (Supplementary Table S2). To genetically measure the efficacy of programmable RNA targeting, we targeted three genes known to produce visible phenotypes when expression is disrupted, including: *white* (*w*), *Notch* (*N*), and *yellow* (*y*) (Biessmann 1985; Sullivan and Sullivan 1975; Kidd, Kelley, and Young 1986; Dan L. Lindsley and E. H. Grell, n.d.). To target these genes with CasRx, we designed a gRNAarray for each gene driven by a ubiquitously expressed polymerase-3 U6 (U6:3) promoter (Kandul et al. 2019; Port et al. 2014), (Supplementary Fig. S2; Supplementary Table S2). Each array consisted of four transcript-targeting spacers (30 nt in length) each positioned between CasRx-specific direct repeats (36 nt in length) with a conserved 5'-AAAAC motif designed to be processed by either CasRx or dCasRx (Konermann et al. 2018) (Supplementary Fig. S2). For each gRNAarray, we site-specifically integrated the transgene at an available φC31 docking site located on the third

chromosome (site 8622) and established a homozygous transgenic line (Supplementary Fig. S2; Supplementary Table S2).

2.3.2 Programmable RNA targeting of endogenous target genes

To assess the efficacy of programmable RNA targeting by CasRx, we conducted bidirectional genetic crosses between homozygous gRNAarray (+/+; gRNAarray/gRNAarray) expressing flies crossed to either Ubiq-CasRx (Ubiq-CasRx/CyO; +/+) or Ubiq-dCasRx (Ubiq-dCasRx/CyO; +/+) expressing flies (Fig. 2A). When crossed to Ubiq-CasRx, we were able to obtain highly penetrant (100%) expected visible eye pigmentation disruption phenotypes exclusively in transheterozygotes (Ubiq-CasRx/+; gRNAarray /+) for *w* suggesting that CasRx exhibits programmable on-target RNA cleavage capabilities (Fig. 2B and C; Supplementary Table S10). However, while we expected Mendelian transheterozygote inheritance rates to be 50%, the recorded inheritance rates were significantly lower than expected (12.9%), suggesting some degree of unexpected toxicity leading to lethality (Fig. 2B; Supplementary Fig. S4; Supplementary Table S10). Moreover, when targeting *y* or *N*, Ubiq-CasRx transheterozygotes (Ubiq-CasRx/+; gRNAarray /+) were 100% lethal and did not develop beyond the second instar larval stage (Supplementary Fig. S5A and C). This was expected for *N* as there are many examples of lethal alleles for this gene (Micchelli and Perrimon 2006; Leonardi et al. 2011; Simón et al. 2014); however, mutations in *y* should be recessive viable with defective chitin pigmentation producing yellow cuticle phenotypes (Geyer, Green, and Corces 1990). It is worth noting that the corresponding Ubiq-dCasRx transheterozygote controls also showed less than 50% inheritance rates, though less severe than the Ubiq-CasRx transheterozygote experimental group. This indicates that the CasRx system may introduce toxicity when expressed at the

organismal level. Furthermore, we were unable to obtain visual phenotypes in transheterozygotes (Ubiq-dCasRx/+; gRNAarray /+) from our negative control crosses using all arrays tested, indicating that a catalytically active form of the ribonuclease is necessary for visual phenotypes to be observed (Fig. 2C). Taken together, these results indicate that the catalytically active form of the CasRx ribonuclease can generate phenotypes, although unexpected toxicity which resulted in lethality (only in the presence of the CasRx and the array) were also observed.

2.3.3 Tissue-Specific RNA Targeting by CasRx

Given the above toxicity when ubiquitously expressed, we next explored the efficiency of CasRx when expression was restricted to specific cell types and tissues. We leveraged the classical binary Gal4/UAS system which enables targeted gene expression (Brand and Perrimon 1993). To develop this system, we generated two transgenes using the UASt promoter (Brand and Perrimon 1993) to drive expression of either CasRx (UASt-CasRx) or dCasRx (UASt-dCasRx) as a negative control (Supplementary Fig. S2). These transgenes were integrated site-specifically using a φC31 docking site located on the second chromosome (site 8621), and these stocks were homozygous viable (Supplementary Fig. S2; Supplementary Table S2). To activate CasRx expression in specific tissues, we used available Gal4 driver lines that restricted expression to either the eye (GMR-Gal4) (Saj et al. 2010) or the wing and body (yellow-Gal4) (Massey et al. 2019); (Supplementary Table S2). These lines were crossed to the same homozygous gRNAarray lines described above targeting *w, y*, or *N* (Supplementary Fig. S2; Supplementary Table S2). To test this system, we performed a two-step genetic crossing scheme to generate F2 triple transheterozygotes (either UASt-CasRx/+; gRNAarray/Gal4 or UASt-dCasRx/+; gRNAarray/Gal4) (Fig. 1A). This consisted of initially crossing homozygous

gRNAarray (gRNAarray/gRNAarray) expressing flies to heterozygous, double-balanced UASt-CasRx (UASt-CasRx/Cyo; TM6/+) flies, or for the negative control, heterozygous, double-balanced UASt-dCasRx (UASt-dCasRx/Cyo; TM6/+) flies. The second step was to cross the F1 transheterozygote males expressing both a CasRx ribonuclease and the gRNAarray (UASt-CasRx/+; gRNAarray/TM6 or UASt-dCasRx/+; gRNAarray/TM6) to respective homozygous Gal4 driver lines to generate F2 triple transheterozygotes (UASt-CasRx/+; gRNAarray/Gal4 or UASt-dCasRx/+; gRNAarray/Gal4) to be scored for phenotypes (Fig. 1A).

From these crosses, our results indicated that tissue-specific expression of CasRx can indeed result in phenotypes, though this was also accompanied by tissue-specific cell death or organismal lethality, similar to previous observations of ubiquitous CasRx expression described above. For example, of the expected 25% Mendelian inheritance rates from the F1 cross between gRNAw (UASt-CasRx/+; gRNAw/TM6) and GMR-Gal4 (+/+; GMR-Gal4/GMR-Gal4), we observed survival of only 0.57% viable F2 triple transheterozygotes (UASt-CasRx/+; gRNAw/GMR-Gal4), all of which displayed severe eye specific pigmentation and morphology phenotypes (Fig. 1B andC; Supplementary Fig. S6; Supplementary Table S11). This gRNA*w* F2 triple transheterozygote inheritance rate was significantly lower than the corresponding negative control F2 triple transheterozygote (UASt-dCasRx/+; gRNAw/GMR-Gal4) inheritance rate, which was closer to the expected 25% Mendelian inheritance (27.6%) (Supplementary Fig. S6; Supplementary Table S11). Moreover, using the same Gal4 driver (GMR-Gal4), a significant difference in inheritance was also observed for *N* targeting, which resulted in 100% lethality in F2 triple transheterozygotes (UASt-CasRx/+; gRNAN/GMR-Gal4) compared to the 29.3% inheritance rate for the negative control F2 triple transheterozygotes (UASt-dCasRx/+; gRNAN/GMR-Gal4) (Fig. 1B and C; Supplementary Fig. S6; Supplementary Table S11).

Finally, when targeting *y* using the yellow-Gal4 driver (+/+; y-Gal4/y-Gal4), we observed marginal chitin pigment reduction as small patches of yellow cuticle on the thorax and abdomen in F2 triple transheterozygotes (UASt-CasRx/+; gRNAy/y-Gal4) (Fig. 1C, arrows) which is a phenotype that would be expected when *y* is disrupted. Similar to crosses described above, the F2 triple transheterozygote (UASt-CasRx/+; gRNAy/y-Gal4) inheritance was significantly lower (2.67%) when compared to the control F2 triple transheterozygote (UASt-dCasRx/+; gRNAy/y-Gal4) inheritance (25.2%), indicating a substantial degree of lethality (Fig. 1B and C; Supplementary Fig. S6; Supplementary Table S11). In all negative control crosses, phenotypes were not observed in F2 triple transheterozygotes (UASt-dCasRx/+; gRNAarray/Gal4) again indicating that functional catalytic residues of the HEPN motifs are necessary for generating phenotypes observed (Fig. 2B and C; Supplementary Table S11). Taken together, these results demonstrate that tissue-specific expression of CasRx using the classical Gal4/UAS approach can result in phenotypes. However, as seen in the ubiquitous binary approach above, toxicity and lethality phenotypes were also observed again limiting the utility of the system.

2.3.4 Criteria for CasRx-mediated phenotypes

Because CasRx on-target cleavage resulted in unexpected lethality we set out to determine the importance of target sequence availability to CasRx-mediated lethality. To do so, we opted to target a gene that is not natively expressed in flies. Therefore, we generated a green fluorescent protein (GFP) reporter assay to assess the necessity of a target sequence in CasRx-mediated lethality while simultaneously visualizing on-target transcript reduction. We designed a binary GFP reporter construct consisting of both a CasRx gRNAarray targeting GFP along with GFP expression driven by the broadly expressing OpIE2 promoter (gRNA*GFP*) (Fig. 3; Supplementary Fig. S2; Supplementary Table S2) (Pfeifer et al. 1997). We established a

homozygous transgenic line (+/+; gRNA*GFP*-OpIE2-GFP/gRNA*GFP*-OpIE2-GFP) by site-specifically integrating the construct at an available φC31 docking site located on the 3rd chromosome (site 8622) (Supplementary Fig. S2; Supplementary Table S2). To test for GFP transcript targeting, we performed bidirectional crosses between homozygous flies expressing gRNA*GFP* (+/+; gRNA*GFP*-OpIE2-GFP/gRNA*GFP*-OpIE2-GFP) to heterozygous Ubiq-CasRx-expressing flies (Ubiq-CasRx/CyO; +/+) or heterozygous Ubiq-dCasRx-expressing flies (Ubiq-dCasRx/CyO; +/+) as a negative control (Fig. 3A). With this assay, we observed 100% larval lethality for F1 transheterozygotes (Ubiq-CasRx/+; gRNA*GFP*-OpIE2-GFP/+), while larval lethality was eliminated in F1 progeny that did not inherit Ubiq-CasRx (Cyo/+; gRNAGFP-OpIE2-GFP/+) in addition to the transheterozygote controls (Ubiq-dCasRx/+; gRNA*GFP*-OpIE2-GFP/+). Lethality was also observed regardless of the maternal or paternal deposition of CasRx (Fig. 3B; Supplementary Table S10). Given that GFP expression was also visible in larvae, we monitored the development of the F1 progeny and observed that Ubiq-CasRx transheterozygotes survived only to the first instar developmental stage, but not beyond (Supplementary Fig. S5). Given this survival, we imaged first instar transheterozygote (Ubiq-CasRx/+; gRNA*GFP*-OpIE2-GFP/+) larvae and observed near-complete reduction in GFP expression for Ubiq-CasRx transheterozygote larvae as compared to Ubiq-dCasRx transheterozygote (Ubiq-dCasRx/+; gRNA*GFP*-OpIE2-GFP/+) control larvae indicating robust CasRx mediated target transcript (GFP) reduction (Fig. 3C). Taken together, these results suggest that CasRx possesses programmable RNA-targeting activity, and the lethality is dependent upon the availability of a guide RNA and a target sequence as well as enzymatic RNA cleavage mediated by the positively charged residues of CasRx HEPN domains.

2.3.5 Quantification of CasRx-mediated on/off-target activity

We next aimed to quantify both the on- and potential off-target transcript reduction rates. To do this, we analyzed all gRNAarray target genes from our binary crosses producing either highly penetrant, visible phenotypes (*w*) or lethal phenotypes (*N*, *y*, and *GFP*) (Supplementary Table S12). To do so, we implemented whole-transcriptome RNAseq analysis comparing F1 Ubiq-CasRx transheterozygotes (Ubiq-CasRx/+; gRNAarray /+) to control F1 Ubiq-dCasRx transheterozygotes (Ubiq-dCasRx/+; gRNAarray /+) (Fig. 2A, red box; Fig. 3A, red box; Supplementary Table S12). Using the available transcriptome data of *D. melanogaster* (modENCODE) (Graveley et al. 2011), we extracted total RNA stages of development when high transcript expression levels were expected for each target gene with the exception of *GFP*, where we sequenced first instar larvae (Supplementary Fig. S3; Supplementary Table S12). In total, we analyzed 24 samples (Supplementary Table S12). From our bioinformatic analysis, we found reduced target transcript expression (Fig. 4A and B). For example, of the four target genes, CasRx was able to target and significantly reduce (1.5%–2.9%; Supplementary Table S13) the target transcript expression of three genes compared with dCasRx controls *N*, *y*, and *GFP* (Fig. 4B; Supplementary Tables S3, S4, S5, S6, S7, S8, S9). Although we did not observe significant transcript reduction targeting *w* we did consistently observe relative expression reduction by comparing Ubiq-CasRx samples to Ubiq-dCasRx controls, indicating some degree of on-target reduction which likely contributes to the phenotypes observed (Fig. 4B; Supplementary Tables S3, S4, S5, S6, S7, S8, S9). We also quantified the number of genes with significantly misexpressed transcripts by comparing Ubiq-CasRx to Ubiq-dCasRx using DESeq250 (Fig. 4A, red dots; Supplementary Tables S5, S6, S7, S8, S9).

Across all gene targets, we observed some evidence of potential off-target activity, which we define as significantly misexpressed genes between CasRx and dCasRx samples. The observed potential off-target activity was demonstrated by significant changes in the gene-expression levels of numerous non-target transcripts. The number of significantly differentially expressed non-target transcripts in each group are: 253 (*w*), 300 (*N*), 41 (*y*), and 5,880 (*GFP*), representing 1.4% (*w*), 1.7% (*N*), 0.23% (*y*), and 33% (*GFP*) of the total transcripts (Fig. 4A and C; Supplementary Tables S5, S6, S7, S8, S9). Taking a closer look at the gene-expression profiles of the four gene targets, we found that a total of 6,082 transcripts (out of 17,779) displayed significant expression level changes in at least one of the six CasRx-expressing groups compared to their corresponding dCasRx-expressing control group (Supplementary Tables S5, S6, S7, S8, S9). Among the 6,082 misexpressed transcripts, 5,722 transcripts are affected by only one of the four genes targeted when CasRx is present, 334 transcripts are affected by two gene targets, 20 transcripts are affected by three gene targets, and 6 transcripts are affected by four gene targets simultaneously (Supplementary Tables S5, S6, S7, S8, S9). As targeting exogenously introduced GFP induces 33% of the endogenous transcripts to be misexpressed, suggesting that at the organismal level CasRx system has the risk of resulting in high off-target activity and collateral tissue damages resulting in observed lethality. This quantitative analysis of CasRx-mediated transcript reduction provides evidence of CasRx ribonuclease capabilities in flies, while also identifying potential off-target effects resulting in significantly misexpressed non-target genes. That said, it should be emphasized that these off-target results should be taken with a grain of salt as we are unable to tease apart pleiotropic effects that could also contribute to misexpression of non-target transcripts and therefore a more comprehensive characterization of CasRx mediated off-targets should be conducted in the future.

2.4 Discussion

Our results demonstrate that CasRx has some potential for programmable RNA targeting in flies, as we did observe some expected phenotypes for each target transcript, including *GFP* (lethality; reduction in *GFP* expression), *N* (lethality), *y* (lethality; yellow patches on cuticle and thorax) and *w* (white eyes and necrosis in eyes for Gal4) . Specifically, RNA targeting was demonstrated with ubiquitous, inducible, and tissue-specific CasRx expression systems against native and synthetic RNA targets, which are prerequisites for enabling comprehensive studies of gene function. However, we did also consistently observe both cellular toxicity from the ubiquitous expression of CasRx and dCasRx as we could not generate homozygous strains for either, and unexpected lethality and tissue necrosis, presumably due collateral off target effects which have been a feature previously observed for many CRISPR ribonucleases including CasRx (Abudayyeh et al. 2016; Konermann et al. 2018; Smargon et al. 2017; East-Seletsky et al. 2016; Yan et al. 2018; Meeske, Nakandakari-Higa, and Marraffini 2019). Nevertheless, in both bidirectional and Gal4/UAS crosses, we were able to obtain visible phenotypes as well as quantitative evidence (e.g., RNAseq data demonstrating a reduction in target gene expression) indicating that the CasRx is capable of targeting and degrading target RNA in flies. It is interesting to note that for one of the targeted genes (*w*), while the observed phenotype indicated consistent on-target transcript reduction, DESeq2 analysis did not reveal significant on-target reduction, which may be due to the timing of sample collection for RNAseq since expression levels of these genes vary over development.

Notwithstanding, we were able to obtain expected visual phenotypes in addition to significant on-target CasRx mediated transcript reduction for three of the targeted genes: *y*, *N*, and *GFP.* Interestingly, transheterozygotes (Ubiq-CasRx/+; gRNAarray /+) for *y*, *N*, and *GFP*

also had many other misexpressed non-target genes, possibly indicating that target cleavage results in increased collateral off-target activity that is detrimental to development as these individuals were adult lethal. For example, targeting *GFP*, a nonessential gene, produced the largest number of misexpressed genes as well as the most significant fold change in expression compared to all other gene targets analyzed. Additionally, because *Gadd45*, a gene involved in cellular arrest and apoptosis in *D. melanogaster (Peretz, Bakhrat, and Abdu 2007)*, was also significantly misexpressed in four samples (*w*, *N*, *y*, and *GFP*) (Supplementary Tables S5, S6, S7, S8, S9), it is possible that CasRx cleavage may result in an increased level of misexpressed genes leading to lethality or cellular apoptosis. Moreover, for the off target analysis for most of the target genes less than 300 (1.7%) other genes were misexpressed, however for GFP we found 5880 (33%) of genes misexpressed and it remains unclear whether this is a result of guide specific off-target, pleiotropic effects, or simply bystander cleavage (i.e. collateral off targeting).

Through this study, we identified two main factors contributing to CasRx-mediated lethality: (i) the catalytic activity of the CasRx HEPN domains, as lethality and tissue necrosis phenotypes were eliminated in dCasRx compared to CasRx crosses, and (ii) the presence of the guide RNA and target transcript resulting in on-target cleavage, as lethality was only observed when crossing Ubiq-CasRx-expressing flies to gRNAGFP-expressing flies. These results recapitulate previous mechanistic analyses of CasRx and other Cas13 ribonucleases, demonstrating that collateral off-target activity following targeted transcript cleavage is a native feature of Cas13 ribonucleases (Abudayyeh et al. 2016; Konermann et al. 2018; Smargon et al. 2017; East-Seletsky et al. 2016; Yan et al. 2018; Meeske, Nakandakari-Higa, and Marraffini 2019). While this feature may not be desirable for generating tools for targeting specific transcripts of genes, this may be useful for generating sensors that get activated in response to a

target transcript (e.g., viral target) such as diagnostic tools that could alert the presence of a nucleic acid target and activate a marker, or even organismal lethality acting as an in vivo ribonucleic acid sensor.

Taken together, further optimization will be required to increase the CasRx on-target cleavage rates and decrease cellular toxicity and off-target effects, but this is the first demonstration of a genetically encoded programmable RNA-targeting Cas system in *D. melanogaster*. In the future, optimization of the strength and timing of CasRx expression could mitigate some of the off-target-associated lethality in this system. Stricter and more tunable regulation of CasRx expression may also improve phenotype penetrance as it appears to be dosage dependent in both our system and other CasRx systems (Kushawah et al. 2020). For example, the phenotypes of *y* varied by expression, with ubiquitous expression of CasRx resulting in a Ubiq-CasRx/+; gRNAy /+ lethal phenotype and embryo and wing and body specific expression mitigated lethality phenotype seen in Ubiq-CasRx expression. Optimization of gRNA design may further improve these systems as CasRx gRNAs have been shown to have variable knockdown efficiency (Kushawah et al. 2020; Wessels et al. 2019). Nevertheless, this is an important first step towards making transcriptome engineering a viable *in vivo* technology and provides a foundation for future experiments to mitigate the off-target and toxic attributes of the enzyme to make a new, viable tool in the expanding gene-editing toolbox.

2.5 Methods

2.5.1 Design and assembly of constructs

To select the CasRx target sites, target genes were analyzed to identify 30-nucleotide (nt) regions that had no poly-U stretches greater than four base pairs, had GC base content between 30% and 70%, and were not predicted to form strong RNA hairpin structures. Care was also taken to select target sites in RNA regions that were predicted to be accessible, such as regions without predicted RNA secondary or tertiary structure (Supplementary Fig. S1). All RNA folding/hairpin analysis was performed using mFold (Zuker 2003). For transgenic gRNA arrays, four targets per gene were selected to ensure efficient targeting. We assembled four CasRx- and catalytically inactive negative control (dCasRx)–expressing constructs under the control of one of two promoters: Ubiquitin-63E (Ubiq) or the original Upstream Activation Sequence (UAS) promoter developed in Brand and Perrimon (Brand and Perrimon 1993) (Ubiq-CasRx, Ubiq-dCasRx, UASt-CasRx, UASt-dCasRx) using the Gibson enzymatic assembly method (Gibson et al. 2009). A base vector (Addgene plasmid 112686) containing piggyBac and an attB-docking site, the Ubiq promoter fragment, SpCas9-T2A-GFP, and the Opie2-dsRed transformation marker was used as a template to build all four constructs (Kandul et al. 2019). To assemble constructs OA-1050E (Addgene plasmid 132416, Ubiq-CasRx), and OA-1050R (Addgene plasmid 132417, Ubiq-dCasRx), the SpCas9-T2A-GFP fragment was removed from the base vector by cutting with restriction enzymes SwaI and PacI and replaced with CasRx and dCasRx fragments amplified with primers 1050E.C3 and 1050E.C4 (Supplementary Table S1) from constructs pNLS-RfxCas13d-NLS-HA (pCasRx) and pNLS-dRfxCas13d-NLS-HA (pdCasRx),4 respectively. To assemble constructs OA-1050L (Addgene plasmid 132418, UASt-CasRx) and OA-1050S (Addgene plasmid 132419, UASt-dCasRx), the base vector

described above (Addgene plasmid 112686) was digested with restriction enzymes NotI and PacI to remove the Ubiq promoter and SpCas9-T2A-GFP fragments. Then the UASt promoter fragment and CasRx or dCasRx fragments were cloned in. The UASt promoter fragment was amplified from plasmid pJFRC8132 using primers 1041.C9 and 1041.C11 (Supplementary Table S1). The CasRx and dCasRx fragments were amplified with primers 1050L.C1 and 1050E.C4 (Supplementary Table S1) from constructs pCasRx and pdCasRx, respectively.

We designed four constructs, each carrying a four-gRNA array: OA-1050G (Addgene plasmid 132420), OA-1050I (Addgene plasmid 132421), OA-1050J (Addgene plasmid 133304), and OA-1050Z4 (Addgene plasmid 132425), targeting transcripts of *white, Notch, GFP,* and *yellow*, respectively. To generate a base plasmid, OA-1043, which was used to build all constructs carrying the four-gRNA array, Addgene plasmid 112688 containing the *miniwhite* gene as a marker, an attB-docking site, a *D. melanogaster* polymerase-3 U6 (U6:3) promoter fragment, and a gRNA fragment with a target, scaffold, and terminator sequence was digested with restriction enzymes AscI and XbaI to remove the U6:3 promoter and gRNA fragments. Then, the U6:3 promoter fragment was amplified from the same Addgene plasmid 112688 with primers 1043.C1 and 1043.C23 (Supplementary Table S1) and was cloned back using Gibson enzymatic assembly. To generate constructs OA-1050G, OA-1050I, and OA-1050Z4, plasmid OA-1043 was digested with restriction enzymes PstI and NotI. Then, a fragment that contained arrays of four tandem variable gRNAs (comprised of a 36-nt direct repeat [DR] and a 30-nt spacer) corresponding to different target genes followed by an extra DR and a seven-thymine terminator was synthesized and subcloned into the digested backbone using Gene Synthesis (GenScript USA, Inc.). To generate construct OA-1050J, a fragment containing arrays of four tandem variable gRNAs targeting *GFP* with an extra DR and a seven-thymine terminator

followed by the OpIE2-GFP marker was synthesized and subcloned into the above digested OA-1043 backbone using Gene Synthesis (GenScript USA, Inc.). We have also made all plasmids and sequence maps available for download and/or order from Addgene using the identification numbers listed in Supplementary Fig. S2 and Supplementary Table S2.

2.5.2 Fly genetics and imaging

Flies were maintained under standard conditions at 26°C. Embryo injections were performed at Rainbow Transgenic Flies, Inc. All CasRx and dCasRx-expressing lines were generated by site-specifically integrating our constructs at available φC31 integration sites on the 2nd chromosome (sites 8621 [UAS/-(d)CasRx] and attp40w [Ubiq-(d)CasRx]). Homozygous lines were established for UASt-CasRx and UASt-dCasRx, and heterozygous balanced lines were established for Ubiq-CasRx and Ubiq-dCasRx (over Curly of Oster: CyO). All gRNAarray-expressing lines were generated by site-specific integrating constructs at an available φC31 integration site on the third chromosome (site 8622). Homozygous lines were established for all gRNAarray-expressing flies.

To genetically assess the efficiency of CasRx ribonuclease activity, we bidirectionally crossed Ubiq-CasRx- and Ubiq-dCasRx-expressing lines to gRNAarray-expressing lines at 26°C. F1 transheterozygotes were scored for the inheritance and penetrance of observable phenotypes. Embryo, larvae, and pupae counts were preceded by crossing male Ubiq-CasRx- or Ubiq-dCasRx-expressing flies to female gRNAarray-expressing flies. Flies were incubated at 26°C for 48 h with yeast to induce embryo laying. Flies were then transferred to embryo collection chambers containing yeast-smeared grape-juice plates and were incubated at 26°C overnight (16 h). The grape-juice plates were then removed, the embryos were counted, and the

grape-juice plates were incubated for 24 h at 26°C. Total larvae and transheterozygote larvae were then counted, and the grape-juice plates were transferred to jars and incubated at 26°C. Once all larvae reached the pupal stage, total, and transhet pupae were counted. Finally, total adult flies and total adult transheterozygotes were counted 20 days post initial lay. Each genetic cross was set using 5 male and 10 female (paternal CasRx) or 4 male and 8 female (maternal CasRx) flies in triplicate.

To investigate the tissue-specific activity of CasRx, we designed a two-step crossing scheme to generate F2 triple transheterozygotes (Fig. 1A). First, we crossed double-balanced UASt-CasRx- or UASt-dCasRx-expressing flies (male) to homozygous gRNAarray-expressing flies (female) to generate F1 transheterozygote males carrying the TM6-balancer chromosome. The F1 transheterozygote males carrying TM6 were then crossed with a Gal4–driver-expressing line. F2 triple transheterozygous inheritance and phenotype penetrance was scored based on visible phenotypes manifesting in flies F2 flies with red eyes, a lack of the TM6 balancer chromosome, and red fluorescent protein (dsRed) expression. Marked by the presence of dsRed (for UASt-CasRx or UASt-dCasRx), red eyes (to mark the gRNA), and the lack of TM6, F2 triple transheterozygotes' inheritance and phenotype penetrance was scored. Each cross was set using 1 female and 10 male flies in triplicate. The flies were imaged on the Leica M165FC fluorescent stereomicroscope equipped with a Leica DMC4500 color camera. Image stacks of adult flies were taken in Leica Application Suite X (LAS X) and compiled in Helicon Focus 7. Stacked images were then cropped and edited in Adobe Photoshop CC 2018.

2.5.3 Illumina RNA sequencing

The total RNA was extracted from F1 transheterozygous flies at different developmental stages based on the expression data available through modENCODE analysis (Supplementary Fig. S3). For guide RNA targeting the *white* gene (gRNA*w*) flies, transheterozygous adult heads were removed one day after emerging and were frozen at −80°C. For guide RNA targeting the *yellow* gene (gRNA*y*) flies, the flies were incubated in vials for 48 h with yeast to induce embryo laying. The flies were then transferred to embryo collection chambers containing yeast-smeared grape-juice plates and incubated at 26°C for 3 h. The flies were then removed, and the embryos were aged on the grape-juice plates (gRNA*y* = 17 h; 17–20 h total). The embryos were removed from the grape-juice plates, washed with deionized H2O, and frozen at −80°C. The guide RNA targeting the *Notch* gene (gRNA*N*) and guide RNA targeting the *Green Fluorescent Protein* gene (gRNAGFP) flies were incubated in vials with yeast for 48 h to induce embryo laying. The flies were then transferred to a new vial and allowed to lay overnight (16 h). The adults were removed, and the vials were incubated at 26°C for 24 h. Transheterozygote first instar larvae were then picked (based on dsRed expression) and frozen at −80°C. For all samples, the total RNA was extracted using Qiagen RNeasy Mini Kit (Qiagen 74104). Following extraction, the total RNA was treated with Invitrogen TurboTM DNase (Invitrogen AM2238). The RNA concentration was analyzed using a Nanodrop One*C* UV-vis spectrophotometer (ThermoFisher ND-ONEC-W). The RNA integrity was assessed using an RNA 6000 Pico Kit for Bioanalyzer (Agilent Technologies 5067-1513). The RNA-seq libraries were constructed using NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB E7770) following the manufacturer's instructions.33 The libraries were sequenced on an Illumina HiSeq2500 in single read mode with a read length of 50 nt and a sequencing depth of 20 million reads per library following the

manufacturer's instructions. Base calls were performed with RTA 1.18.64 followed by conversion to FASTQ with bcl2fastq 1.8.4. We sequenced and analyzed three replicates for all CasRx and dCasRx samples. In total, we sequenced and analyzed 24 samples: 12 CasRx experimental samples and 12 dCasRx control samples. All raw sequencing data can be accessed at the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA; submission ID SUB6818910 [BioProject PRJNA600654]).

2.5.4 Bioinformatics: Quantification and differential expression analysis

Reads were mapped to the *D. melanogaster* genome (BDGP release 6; GenBank accession GCA_000001215.4) supplemented with cDNA sequences of CasRx and GFP using the default parameters of STAR aligner (Dobin et al. 2013) with the addition of the "–outFilterType BySJout" filtering option and "–sjdbGTFfile Drosophila_melanogaster.BDGP6.22.97.gtf" gene transfer format (GTF) file downloaded from ENSEMBL. The expression levels were determined with featureCounts (Liao, Smyth, and Shi 2014) using "-t exon -g gene_id -M -O –fraction" options. The raw transcript counts were normalized to transcripts per million (TPM), which were calculated from count data using the "addTpmFpkmToFeatureCounts.pl" Perl script (see Supplementary Material, Supplementary File S1). The raw count and TPM data are available in Supplementary Tables S3 and S4. To further explore CasRx-induced differential gene expression profiles, we used the maximum *a posteriori* method with the original shrinkage estimator in the DESeq2 pipeline to estimate the transcript logarithmic fold change (Geyer, Green, and Corces 1990). The Wald test with the Benjamini-Hochberg correction was used for statistical inference. The analysis script can be found in the Supplementary Material (Supplementary File S2), and the analyzed results are in Supplementary Tables S5, S6, S7, S8, S9. Per the DESeq2 analysis

requirements, some values are shown as NA (not applicable) for the following reasons: (1) if all samples for a given transcripts have 0 transcript counts, this transcript's baseMean will be 0 and its logarithmic fold change, *P* value, and padj will be set to NA; (2) if one replicate of a transcript is an outlier with an extreme count (detected by Cook's distance), this transcript's *P* value and padj will be set to NA; or (3) if a transcript is found to have a low mean normalized count after automatic independent filtering, this transcript's padj will be set to NA.

2.6 Acknowledgements

This chapter, in full, is a reprint of the material as it appears in The CRISPR Journal. Anna B. Buchman*, Daniel J Brogan*, Ruichen Sun*, Ting Yang, Igor Antoshechkin, Patrick D. Hsu, Omar S. Akbari. "Programmable RNA-targeting using CasRx in flies". The CRISPR Journal. (2020). [* = Equal contribution]. The dissertation/thesis author was the co-first-author and primary investigator of this paper.

2.7 Figures

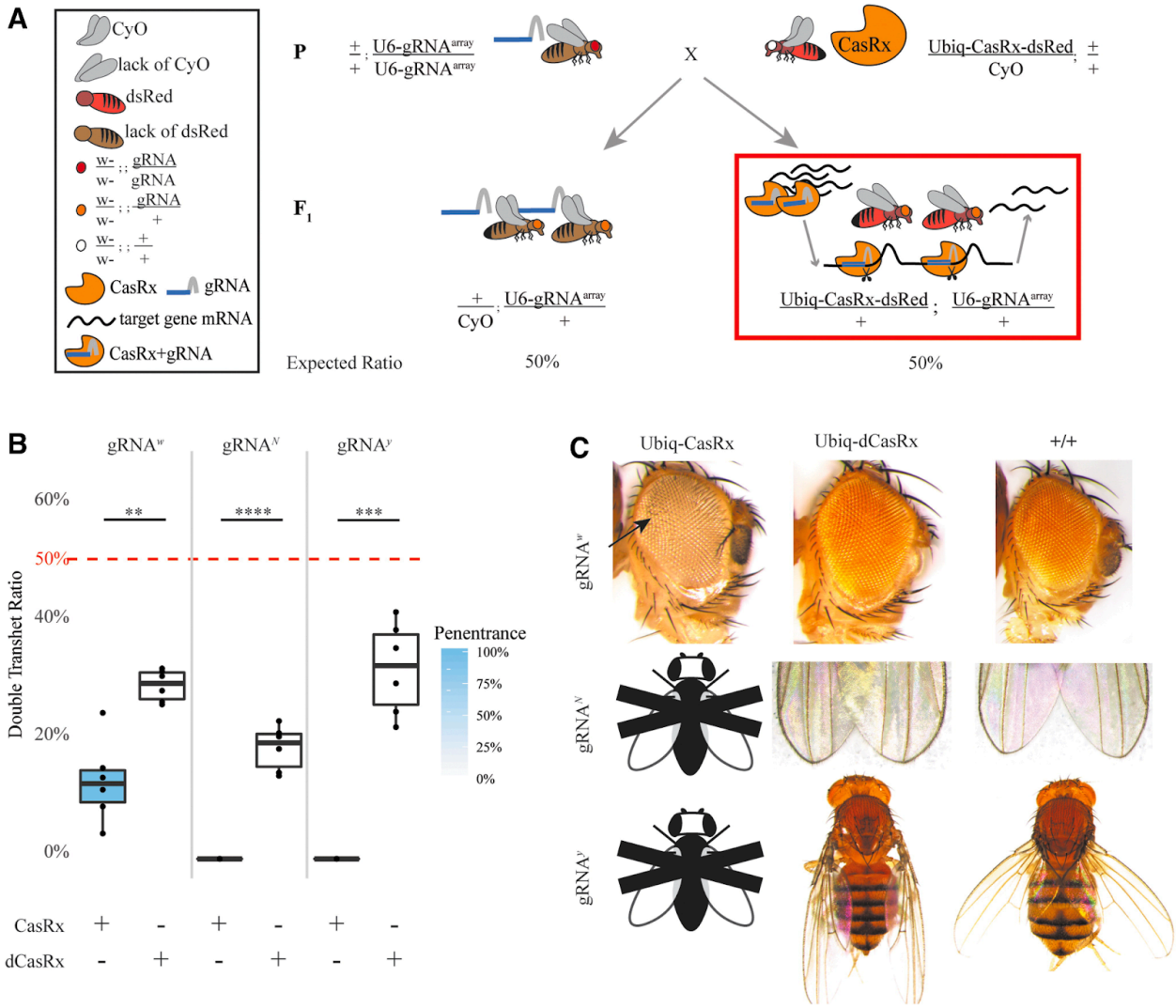


Figure 2.1: Genetic assessment of programmable CasRx-mediated transcript reduction in flies. (A) Representative genetic crossing schematic to generate transheterozygotes. (B) Inheritance and penetrance rates of transheterozygous flies inheriting both Ubiq-CasRx (or Ubiq-dCasRx) and gRNAarray corresponding to the red box in panel A. Phenotype penetrance rate is depicted by blue shading in the box plot. Significant differences in inheritance between CasRx and dCasRx groups were observed in all three groups (*P* values: gRNA*w* = 0.00135; gRNA*N* = 0.00006; gRNA*y* = 0.00016). (C) Brightfield images of transheterozygous flies with representative phenotypes for each cross. Corresponding genotype for each image is dictated by the combination of constructs on top of the columns and the side of the rows. Arrows point to tissue necrosis in the eye. Black and white fly with "X" represents lethality phenotype where no transheterozygote adults emerged. dCasRx, catalytically inactive negative control CasRx.

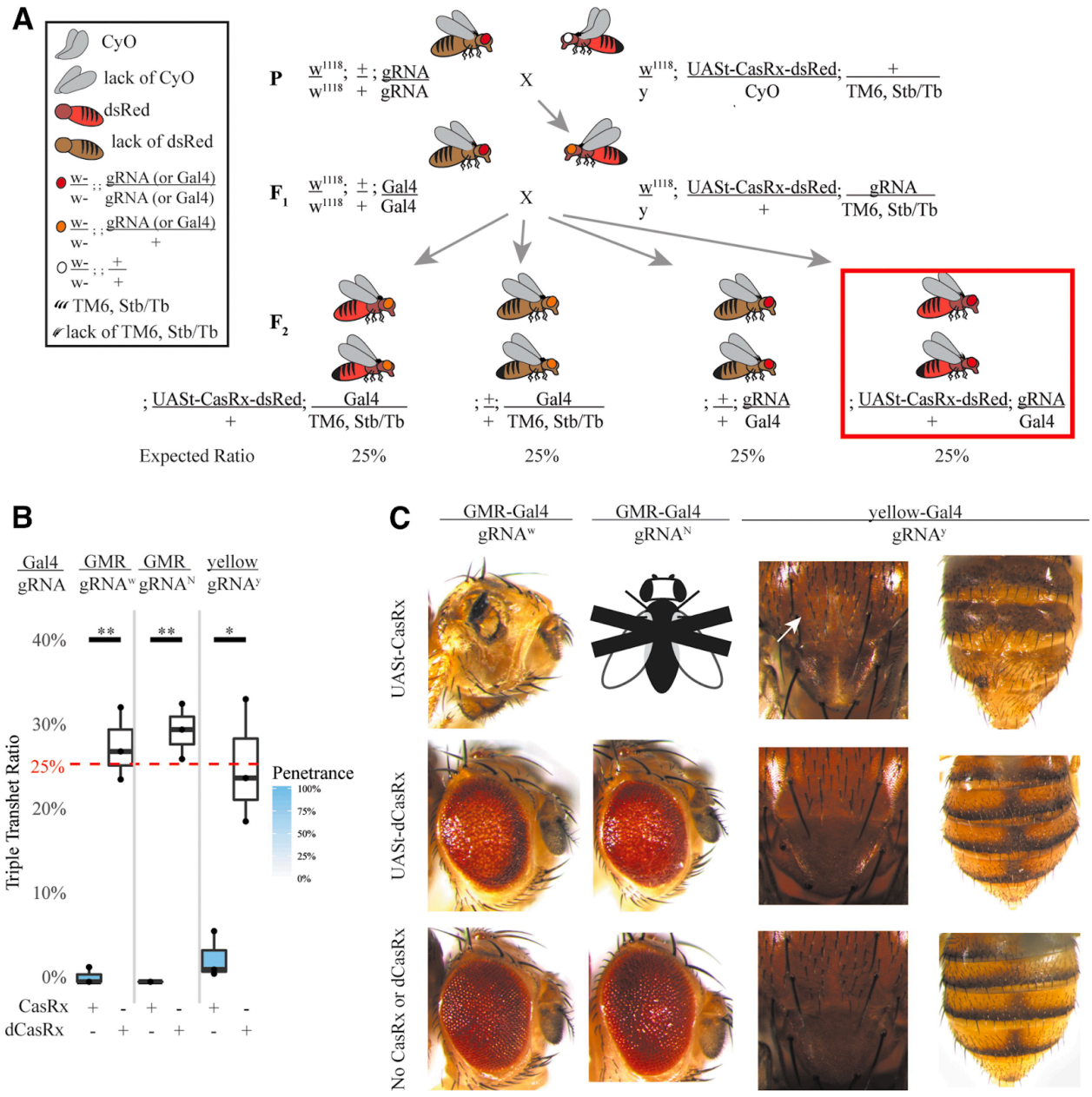


Figure 2.2: CasRx-mediated target transcript reduction in restricted tissue types using the binary Gal4/UAS system. (A) Representative genetic crossing schematic. (B) Inheritance rates of triple transheterozygous flies inheriting three transgenes (UASt-CasRx or UASt-dCasRx, gRNAarray, and Gal4-driver), corresponding to flies highlighted in the red box in panel A. Significant differences in inheritance between CasRx and dCasRx groups were observed in all three gene targets (gRNA*w*, $P = 0.00595$; gRNA*N*, $P = 0.00402$; gRNA*y*, $P = 0.02205$). (C) Phenotypes of the triple transheterozygous flies. The white arrow identifies chitin pigment reduction in the thorax resulting from *y* targeting. Black and white fly with "X" represents a lethal phenotype with no live adults able to be scored or imaged. CasRx, Cas ribonuclease; gRNAarray, guide RNA array; gRNA*N*, guide RNA targeting the *Notch* gene; gRNA*w*, guide RNA targeting the *white* gene; gRNA*y*, guide RNA targeting the *yellow* gene.

Figure 2.3: Robust CasRx-mediated reduction of GFP. (A) Representative bidirectional genetic crossing schematic. (B) Box plot of transheterozygote inheritance rates resulting from bidirectional crosses between Ubiq-CasRx (or Ubiq-dCasRx) and gRNA*GFP*-OpIE2-GFP flies. (C) Images of F1 larvae from paternal crosses demonstrating significant reduction in GFP expression for transheterozygous larvae expressing both Ubiq-CasRx and gRNA*GFP*-OpIE2-GFP compared to control transheterozygotes expressing Ubiq-dCasRx and gRNA*GFP*-OpIE2-GFP or without expressing a CasRx protein. (Left-right) Ubiq-CasRx/gRNAGFP transheterozygous larvae, heterozygous gRNA*GFP* larvae from Ubiq-CasRx cross, Ubiq-dCasRx/gRNAGFP transheterozygous larvae, heterozygous gRNA*GFP* larvae from Ubiq-dCasRx cross. CyO, ; dsRed, red fluorescent protein; GFP, green fluorescent protein; M, maternal inheritance of CasRx; P, paternal inheritance of CasRx; RFP, red fluorescent protein.

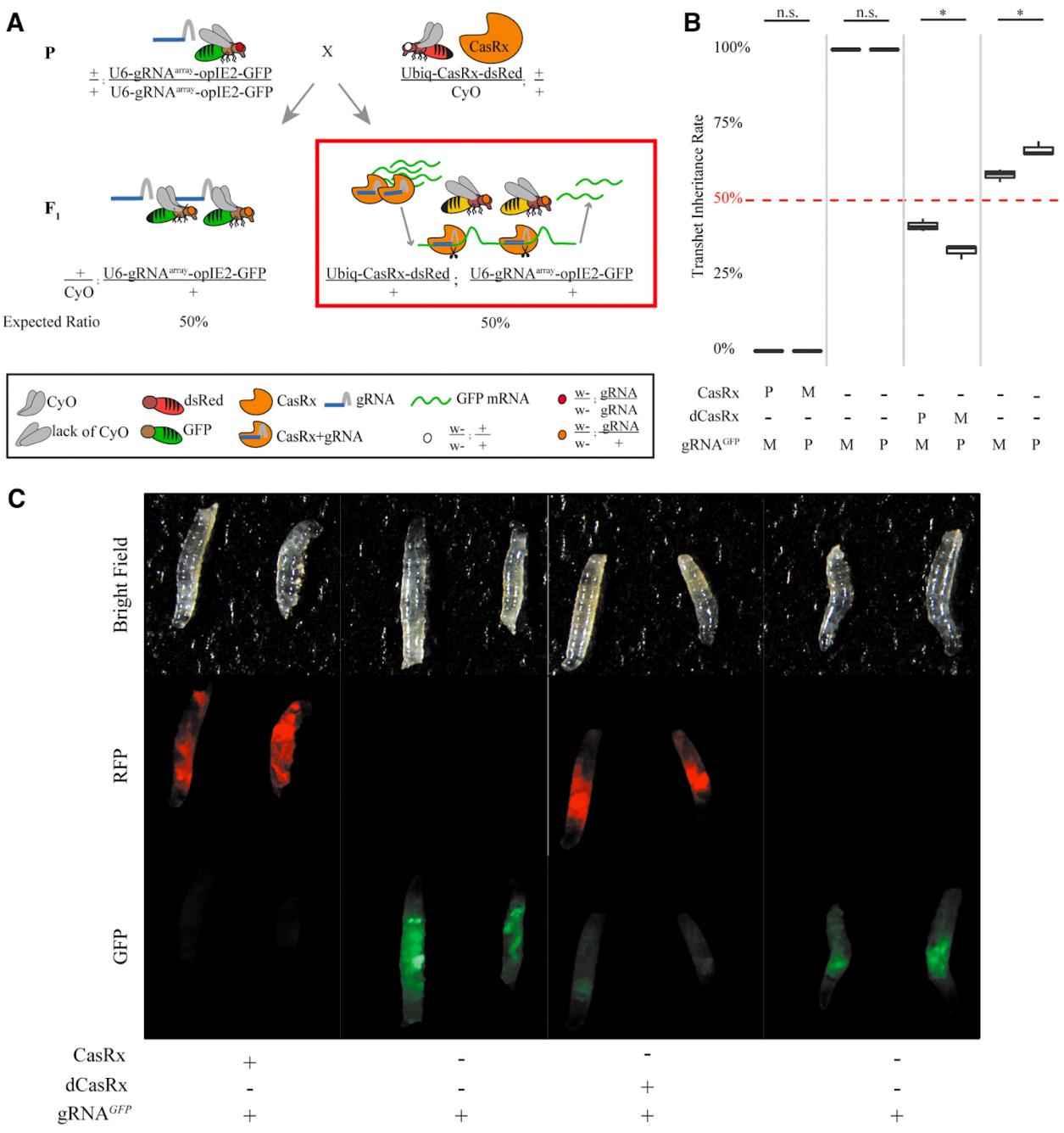
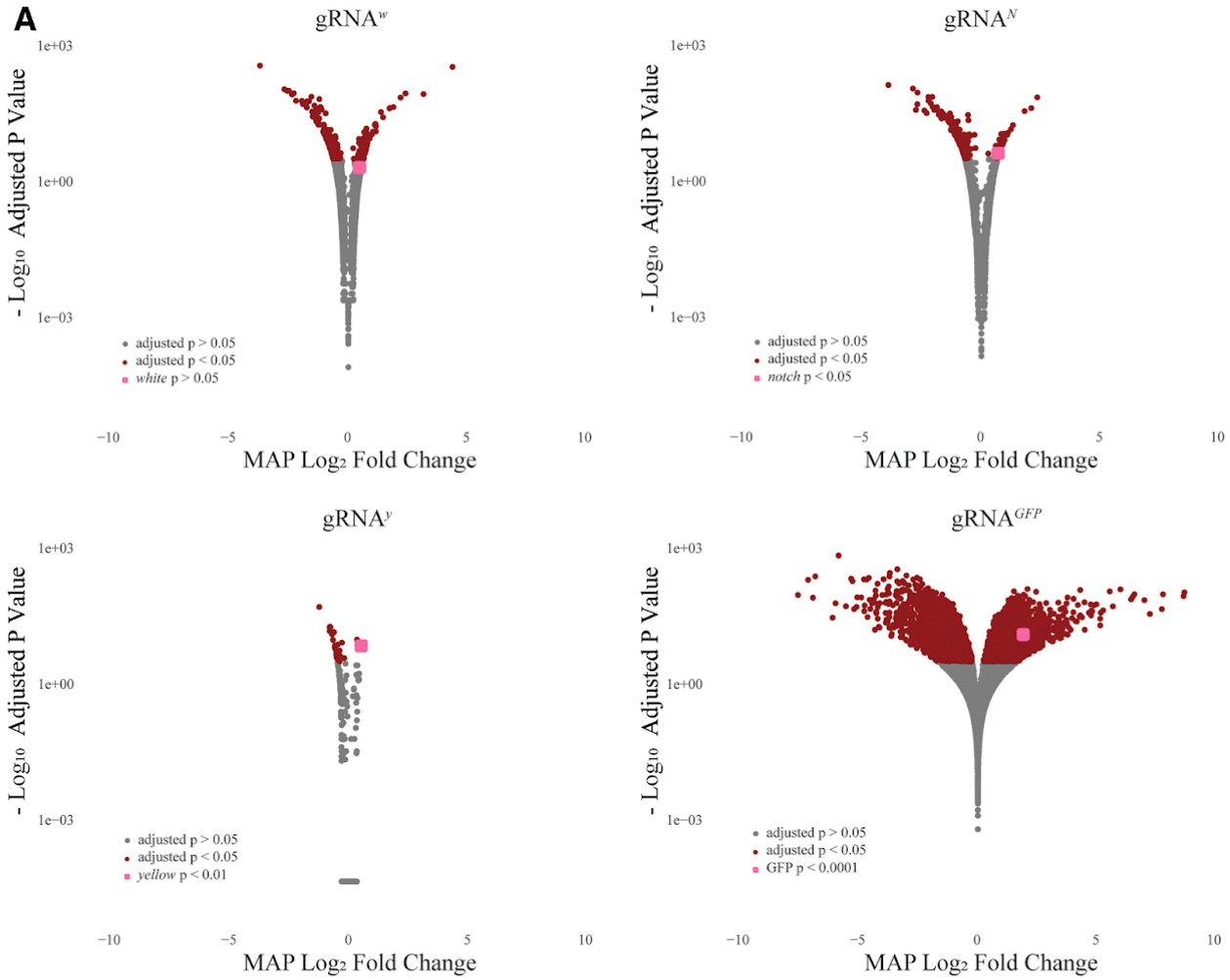
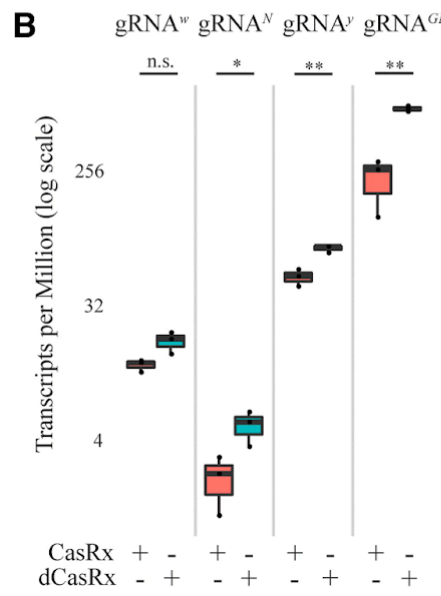
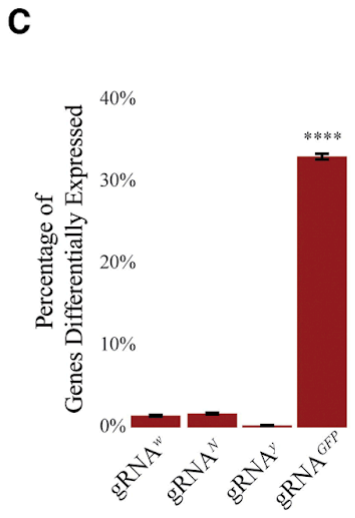

A
P
X
CasRx
Ubiq-CasRx-dsRed
CyO
F1
Expected Ratio
50%
50%
Ubiq-CasRx-dsRed
U6-gRNAarray-opIE2-GFP
CyO
lack of CyO
dsRed
GFP
CasRx
gRNA
GFP mRNA
CasRx+gRNA
B
n.s.
n.s.
*
*
100%
75%
50%
25%
0%
Transhet Inheritance Rate
CasRx
dCasRx
gRNAGFP
C
Bright Field
RFP
GFP
CasRx
dCasRx
gRNAGFP

Figure 2.4: Quantification of potential CasRx-mediated on/off target activity. (A) Maximum *a posteriori* estimates for the logarithmic fold change of transcripts. DESeq2 pipeline was used for estimating shrunken Maximum *a posteriori* logarithmic fold changes. Wald test with Benjamini-Hochberg correction was used for statistical inference. Grey dots represent transcripts not significantly differentially expressed between Ubiq-CasRx and Ubiq-dCasRx group ($P > 0.05$). Red dots represent transcripts significantly differentially expressed between CasRx and dCasRx group ($P < 0.05$). Pink dot identifies the respective CasRx target gene for each analysis (*P*-value indicated in th/by inset). (B) Transcript expression levels (TPM) of transcripts targeted with CasRx or dCasRx. Student's *t*-test was used to calculate significance (w: $P = 0.07$; N: $P = 0.04$; y: $P = 0.006$; GFP: $P = 0.008$). (C) Percentage of transcripts significantly differentially expressed resulting from CasRx cleavage. A pairwise two-sample test for independent proportions with Benjamini-Hochberg correction was used to calculate significance. LFC, logarithmic fold change; MAP, Maximum *a posteriori*.


A
gRNA^w
gRNA^N
gRNA^y
gRNA^GFP
- Log10 Adjusted P Value
1e+03
1e+00
1e−03
adjusted p > 0.05
adjusted p < 0.05
white p > 0.05
notch p < 0.05
yellow p < 0.01
GFP p < 0.0001
−10
−5
0
5
10
MAP Log2 Fold Change


B
gRNA^w gRNA^N gRNA^y gRNA^GFP
n.s.
*
**
**
Transcripts per Million (log scale)
256
32
4
CasRx + - + - + - + -
dCasRx - + - + - + - +


C
Percentage of Genes Differentially Expressed
40%
30%
20%
10%
0%
****
gRNA^w
gRNA^N
gRNA^y
gRNA^GFP

2.8 Supplemental figures

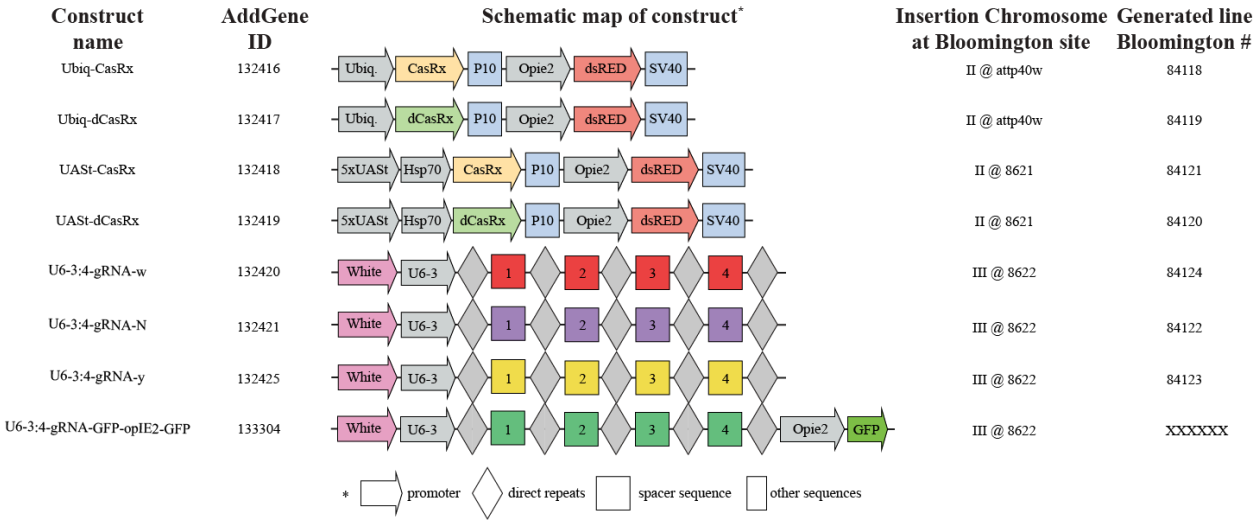


Figure S2.1: Schematic representation of constructs generated for this study. All constructs used in this study are depicted here along with addgene ID, insertion site, and Bloomington stock number.

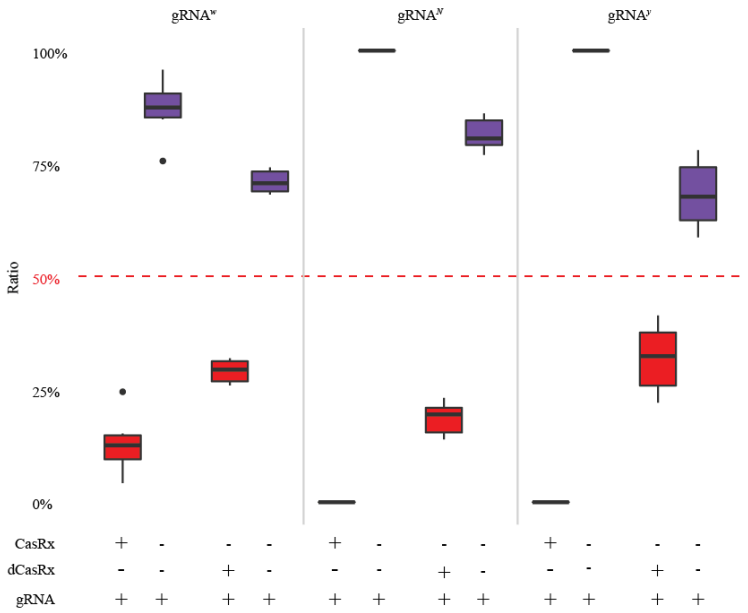


Figure S2.2: Complete inheritance data for bidirectional crosses. Complete inheritance plot for the bidirectional crosses featured in Fig 1. The plot includes all genotypes scored in all crosses between either Ubiq-CasRx or Ubiq-dCasRx and a respective gRNA^{array}. In all crosses, gRNA^{array}-only inheritance is dramatically higher than transheterozygote inheritance rates, including for Ubiq-dCasRx crosses.

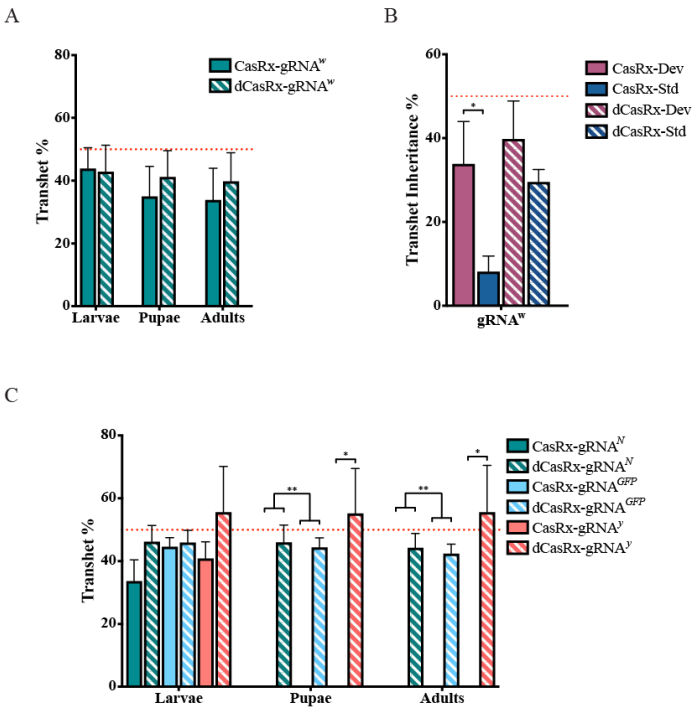


Figure S2.3: Development-related inheritance and lethality of Ubiq-CasRx and Ubiq-dCasRx transheterozygotes. (A-B) Transheterozygote percentages at larval, pupal, and adult developmental periods for each gRNA^{array} that produced an observable phenotype (*w*). There were no significant differences in inheritance. (C) Transheterozygote percentages through larval, pupal, and adult development periods for each gRNA^{array} that produced a lethal phenotype (*N*, *y*, *GFP*). No Ubiq-CasRx transheterozygotes developed beyond larvae.

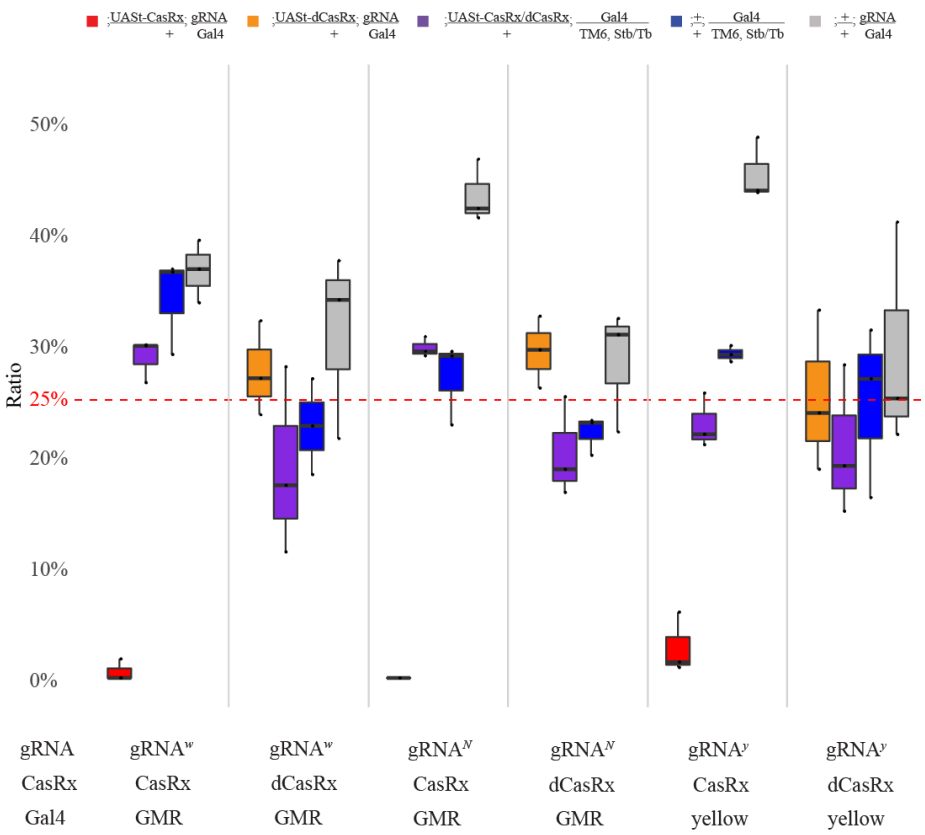


Figure S2.4: Complete inheritance data for binary Gal4/UAS crosses. The plot includes all genotypes scored in all crosses for UASt-CasRx and UASt-dCasRx. For all five gRNA^{array} targets, the inheritance of transheterozygous progeny expressing UASt-CasRx, a Gal4 driver, and a gRNA^{array} are lower compared to the other non-transheterozygous flies and to their corresponding dCasRx control group expressing UASt-dCasRx, a Gal4 driver, and a gRNA^{array}.

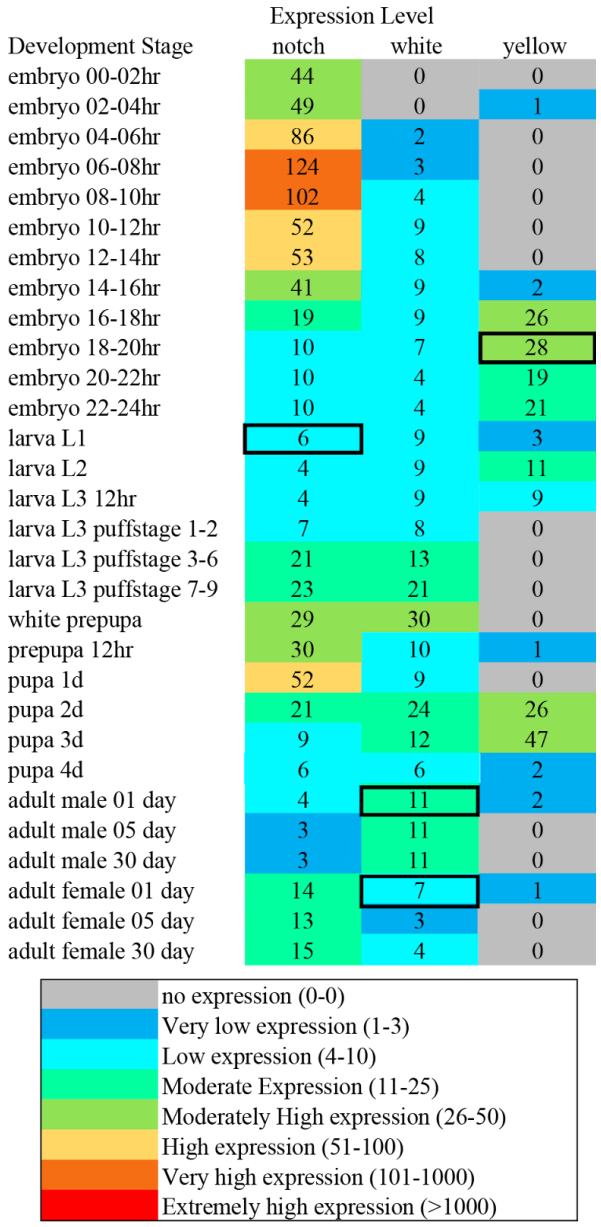


Figure S2.5: modENCODE transcript expression relative to *Drosophila melanogaster* development. Black box indicates which developmental period was chosen for RNA sequencing of samples for the analysis of CasRx-mediated transcript reduction in Ubiq-CasRx vs Ubiq-dCasRx comparison. Not included: GFP 1st instar larvae were chosen for analysis of *GFP* transcript reduction.

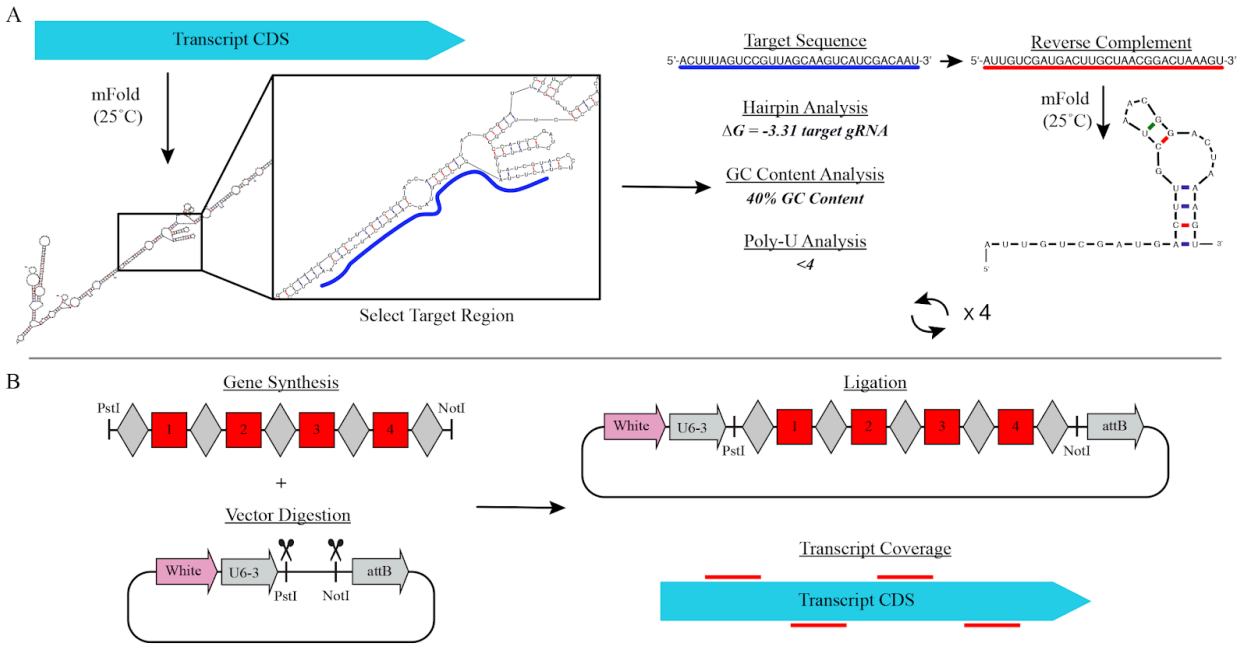


Figure S2.6: CasRx-gRNA^{array} transcript target selection and construct generation. (A) Schematic representing the workflow for gRNA choice. The transcript coding sequence (CDS) for a gene of interest (GOI) was entered into the mFold database (condition: 25°C) where predictive analysis identified the most probable secondary and tertiary folding of the entire transcript. We then chose specific regions that were predicted to be easily accessible for CasRx targeting (blue line), contained GC content between 30% and 70%, and possessed no poly-U stretches longer than 4 nt. We then converted the target sequence into the reverse complement (red line) and entered this spacer sequence into mFold (condition: 25°C) for hairpin analysis. This was repeated until four optimal target sites were selected. (B) Generation of gRNA^{array} construct. dsDNA was first synthesized to contain four spacer and five DR sequences with specific restriction sites present on the 5' and 3' end of the DNA. Simultaneously, the vector backbone containing the miniwhite marker, a U6:3 promoter fragment, and an attB site was digested using the corresponding restriction sites of the dsDNA gene fragment. The two pieces were then ligated together to generate a CasRx gRNA^{array} covering the majority of the transcript for the GOI.

2.9 Supplemental tables

Table S2.2: Complete data set for the Ubiq-CasRx and Ubiq-dCasRx bidirectional crosses. Absolute counts of inheritance and phenotype penetrance for maternal and paternal inheritance of Ubiq-CasRx and Ubiq-dCasRx crosses to gRNA^{array}-expressing flies. Each cross (paternal and maternal) was done in triplicate.

Cross	Cross (Cas Insertion Site)	gRNA Transgene	CasRx Transgene	CasRx Transgene Insertion Site	Number of Double Transhet	Double Transhet with Phenotype	Number of gRNA-only Progeny	Double Transhet Ratio	gRNA-only Ratio	Total Progeny	Cross Type
Ubiq-CasRx x gRNA-white	1050Ga1050f: 2(attp5)/+	U6-3-gRNA.w	Ubiq-CasRx	attp40w	10		10	104	8.771929825	114	Paternal
Ubiq-CasRx x gRNA-white	1050Ga1050f: 2(attp5)/+	U6-3-gRNA.w	Ubiq-CasRx	attp40w	7		7	158	4.242424242	165	Paternal
Ubiq-CasRx x gRNA-white	1050Ga1050f: 2(attp5)/+	U6-3-gRNA.w	Ubiq-CasRx	attp40w	22		22	168	11.57894737	190	Paternal
Ubiq-CasRx x gRNA-notch	1050Ga1050f: 2(attp40)/+	U6-3-gRNA.N	Ubiq-CasRx	attp40w	0		0	213	0	213	Paternal
Ubiq-CasRx x gRNA-notch	1050Ga1050f: 2(attp40)/+	U6-3-gRNA.N	Ubiq-CasRx	attp40w	0		0	195	0	195	Paternal
Ubiq-CasRx x gRNA-notch	1050Ga1050f: 2(attp40)/+	U6-3-gRNA.N	Ubiq-CasRx	attp40w	0		0	197	0	197	Paternal
Ubiq-CasRx x gRNA-yellow	1050ZAa1050f: 2(attp40)/+	U6-3-gRNA.y	Ubiq-CasRx	attp40w	0		0	199	0	199	Paternal
Ubiq-CasRx x gRNA-yellow	1050ZAa1050f: 2(attp40)/+	U6-3-gRNA.y	Ubiq-CasRx	attp40w	0		0	256	0	256	Paternal
Ubiq-CasRx x gRNA-yellow	1050ZAa1050f: 2(attp40)/+	U6-3-gRNA.y	Ubiq-CasRx	attp40w	0		0	167	0	167	Paternal
Ubiq-CasRx x gRNA-GFP	1050Ja1050f: 2(attp40)/+	U6-3-gRNA-GFP	Ubiq-CasRx	attp40w	0		0	185	0	185	Paternal
Ubiq-CasRx x gRNA-GFP	1050Ja1050f: 2(attp40)/+	U6-3-gRNA-GFP	Ubiq-CasRx	attp40w	0		0	173	0	173	Paternal
Ubiq-CasRx x gRNA-GFP	1050Ja1050f: 2(attp40)/+	U6-3-gRNA-GFP	Ubiq-CasRx	attp40w	0		0	160	0	160	Paternal
Ubiq-dCasRx x gRNA-white	1050Ga1050R(attp40)/+	U6-3-gRNA.w	Ubiq-dCasRx	attp40w	41		115	26.28205128	73.71794872	156	Paternal
Ubiq-dCasRx x gRNA-white	1050Ga1050R(attp40)/+	U6-3-gRNA.w	Ubiq-dCasRx	attp40w	56		127	30.60109979	69.3989071	183	Paternal
Ubiq-dCasRx x gRNA-white	1050Ga1050R(attp40)/+	U6-3-gRNA.w	Ubiq-dCasRx	attp40w	51		109	31.875	68.125	160	Paternal
Ubiq-dCasRx x gRNA-notch	1050Ja1050R(attp40)/+	U6-3-gRNA.N	Ubiq-dCasRx	attp40w	28		166	14.43298969	85.56701031	194	Paternal
Ubiq-dCasRx x gRNA-notch	1050Ja1050R(attp40)/+	U6-3-gRNA.N	Ubiq-dCasRx	attp40w	44		195	18.41004184	81.58995816	239	Paternal
Ubiq-dCasRx x gRNA-notch	1050Ja1050R(attp40)/+	U6-3-gRNA.N	Ubiq-dCasRx	attp40w	20		124	13.88888889	86.11111111	144	Paternal
Ubiq-dCasRx x gRNA-yellow	1050ZAa1050R(attp40)/+	U6-3-gRNA.y	Ubiq-dCasRx	attp40w	31		156	21.66435516	78.33564484	187	Paternal
Ubiq-dCasRx x gRNA-yellow	1050ZAa1050R(attp40)/+	U6-3-gRNA.y	Ubiq-dCasRx	attp40w	30		106	22.05882353	77.94117647	136	Paternal
Ubiq-dCasRx x gRNA-yellow	1050ZAa1050R(attp40)/+	U6-3-gRNA.y	Ubiq-dCasRx	attp40w	62		149	29.38388626	70.61611374	211	Paternal
Ubiq-dCasRx x gRNA-GFP	1050Ja1050R(attp40)/+	U6-3-gRNA-GFP	Ubiq-dCasRx	attp40w	80		102	43.95604396	56.04395604	182	Paternal
Ubiq-dCasRx x gRNA-GFP	1050Ja1050R(attp40)/+	U6-3-gRNA-GFP	Ubiq-dCasRx	attp40w	87		125	41.83773585	58.96226415	212	Paternal
Ubiq-dCasRx x gRNA-GFP	1050Ja1050R(attp40)/+	U6-3-gRNA-GFP	Ubiq-dCasRx	attp40w	86		130	39.81481481	60.18518519	216	Paternal
Ubiq-CasRx x gRNA-white	1050Ga1050f: 2(attp5)/+	U6-3-gRNA.w	Ubiq-CasRx	attp40w	20		127	13.60544218	86.39455782	147	Maternal
Ubiq-CasRx x gRNA-white	1050Ga1050f: 2(attp5)/+	U6-3-gRNA.w	Ubiq-CasRx	attp40w	21		117	15.21739913	84.78260087	138	Maternal
Ubiq-CasRx x gRNA-white	1050Ga1050f: 2(attp5)/+	U6-3-gRNA.w	Ubiq-CasRx	attp40w	33		102	24.44444444	75.55555556	135	Maternal
Ubiq-CasRx x gRNA-notch	1050Ja1050f: 2(attp40)/+	U6-3-gRNA.N	Ubiq-CasRx	attp40w	0		134	0	100	134	Maternal
Ubiq-CasRx x gRNA-notch	1050Ja1050f: 2(attp40)/+	U6-3-gRNA.N	Ubiq-CasRx	attp40w	0		122	0	100	122	Maternal
Ubiq-CasRx x gRNA-notch	1050Ja1050f: 2(attp40)/+	U6-3-gRNA.N	Ubiq-CasRx	attp40w	0		148	0	100	148	Maternal
Ubiq-CasRx x gRNA-yellow	1050ZAa1050f: 2(attp40)/+	U6-3-gRNA.y	Ubiq-CasRx	attp40w	0		117	0	100	117	Maternal
Ubiq-CasRx x gRNA-yellow	1050ZAa1050f: 2(attp40)/+	U6-3-gRNA.y	Ubiq-CasRx	attp40w	0		102	0	100	102	Maternal
Ubiq-CasRx x gRNA-yellow	1050ZAa1050f: 2(attp40)/+	U6-3-gRNA.y	Ubiq-CasRx	attp40w	0		129	0	100	129	Maternal
Ubiq-CasRx x gRNA-GFP	1050Ja1050f: 2(attp5)/+	U6-3-gRNA-GFP	Ubiq-CasRx	attp40w	0		106	0	100	106	Maternal
Ubiq-CasRx x gRNA-GFP	1050Ja1050f: 2(attp5)/+	U6-3-gRNA-GFP	Ubiq-CasRx	attp40w	0		136	0	100	136	Maternal
Ubiq-CasRx x gRNA-GFP	1050Ja1050f: 2(attp5)/+	U6-3-gRNA-GFP	Ubiq-CasRx	attp40w	0		115	0	100	115	Maternal
Ubiq-dCasRx x gRNA-white	1050Ga1050R(attp40)/+	U6-3-gRNA.w	Ubiq-dCasRx	attp40w	54		118	31.39534884	68.60465116	172	Maternal
Ubiq-dCasRx x gRNA-white	1050Ga1050R(attp40)/+	U6-3-gRNA.w	Ubiq-dCasRx	attp40w	45		129	25.86206897	74.13793103	174	Maternal
Ubiq-dCasRx x gRNA-white	1050Ga1050R(attp40)/+	U6-3-gRNA.w	Ubiq-dCasRx	attp40w	49		125	28.16091954	71.83908046	174	Maternal
Ubiq-dCasRx x gRNA-notch	1050Ja1050R(attp40)/+	U6-3-gRNA.N	Ubiq-dCasRx	attp40w	30		100	23.07692308	76.92307692	130	Maternal
Ubiq-dCasRx x gRNA-notch	1050Ja1050R(attp40)/+	U6-3-gRNA.N	Ubiq-dCasRx	attp40w	26		101	20.47256098	79.52743902	127	Maternal
Ubiq-dCasRx x gRNA-notch	1050Ja1050R(attp40)/+	U6-3-gRNA.N	Ubiq-dCasRx	attp40w	28		101	21.8523153	78.8476842	133	Maternal
Ubiq-dCasRx x gRNA-yellow	1050ZAa1050R(attp40)/+	U6-3-gRNA.y	Ubiq-dCasRx	attp40w	81		115	41.32653061	58.67346939	196	Maternal
Ubiq-dCasRx x gRNA-yellow	1050ZAa1050R(attp40)/+	U6-3-gRNA.y	Ubiq-dCasRx	attp40w	69		111	38.33333333	61.66666667	180	Maternal
Ubiq-dCasRx x gRNA-yellow	1050ZAa1050R(attp40)/+	U6-3-gRNA.y	Ubiq-dCasRx	attp40w	60		110	35.29411765	64.70588235	170	Maternal
Ubiq-dCasRx x gRNA-GFP	1050Ja1050R(attp40)/+	U6-3-gRNA-GFP	Ubiq-dCasRx	attp40w	48		110	30.37974684	69.62025316	158	Maternal
Ubiq-dCasRx x gRNA-GFP	1050Ja1050R(attp40)/+	U6-3-gRNA-GFP	Ubiq-dCasRx	attp40w	83		156	34.72803347	65.27196653	239	Maternal
Ubiq-dCasRx x gRNA-GFP	1050Ja1050R(attp40)/+	U6-3-gRNA-GFP	Ubiq-dCasRx	attp40w	85		162	34.41295547	65.58704453	247	Maternal

Table S2.3: Complete data set for the Gal4/UASt-CasRx or Gal4/UASt-dCasRx crosses. Absolute counts of inheritance and phenotype penetrance for the F_2 generation resulting from F_1 transheterozygote males expressing either UASt-CasRx/gRNA^{array} or UASt-dCasRx/gRNA^{array} crossed to Gal4 driver lines.

Cross	Cross (Internal Reference)	Gal4 Transgene	gRNA Transgene	CasRx Transgene	Replicate No.	Number of Progeny	Number of Triple Transhets	Number of Triple Transhets Manifesting Distinct Phenotypes	Triple Transhets Ratio	Distinct Phenotype Penetrance Among Triple Transhets	Number of CasRx, Stubble	Number of Non-CasRx, Stubble	Number of Non-CasRx, Non-Stubble	Ratio of CasRx, Stubble	Ratio of Non-CasRx, Stubble	Ratio of Non-CasRx, Non-Stubble	gRNA Target	Cross Type	Mean Transhet Penetrance
GMR-Gal4 -> UASt-CasRx_gRNA.white	29967.x.1050G1050Gl4821.GMR2	GMR2	10x5-4_gRNA.w	UASt-CasRx	1	214	4	4	0.018692	1	57	68	87	0.279245	0.332917	0.425837	white	GMR2-white	-
GMR-Gal4 -> UASt-CasRx_gRNA.white	29967.x.1050G1050Gl4821.GMR2	GMR2	10x5-4_gRNA.w	UASt-CasRx	2	181	0	0	NA		54	66	61	0.298343	0.364641	0.337017	white	GMR2-white	-
GMR-Gal4 -> UASt-CasRx_gRNA.white	29967.x.1050G1050Gl4821.GMR2	GMR2	10x5-4_gRNA.w	UASt-CasRx	3	98	0	0	NA		26	36	26	0.265306	0.367347	0.367347	white	GMR2-white	-
GMR-Gal4 -> UASt-dCasRx_gRNA.white	29967.x.1050G1050Gl4821.GMR2	GMR2	10x5-4_gRNA.w	UASt-dCasRx	1	93	22	0	0.236559		25	25	20	0.27957	0.28491	0.215054	white	GMR2-white	0
GMR-Gal4 -> UASt-dCasRx_gRNA.white	29967.x.1050G1050Gl4821.GMR2	GMR2	10x5-4_gRNA.w	UASt-dCasRx	2	144	28	0	0.194444		38	39	41	0.29577	0.302817	0.316901	white	GMR2-white	0
GMR-Gal4 -> UASt-dCasRx_gRNA.white	29967.x.1050G1050Gl4821.GMR2	GMR2	10x5-4_gRNA.w	UASt-dCasRx	3	106	34	0	0.320755		12	24	36	0.113208	0.226415	0.339623	white	GMR2-white	0
GMR-Gal4 -> UASt-CasRx_gRNA.notch	29967.x.1050G1050Gl4821.GMR2	GMR2	10x5-4_gRNA.N	UASt-CasRx	1	150	0	0	NA		44	44	62	0.293333	0.293333	0.413333	notch	GMR2-notch	-
GMR-Gal4 -> UASt-CasRx_gRNA.notch	29967.x.1050G1050Gl4821.GMR2	GMR2	10x5-4_gRNA.N	UASt-CasRx	2	83	0	0	NA		24	24	35	0.289157	0.289157	0.421687	notch	GMR2-notch	-
GMR-Gal4 -> UASt-CasRx_gRNA.notch	29967.x.1050G1050Gl4821.GMR2	GMR2	10x5-4_gRNA.N	UASt-CasRx	3	88	0	0	NA		27	20	41	0.306818	0.227273	0.465909	notch	GMR2-notch	-
GMR-Gal4 -> UASt-dCasRx_gRNA.notch	29967.x.1050I1050N5821.GMR2	GMR2	10x5-4_gRNA.N	UASt-dCasRx	1	31	8	0	0.258065		0	7	8	0.258065	0.258065	0.258065	notch	GMR2-notch	0
GMR-Gal4 -> UASt-dCasRx_gRNA.notch	29967.x.1050I1050N5821.GMR2	GMR2	10x5-4_gRNA.N	UASt-dCasRx	2	120	39	0	0.325		0	20	24	0.166667	0.2	0.308333	notch	GMR2-notch	0
GMR-Gal4 -> UASt-dCasRx_gRNA.notch	29967.x.1050I1050N5821.GMR2	GMR2	10x5-4_gRNA.N	UASt-dCasRx	3	96	25	0	0.260417		0	18	22	0.1875	0.229167	0.322917	notch	GMR2-notch	0
y-Gal4 -> UAS-CasRx_gRNA.yellow	44373.x.1050Z41050Gl4821.GMR2	GMR2	10x5-4_gRNA.y	UAS-CasRx	1	270	3	3	0.011111		1	44	61	0.003704	0.162963	0.633333	yellow	y2-yellow	0.933333
y-Gal4 -> UAS-CasRx_gRNA.yellow	44373.x.1050Z41050Gl4821.GMR2	GMR2	10x5-4_gRNA.y	UAS-CasRx	2	169	10	8	0.059172	0.8	37	48	74	0.218935	0.284024	0.43787	yellow	y2-yellow	0.933333
y-Gal4 -> UAS-CasRx_gRNA.yellow	44373.x.1050Z41050Gl4821.GMR2	GMR2	10x5-4_gRNA.y	UAS-CasRx	3	211	2	2	0.009479		1	54	63	0.004739	0.255924	0.298578	yellow	y2-yellow	0.933333
y-Gal4 -> UAS-dCasRx_gRNA.yellow	44373.x.1050Z41050Gl5962.GMR2	GMR2	10x5-4_gRNA.y	UASt-dCasRx	1	227	75	0	0.330396		0	34	61	0.22467	0.268722	0.251101	yellow	y2-yellow	0
y-Gal4 -> UAS-dCasRx_gRNA.yellow	44373.x.1050Z41050Gl5962.GMR2	GMR2	10x5-4_gRNA.y	UASt-dCasRx	2	77	4	0	0.051948		0	3	7	0.038961	0.1125	0.611526	yellow	y2-yellow	0
y-Gal4 -> UAS-dCasRx_gRNA.yellow	44373.x.1050Z41050Gl5962.GMR2	GMR2	10x5-4_gRNA.y	UASt-dCasRx	3	105	25	0	0.238095		0	20	17	0.190476	0.161905	0.469524	yellow	y2-yellow	0
y-Gal4 -> UAS-dCasRx_gRNA.yellow	44373.x.1050Z41050Gl5962.GMR2	GMR2	10x5-4_gRNA.y	UASt-dCasRx	4	40	17	0	0.425		0	4	15	0.1	0.375	0.375	yellow	y2-yellow	0

Table S2.4: Illumina RNA sequencing whole-transcriptome analysis samples. List of samples, in triplicate, analyzed for quantification of CasRx-mediated transcript reduction in comparison to dCasRx. The genotype, development stage or tissue type, and corresponding vectors are provided (Experimental = Ubiq-CasRx, Control = Ubiq-dCasRx).

Name	Genotype	Related Vector Name (s)	Sample Type
CasRx_w_adult_heads-R1	Ubiq-CasRx/+; gRNA(w)/+	OA-1050E.2/OA-1050G	Experimental
dCasRx_w_adult_heads-R1	Ubiq-dCasRx/+; gRNA(w)/+	OA-1050R/OA-1050G	Control
CasRx_w_adult_heads-R2	Ubiq-CasRx/+; gRNA(w)/+	OA-1050E.2/OA-1050G	Experimental
dCasRx_w_adult_heads-R2	Ubiq-dCasRx/+; gRNA(w)/+	OA-1050R/OA-1050G	Control
CasRx_w_adult_heads-R3	Ubiq-CasRx/+; gRNA(w)/+	OA-1050E.2/OA-1050G	Experimental
dCasRx_w_adult_heads-R3	Ubiq-dCasRx/+; gRNA(w)/+	OA-1050R/OA-1050G	Control
CasRx_N_larvae-R1	Ubiq-CasRx/+; gRNA(N)/+	OA-1050E.2/OA-1050I	Experimental
dCasRx_N_larvae-R1	Ubiq-dCasRx/+; gRNA(N)/+	OA-1050R/OA-1050I	Control
CasRx_N_larvae-R2	Ubiq-CasRx/+; gRNA(N)/+	OA-1050E.2/OA-1050I	Experimental
dCasRx_N_larvae-R2	Ubiq-dCasRx/+; gRNA(N)/+	OA-1050R/OA-1050I	Control
CasRx_N_larvae-R3	Ubiq-CasRx/+; gRNA(N)/+	OA-1050E.2/OA-1050I	Experimental
dCasRx_N_larvae-R3	Ubiq-dCasRx/+; gRNA(N)/+	OA-1050R/OA-1050I	Control
CasRx_GFP_larvae-R1	Ubiq-CasRx/+; gRNA(GFP)/+	OA-1050E.2/OA-1050J	Experimental
dCasRx_GFP_larvae-R1	Ubiq-dCasRx/+; gRNA(GFP)/+	OA-1050R/OA-1050J	Control
CasRx_GFP_larvae-R2	Ubiq-CasRx/+; gRNA(GFP)/+	OA-1050E.2/OA-1050J	Experimental
dCasRx_GFP_larvae-R2	Ubiq-dCasRx/+; gRNA(GFP)/+	OA-1050R/OA-1050J	Control
CasRx_GFP_larvae-R3	Ubiq-CasRx/+; gRNA(GFP)/+	OA-1050E.2/OA-1050J	Experimental
dCasRx_GFP_larvae-R3	Ubiq-dCasRx/+; gRNA(GFP)/+	OA-1050R/OA-1050J	Control
CasRx_y_embryos-R1	Ubiq-CasRx/+; gRNA(y)/+ & CyO/+; gRNA(y)/+	OA-1050E.2/OA-1050Z4	Experimental
dCasRx_y_embryos-R1	Ubiq-dCasRx/+; gRNA(y)/+ & CyO/+; gRNA(y)/+	OA-1050R/OA-1050Z4	Control
CasRx_y_embryos-R2	Ubiq-CasRx/+; gRNA(y)/+ & CyO/+; gRNA(y)/+	OA-1050E.2/OA-1050Z4	Experimental
dCasRx_y_embryos-R2	Ubiq-dCasRx/+; gRNA(y)/+ & CyO/+; gRNA(y)/+	OA-1050R/OA-1050Z4	Control
CasRx_y_embryos-R3	Ubiq-CasRx/+; gRNA(y)/+ & CyO/+; gRNA(y)/+	OA-1050E.2/OA-1050Z4	Experimental
dCasRx_y_embryos-R3	Ubiq-dCasRx/+; gRNA(y)/+ & CyO/+; gRNA(y)/+	OA-1050R/OA-1050Z4	Control

Table S2.13: Primers used for vector construction. A list of primers and their respective sequences used to generate the constructs used in this study.

Construct	Description	Primer	Primer Sequence (5' to 3')	PCR Template
OA-1050E	CasRx	1050E.C3	TACTAATTTTCCACATCTCTATTTTGACCCGCAGATTAATTAATGAGCCCCAAGAAGAA	pNLS-RfxCas13d-NLS-HA (pCasRx)
		1050E.C4	CAATTGATTTGTTATTTTAAAAACGATTCATTCTAGCTAGCTTAAGCGTAATCTGGAACA	
OA-1050R	dCasRx	1050E.C3	TACTAATTTTCCACATCTCTATTTTGACCCGCAGATTAATTAATGAGCCCCAAGAAGAA	pNLS-dRfxCas13d-NLS-HA (pdCasRx)
		1050E.C4	CAATTGATTTGTTATTTTAAAAACGATTCATTCTAGCTAGCTTAAGCGTAATCTGGAACA	
OA-1050L	UASt promoter	1041.C9	GCGGGTTCTCGACGGTCACGGCGGGCATGTCGACGCGGCCGCAACCAACAACACTAGTAG	pJFRC81
		1041.C11	CTGGCCTCCACCTTTCTCTTCTTCTTGGGGCTCATGTTTAAACCCAATTCCCTATTCAGA	
	CasRx	1050L.C1	AATACAAGAAGAGAACTCTGAATAGGGAATTGGGTTTAAACATGAGCCCCAAGAAGAA	pCasRx
		1050E.C4	CAATTGATTTGTTATTTTAAAAACGATTCATTCTAGCTAGCTTAAGCGTAATCTGGAACA	
OA-1050S	UASt promoter	1041.C9	GCGGGTTCTCGACGGTCACGGCGGGCATGTCGACGCGGCCGCAACCAACAACACTAGTAG	pJFRC81
		1041.C11	CTGGCCTCCACCTTTCTCTTCTTCTTGGGGCTCATGTTTAAACCCAATTCCCTATTCAGA	
	dCasRx	1050L.C1	AATACAAGAAGAGAACTCTGAATAGGGAATTGGGTTTAAACATGAGCCCCAAGAAGAA	pdCasRx
		1050E.C4	CAATTGATTTGTTATTTTAAAAACGATTCATTCTAGCTAGCTTAAGCGTAATCTGGAACA	
OA-1043	U6:3 promoter	1043.C1	GGGAATTGGGAATTGGGCAATATTTAAATGGCGGCGCGCCGAATTCTTTTTTGCTCACCT	Addgene plasmid #112688
		1043.C23	ACACTAGTGGATCTCTAGAGGTACCGTTGCGGCCGCAAAAAAGTTGTAATAGCCCCTCAAAACTGGACCTTCCACAACTGCAGCCGACGTTAAATTGAA	
OA-1052B	Ubiq promoter	1052B.C1	GGGAATTGGGCAATATTTAAATGGCGGCTGCAGCGCGCAGATCGCCGAT	Addgene plasmid #112686
		1052B.C2	TTTCTTTATGTTTTTGGCGTCTTCCATCCTAGGTCTGCGGGTCAAAATAGAGATG	
	T2A-eGFP	908A1	ATAAAGGCCAAGAAGGGCGGAAAGATCGCCGTGGAGGGCAGAGGAAGTCTTCTAACATGC	Addgene plasmid #112686
		908A2	TTGTTATTTTAAAAACGATTCATTCTAGGCGATCGCTTACTTGTACAGCTCGTCCATGCC	
	versed Ubiq promo	908A3	ACCGTGACCTACATCGTCGACACTAGTGGATCTCTAGACGCGCAGATCGCCGATG	Addgene plasmid #112686
		908A4	GGATCATAAACTTTCGAAGTCATGCGGCCGCTCTGCGGGTCAAAATAGAGATGT	

2.10 References

Abudayyeh, Omar O., Jonathan S. Gootenberg, Patrick Essletzbichler, Shuo Han, Julia Joung, Joseph J. Belanto, Vanessa Verdine, et al. 2017. "RNA Targeting with CRISPR-Cas13." *Nature* 550 (7675): 280–84. https://doi.org/10.1038/nature24049.

Abudayyeh, Omar O., Jonathan S. Gootenberg, Silvana Konermann, Julia Joung, Ian M. Slaymaker, David B. T. Cox, Sergey Shmakov, et al. 2016. "C2c2 Is a Single-Component Programmable RNA-Guided RNA-Targeting CRISPR Effector." *Science* 353 (6299): aaf5573. https://doi.org/10.1126/science.aaf5573.

Adli, Mazhar. 2018. "The CRISPR Tool Kit for Genome Editing and beyond." *Nature Communications* 9 (1): 1911. https://doi.org/10.1038/s41467-018-04252-2.

Akbari, Omar S., Daniel Oliver, Katie Eyer, and Chi-Yun Pai. 2009. "An Entry/Gateway Cloning System for General Expression of Genes with Molecular Tags in Drosophila Melanogaster." *BMC Cell Biology* 10 (January): 8. https://doi.org/10.1186/1471-2121-10-8.

Biessmann, H. 1985. "Molecular Analysis of the Yellow Gene (y) Region of Drosophila Melanogaster." *Proceedings of the National Academy of Sciences of the United States of America* 82 (21): 7369–73. https://doi.org/10.1073/pnas.82.21.7369.

Brand, A. H., and N. Perrimon. 1993. "Targeted Gene Expression as a Means of Altering Cell Fates and Generating Dominant Phenotypes." *Development* 118 (2): 401–15. https://www.ncbi.nlm.nih.gov/pubmed/8223268.

Buchman, Anna, Stephanie Gamez, Ming Li, Igor Antoshechkin, Hsing-Han Li, Hsin-Wei Wang, Chun-Hong Chen, et al. 2019. "Engineered Resistance to Zika Virus in Transgenic *Aedes Aegypti* Expressing a Polycistronic Cluster of Synthetic Small RNAs." *Proceedings of the National Academy of Sciences of the United States of America* 116 (9): 3656–61. https://doi.org/10.1073/pnas.1810771116.

Buchman, Anna, Stephanie Gamez, Ming Li, Igor Antoshechkin, Hsing-Han Li, Hsin-Wei Wang, Chun-Hong Chen, et al. 2020. "Broad Dengue Neutralization in Mosquitoes Expressing an Engineered Antibody." *PLoS Pathogens* 16 (1): e1008103. https://doi.org/10.1371/journal.ppat.1008103.

Chakraborty, Chiranjib, Seong Lin Teoh, and Srijit Das. 2017. "The Smart Programmable CRISPR Technology: A Next Generation Genome Editing Tool for Investigators." *Current Drug Targets* 18 (14): 1653–63. https://doi.org/10.2174/1389450117666160527142321.

Champer, Jackson, Anna Buchman, and Omar S. Akbari. 2016. "Cheating Evolution: Engineering Gene Drives to Manipulate the Fate of Wild Populations." *Nature Reviews. Genetics* 17 (3): 146–59. https://doi.org/10.1038/nrg.2015.34.

Dan L. Lindsley and E. H. Grell. n.d. "Genetic Variations of Drosophila Melanogaster."

https://publicationsonline.carnegiescience.edu/publications_online/genetic_variations.pdf.

Dietzl, Georg, Doris Chen, Frank Schnorrer, Kuan-Chung Su, Yulia Barinova, Michaela Fellner, Beate Gasser, et al. 2007. “A Genome-Wide Transgenic RNAi Library for Conditional Gene Inactivation in Drosophila.” *Nature* 448 (7150): 151–56. https://doi.org/10.1038/nature05954.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. “STAR: Ultrafast Universal RNA-Seq Aligner.” *Bioinformatics* 29 (1): 15–21. https://doi.org/10.1093/bioinformatics/bts635.

East-Seletsky, Alexandra, Mitchell R. O’Connell, David Burstein, Gavin J. Knott, and Jennifer A. Doudna. 2017. “RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes.” *Molecular Cell* 66 (3): 373–83.e3. https://doi.org/10.1016/j.molcel.2017.04.008.

East-Seletsky, Alexandra, Mitchell R. O’Connell, Spencer C. Knight, David Burstein, Jamie H. D. Cate, Robert Tjian, and Jennifer A. Doudna. 2016. “Two Distinct RNase Activities of CRISPR-C2c2 Enable Guide-RNA Processing and RNA Detection.” *Nature* 538 (7624): 270–73. https://doi.org/10.1038/nature19802.

Franz, Alexander W. E., Irma Sanchez-Vargas, Zach N. Adelman, Carol D. Blair, Barry J. Beaty, Anthony A. James, and Ken E. Olson. 2006. “Engineering RNA Interference-Based Resistance to Dengue Virus Type 2 in Genetically Modified Aedes Aegypti.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (11): 4198–4203. https://doi.org/10.1073/pnas.0600479103.

Geyer, P. K., M. M. Green, and V. G. Corces. 1990. “Tissue-Specific Transcriptional Enhancers May Act in Trans on the Gene Located in the Homologous Chromosome: The Molecular Basis of Transvection in Drosophila.” *The EMBO Journal* 9 (7): 2247–56. https://www.ncbi.nlm.nih.gov/pubmed/2162766.

Gibson, Daniel G., Lei Young, Ray-Yuan Chuang, J. Craig Venter, Clyde A. Hutchison 3rd, and Hamilton O. Smith. 2009. “Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases.” *Nature Methods* 6 (5): 343–45. https://doi.org/10.1038/nmeth.1318.

Graveley, Brenton R., Angela N. Brooks, Joseph W. Carlson, Michael O. Duff, Jane M. Landolin, Li Yang, Carlo G. Artieri, et al. 2011. “The Developmental Transcriptome of Drosophila Melanogaster.” *Nature* 471 (7339): 473–79. https://doi.org/10.1038/nature09715.

Heigwer, Florian, Fillip Port, and Michael Boutros. 2018. “RNA Interference (RNAi) Screening in Drosophila.” *Genetics* 208 (3): 853–74. https://doi.org/10.1534/genetics.117.300077.

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012. “A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity.” *Science* 337 (6096): 816–21. https://doi.org/10.1126/science.1225829.

Kandul, Nikolay P., Junru Liu, Hector M. Sanchez C, Sean L. Wu, John M. Marshall, and Omar S. Akbari. 2019. “Transforming Insect Population Control with Precision Guided Sterile Males with Demonstration in Flies.” *Nature Communications* 10 (1): 84. https://doi.org/10.1038/s41467-018-07964-7.

Kidd, S., M. R. Kelley, and M. W. Young. 1986. “Sequence of the Notch Locus of Drosophila Melanogaster: Relationship of the Encoded Protein to Mammalian Clotting and Growth Factors.” *Molecular and Cellular Biology* 6 (9): 3094–3108. https://doi.org/10.1128/mcb.6.9.3094.

Konermann, Silvana, Peter Lotfy, Nicholas J. Brideau, Jennifer Oki, Maxim N. Shokhirev, and Patrick D. Hsu. 2018. “Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors.” *Cell* 173 (3): 665–76.e14. https://doi.org/10.1016/j.cell.2018.02.033.

Kulkarni, Meghana M., Matthew Booker, Serena J. Silver, Adam Friedman, Pengyu Hong, Norbert Perrimon, and Bernard Mathey-Prevot. 2006. “Evidence of off-Target Effects Associated with Long dsRNAs in Drosophila Melanogaster Cell-Based Assays.” *Nature Methods* 3 (10): 833–38. https://doi.org/10.1038/nmeth935.

Kushawah, Gopal, Luis Hernandez-Huertas, Joaquin Abugattas-Nuñez Del Prado, Juan R. Martinez-Morales, Michelle L. DeVore, Huzaifa Hassan, Ismael Moreno-Sanchez, et al. 2020. “CRISPR-Cas13d Induces Efficient mRNA Knockdown in Animal Embryos.” *Developmental Cell* 54 (6): 805–17.e7. https://doi.org/10.1016/j.devcel.2020.07.013.

Leonardi, Jessica, Rodrigo Fernandez-Valdivia, Yi-Dong Li, Amanda A. Simcox, and Hamed Jafar-Nejad. 2011. “Multiple O-Glucosylation Sites on Notch Function as a Buffer against Temperature-Dependent Loss of Signaling.” *Development* 138 (16): 3569–78. https://doi.org/10.1242/dev.068361.

Liao, Yang, Gordon K. Smyth, and Wei Shi. 2014. “featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features.” *Bioinformatics* 30 (7): 923–30. https://doi.org/10.1093/bioinformatics/btt656.

Markstein, Michele, Chrysoula Pitsouli, Christians Villalta, Susan E. Celniker, and Norbert Perrimon. 2008. “Exploiting Position Effects and the Gypsy Retrovirus Insulator to Engineer Precisely Expressed Transgenes.” *Nature Genetics* 40 (4): 476–83. https://doi.org/10.1038/ng.101.

Massey, Jonathan H., Daayun Chung, Igor Siwanowicz, David L. Stern, and Patricia J. Wittkopp. 2019. “The Yellow Gene Influences Drosophila Male Mating Success through Sex Comb Melanization.” *bioRxiv*. https://doi.org/10.1101/673756.

Mathur, G., I. Sanchez-Vargas, D. Alvarez, K. E. Olson, O. Marinotti, and A. A. James. 2010. “Transgene-Mediated Suppression of Dengue Viruses in the Salivary Glands of the Yellow Fever Mosquito, Aedes Aegypti.” *Insect Molecular Biology* 19 (6): 753–63. https://doi.org/10.1111/j.1365-2583.2010.01032.x.

Ma, Yong, Adrian Creanga, Lawrence Lum, and Philip A. Beachy. 2006. “Prevalence of off-Target Effects in Drosophila RNA Interference Screens.” *Nature* 443 (7109): 359–63. https://doi.org/10.1038/nature05179.

Meeske, Alexander J., Sandra Nakandakari-Higa, and Luciano A. Marraffini. 2019. “Cas13-Induced Cellular Dormancy Prevents the Rise of CRISPR-Resistant Bacteriophage.” *Nature* 570 (7760): 241–45. https://doi.org/10.1038/s41586-019-1257-5.

Micchelli, Craig A., and Norbert Perrimon. 2006. “Evidence That Stem Cells Reside in the Adult Drosophila Midgut Epithelium.” *Nature* 439 (7075): 475–79. https://doi.org/10.1038/nature04371.

Ni, Jian-Quan, Lu-Ping Liu, Richard Binari, Robert Hardy, Hye-Seok Shim, Amanda Cavallaro, Matthew Booker, et al. 2009. “A Drosophila Resource of Transgenic RNAi Lines for Neurogenetics.” *Genetics* 182 (4): 1089–1100. https://doi.org/10.1534/genetics.109.103630.

Ni, Jian-Quan, Michele Markstein, Richard Binari, Barret Pfeiffer, Lu-Ping Liu, Christians Villalta, Matthew Booker, Lizabeth Perkins, and Norbert Perrimon. 2008. “Vector and Parameters for Targeted Transgenic RNA Interference in Drosophila Melanogaster.” *Nature Methods* 5 (1): 49–51. https://doi.org/10.1038/nmeth1146.

Ni, Jian-Quan, Rui Zhou, Benjamin Czech, Lu-Ping Liu, Laura Holderbaum, Donghui Yang-Zhou, Hye-Seok Shim, et al. 2011. “A Genome-Scale shRNA Resource for Transgenic RNAi in Drosophila.” *Nature Methods* 8 (5): 405–7. https://doi.org/10.1038/nmeth.1592.

Peretz, Gabriella, Anna Bakhrat, and Uri Abdu. 2007. “Expression of the Drosophila Melanogaster GADD45 Homolog (CG11086) Affects Egg Asymmetric Development That Is Mediated by the c-Jun N-Terminal Kinase Pathway.” *Genetics* 177 (3): 1691–1702. https://doi.org/10.1534/genetics.107.079517.

Perrimon, Norbert, and Bernard Mathey-Prevot. 2007. “Matter Arising: Off-Targets and Genome-Scale RNAi Screens in Drosophila.” *Fly* 1 (1): 1–5. https://doi.org/10.4161/fly.3601.

Perrimon, Norbert, Jian-Quan Ni, and Lizabeth Perkins. 2010. “In Vivo RNAi: Today and Tomorrow.” *Cold Spring Harbor Perspectives in Biology* 2 (8): a003640. https://doi.org/10.1101/cshperspect.a003640.

Pfeifer, Tom A., Dwayne D. Hegedus, Thomas A. Grigliatti, and David A. Theilmann. 1997. “Baculovirus Immediate-Early Promoter-Mediated Expression of the Zeocin™ Resistance Gene for Use as a Dominant Selectable Marker in Dipteran and Lepidopteran Insect Cell Lines.” *Gene* 188 (2): 183–90. https://doi.org/10.1016/S0378-1119(96)00756-1.

Port, Fillip, Hui-Min Chen, Tzumin Lee, and Simon L. Bullock. 2014. “Optimized CRISPR/Cas Tools for Efficient Germline and Somatic Genome Engineering in Drosophila.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (29): E2967–76. https://doi.org/10.1073/pnas.1405500111.

Saj, Abil, Zeynep Arziman, Denise Stempfle, Werner van Belle, Ursula Sauder, Thomas Horn, Markus Dürrenberger, Renato Paro, Michael Boutros, and Gunter Merdes. 2010. “A Combined Ex Vivo and in Vivo RNAi Screen for Notch Regulators in Drosophila Reveals an Extensive Notch Interaction Network.” *Developmental Cell* 18 (5): 862–76. https://doi.org/10.1016/j.devcel.2010.03.013.

Simón, Rocío, Ricardo Aparicio, Ben E. Housden, Sarah Bray, and Ana Busturia. 2014. “Drosophila p53 Controls Notch Expression and Balances Apoptosis and Proliferation.” *Apoptosis: An International Journal on Programmed Cell Death* 19 (10): 1430–43. https://doi.org/10.1007/s10495-014-1000-5.

Smargon, Aaron A., David B. T. Cox, Neena K. Pyzocha, Kaijie Zheng, Ian M. Slaymaker, Jonathan S. Gootenberg, Omar A. Abudayyeh, et al. 2017. “Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28.” *Molecular Cell*. https://doi.org/10.1016/j.molcel.2016.12.023.

Sullivan, D. T., and M. C. Sullivan. 1975. “Transport Defects as the Physiological Basis for Eye Color Mutants of Drosophila Melanogaster.” *Biochemical Genetics* 13 (9-10): 603–13. https://doi.org/10.1007/bf00484918.

Wessels, Hans-Hermann, Alejandro Méndez-Mancilla, Xinyi Guo, Mateusz Legut, Zharko Daniloski, and Neville E. Sanjana. 2019. “Principles for Rational Cas13d Guide Design.” *bioRxiv*. https://doi.org/10.1101/2019.12.27.889089.

Yan, Winston X., Shaorong Chong, Huaibin Zhang, Kira S. Makarova, Eugene V. Koonin, David R. Cheng, and David A. Scott. 2018. “Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein.” *Molecular Cell* 70 (2): 327–39.e5. https://doi.org/10.1016/j.molcel.2018.02.028.

Yen, Pei-Shi, Anthony James, Jian-Chiuan Li, Chun-Hong Chen, and Anna-Bella Failloux. 2018. “Synthetic miRNAs Induce Dual Arboviral-Resistance Phenotypes in the Vector Mosquito Aedes Aegypti.” *Communications Biology* 1 (February): 11. https://doi.org/10.1038/s42003-017-0011-5.

Zuker, Michael. 2003. “Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction.” *Nucleic Acids Research* 31 (13): 3406–15. https://doi.org/10.1093/nar/gkg595.

Chapter 3: Manipulation of Cas13d collateral cleavage activity for the development of a rapid and sensitive nucleic acid diagnostic assay

3.1 Abstract

The development of an extensive toolkit for potential point-of-care diagnostics that is expeditiously adaptable to new emerging pathogens is of critical public health importance. Recently, a number of novel CRISPR-based diagnostics have been developed to detect SARS-CoV-2. Herein, we outline the development of an alternative CRISPR nucleic acid diagnostic utilizing a Cas13d ribonuclease derived from *Ruminococcus flavefaciens* XPD3002 (CasRx) to detect SARS-CoV-2, an approach we term SENSR (Sensitive Enzymatic Nucleic-acid Sequence Reporter) that can detect attomolar concentrations of SARS-CoV-2. We demonstrate 100% sensitivity in patient-derived samples by lateral flow and fluorescence readout with a detection limit of 45 copy/μL. This technology expands the available nucleic acid diagnostic toolkit which can be adapted to combat future pandemics.

3.2 Introduction

Following emergence from China in late 2019 (Zhu et al. 2020; Mallapaty 2020; Andersen et al. 2020), severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Wang et al. 2020; Zhu et al. 2020) has spread to almost every country despite unprecedented control efforts (WHO 2020). Compared to H1N1, Ebola, MERS, and SARS-CoV-1 outbreaks, this novel coronavirus represents the first pandemic characterized by widespread global transmission coupled with significant mortality. Robust identification and isolation of all infected individuals is essential for curtailing disease transmission. The ongoing SARS-CoV-2 pandemic presents an unparalleled global public health emergency, which spurred the urgent development of molecular

diagnostics and therapeutics for timely patient identification, isolation and treatment. The economic, health, and societal damage wrought by SARS-CoV-2 highlights the importance of expanding and improving on current diagnostic technologies to identify and prevent future pandemics.

Detection of nucleic acids from pathogens is typically accomplished using real-time polymerase chain reaction (RT-qPCR) methodologies. RT-qPCR diagnostics are highly specific and sensitive, detecting nucleic acids consistently at low copy numbers (Vogels et al. 2020). However, these diagnostics are often time-consuming and require specialized laboratory equipment, limiting the ability to expand RT-qPCR diagnostics into point-of-care settings (Rezaei et al. 2020). Therefore, alternative technologies with the potential to yield cost- and time-effective point-of-care diagnostics demands investment.

CRISPR-Cas nucleases can be easily programmed to target nucleic acids in a sequence-specific manner (Jinek et al. 2012; Abudayyeh et al. 2016; Zetsche et al. 2015), making them prime candidates for the detection and diagnosis of genetic material, and therefore comprise the CRISPR diagnostics (CRISPRDx) pipeline (Gootenberg et al. 2017, 2018; Chen et al. 2018; Li et al. 2018). These systems rely on Type II Cas enzymes to physically bind target sequences (Azhar et al. 2021), or collateral cleavage by Type V or Type VI nucleases to detect DNA (Li et al. 2018; Chen et al. 2018; Harrington et al. 2018) or RNA species, respectively (Gootenberg et al. 2017, 2018; Freije et al. 2019). Since pandemic onset, an array of innovative diagnostics and prophylactics relying on these effectors have been adapted to detect, or target, SARS-CoV-2 with unprecedented speed (Mukama et al. 2020; Hajian et al. 2019; Patchsung et al. 2020; Joung et al. 2020; Ding et al. 2020; Broughton, Deng, Yu, Fasching, Servellita, et al. 2020; Zhang, Abudayyeh, and Jonathan 2020; Abbott et al. 2020; Azhar et al. 2021; Ackerman

et al. 2020; Rauch et al. 2021; Metsky et al. 2020; Lucia, Federico, and Alejandra 2020), most notably represented by the DETECTR (DNA Endonuclease Targeted CRISPR Trans Reporter) (Chen et al. 2018; Li et al. 2018) and SHERLOCK (Specific High-Sensitivity Enzymatic Reporter unLOCKing) (Gootenberg et al. 2017, 2018) systems (Summarized in Fig S1, Table S1). To maximize all the capabilities of and expand the CRISPRDx toolkit, it is important to evaluate all Cas enzymes that can complement or supplement existing systems.

Similar to Cas ribonucleases used in CRISPRDx systems, Cas13d effectors, such as RfxCas13d (CasRx), exclusively target RNA species and possess target-dependent collateral cleavage activity (Konermann et al. 2018; Buchman et al. 2020; Yan et al. 2018). Cas13d enzymes do not require a Protospacer Flanking Sequence (PFS) (Konermann et al. 2018; Yan et al. 2018; Abudayyeh et al. 2016; Kellner et al. 2019), thus presenting an advantage for flexible targeting. While the genetic modulatory effects of CasRx have been thoroughly characterized in *Drosophila*, *zebrafish*, and human cells (Konermann et al. 2018; Buchman et al. 2020; Sun et al. 2021; Kushawah et al. 2020), and its putative prophylactic properties against SARS-CoV-2 have been demonstrated (Abbott et al. 2020), its potential as a diagnostic system has not yet been explored.

Herin, we report the first use of CasRx as a molecular diagnostic, developing a unique system we term SENSR (Sensitive Enzymatic Nucleic-acid Sequence Reporter) and demonstrate robust detection of SARS-CoV-2 viral sequences. We exploit the collateral cleavage activity of CasRx to detect SARS-CoV-2 in both synthetic templates as well as in patient-derived samples via fluorescence and paper-based LF readouts. To maximize specificity, we performed an extensive bioinformatic analysis to identify novel conserved and specific viral targets to minimize false-negative and false-positive rates, respectively. We demonstrate that within a

detection limit of 45 copy/μL, SENSR exhibits 100% sensitivity by both fluorescence and LF readouts, detecting SARS-CoV-2 in 60 min or less total reaction time. The detection limit of SENSR is comparable to previously established CRISPRDx systems (Patchsung et al. 2020; Broughton, Deng, Yu, Fasching, Servellita, et al. 2020; Zhang, Abudayyeh, and Jonathan 2020) and the system presents promise for improvement.

3.3 Results

3.3.1 SENSR workflow

Derived from protocols originally developed for (Fig S1, Table S1) Cas13a and Cas13b (Kellner et al. 2019; Gootenberg et al. 2017, 2018; Patchsung et al. 2020), we established a two-step nucleic acid detection protocol using recombinant CasRx (Fig S2). Target sequences are first amplified in a 30 min isothermal preamplification reaction by combining reverse transcription with recombinase polymerase amplification (RT-RPA) (Piepenburg et al. 2006). The RT-RPA reaction produces numerous dsDNA amplicons containing a gRNA target site and an upstream T7 promoter sequence. Next, the RT-RPA reaction is transferred into a second reaction, termed CasRx Cleavage Reaction (CCR), containing a T7 polymerase and the CasRx ribonucleoprotein. In the CCR the dsDNA amplicons are transcribed into ssRNA molecules to be cleaved by CasRx. This initial cleavage event initiates the collateral cleavage activity of CasRx, resulting in cleavage of a bystander ssRNA probe. Collateral cleavage of the probe is analyzed by either fluorescence or lateral flow (LF) readouts, thus indicating the presence or absence of SARS-CoV-2 genomic sequences (Fig 1A, Fig S3).

3.3.2 Target selection and validation

Diagnostics require high specificity to limit the probability of false positives from detection of random nucleic acids. To ensure the specificity of our target sites, we established a bioinformatic pipeline and searched for 30nt sequences conserved across the first 433 published SARS-CoV-2 genomes (available at Genbank on April 7th, 2020), and without homology to other coronaviruses (ViPR, Virus Pathogen Resource, n = 3,164). This search yielded a panel of gRNA target sites (n=8846) less likely to result in false positives or negatives due to sequence constraints (Fig S4, Table S2-S3). Aligning with previously established diagnostics, we selected the *Spike* (*S*) and *Nucleocapsid* (*N*) genes as the targets of SENSR for system validation (Corman et al. 2020; Lu et al. 2020; Patchsung et al. 2020; Broughton, Deng, Yu, Fasching, Servellita, et al. 2020). The bioinformatic analysis revealed multiple specific sequences within the *S*-gene (n = 1568) and the *N*-gene (n = 150), to which we designed three guides for each gene (S: gRNA-S1, -S2, -S3; N: gRNA-N1, -N2, -N3) (Fig 1B, Table S2, S4). To validate these gRNAs against SARS-CoV-2, we generated synthetic ssRNA gene fragments mimicking portions of the *Spike* and *Nucleocapsid* genomic sequences (Fig 1B, Table S4), and assessed CasRx cleavage activity *in vitro*. Initial characterization of on-target cleavage revealed degradation of target transcripts for all guides tested (Fig S5A-B) motivating further assessment of all candidates.

To select the best candidate for each target gene, we tested each gRNA in a preliminary SENSR fluorescence reaction (Fig 1A) against a high-copy number of synthetic template (10,000 copy/μL). Fluorescence data acquired were then analyzed to determine the time each gRNA reached its half-maximum fluorescence (HMF). HMF values were calculated by fitting a non-linear regression to the fluorescence data acquired over 1 hr. The equation for the nonlinear regression was then used to solve for the time HMF occurred (see supporting information). From

the analysis, gRNA-S3 (t = 5.91 min) of the *S*-targeting group and gRNA-N1 (t = 12.55 min) of the *N*-targeting group resulted in the fastest HMF times and were selected for downstream analysis (Fig 1C-D).

3.3.3 Optimization of SENSR

We and others have recently demonstrated that on-target cleavage activates a secondary collateral cleavage property of CasRx (Konermann et al. 2018; Buchman et al. 2020; Sun et al. 2021). We sought to exploit the collateral cleavage activity of CasRx to cleave a bystander fluorescent probe in *trans*, facilitating detection of SARS-CoV-2 by fluorimetry (Fig S3). A previous study has demonstrated that each Cas13 enzyme exhibits select di-nucelotide sequence preferences for collateral cleavage (Gootenberg et al. 2018), however, no such analysis has been performed on CasRx. To develop a probe cleavable by CasRx, we generated ten custom 6 nucleotide ssRNA probes, with variable di-nucleotide sequences, each conjugated to a 5' fluorescent molecule (6-FAM) and a 3' fluorescence quencher (FQ), whereupon separation following cleavage results in detectable fluorescence signal (Fig S6A, Table S4). Following incubation of CasRx with gRNA-N1 in the presence or absence of the synthetic *N* gene, the poly-U and CU/UC probes yielded significant and detectable fluorescence compared to the no-template control (NTC) (poly-U & CU/UC: $p<0.0001$). However, the fluorescence signal increase over noise in the poly-U group was significantly higher than that of the CU/UC group ($p<0.0001$) suggesting a preference for poly-U stretches by CasRx, and motivating the use of the poly-U probe in the remainder of experiments (Fig S6B-C, Table S4).

Following probe selection, we set to optimize the SENSR reaction conditions for the amplification and cleavage reactions. We first evaluated how varying the volume of sample input into the RT-RPA reaction affected the detection of a target. To do so, we diluted the synthetic

Nucleocapsid template to 1,000 copy/µL and added the templates to the RT-RPA reaction at 10%, 20%, 28.5%, 39%, 47%, and 52% v/v (Table S5). After amplification, the samples were then transferred to the CCR and targeted by gRNA-N1. From the analysis we found the 28.5% volume input group resulted in the fastest detection time compared to all other conditions tested (HMF = 7.85 min) and selected 28.5% sample input as our optimized condition (Fig S7A-B). We next set out to determine how varying the volume of RT-RPA transferred into the CCR influenced the speed of detection. We accomplished this by amplifying 1,000 copy/µL of the synthetic *Nucleocapsid* template, transferring varied volumes of RT-RPA into the CCR, and targeting the amplicons with gRNA-N3 (Table S6). From the analysis we found that 1µL and 10µL RT-RPA volume added resulted in the two fastest detection times, respectively (1µL: HMF = 19.23 min; 10µL: HMF = 30.3 min) (Fig S7C). However, upon further analysis, we found 1µL RT-RPA volume added resulted in significantly lower signal over noise compared to 5µL, 7.5µL, and 10µL volume added (5µL: p=0.0131, 7.5µL: p=0.0021, 10µL: p=0.0001) (Fig S7D). We therefore selected 10µL RT-RPA volume added as our optimized condition for a standard protocol.

In an effort to identify the optimal time to detection for SENSR, we further evaluated the system by varying the amplification time and assessing the effect on time to detection. We tested 3 different amplification times (15 min, 30 min, and 45 min) using the best candidate guides identified in the preliminary analysis (gRNA-S3 and gRNA-N1). From the analysis we observed gRNA-S3 resulted in rapid detection regardless of amplification time where the 30 min amplification group demonstrated the fastest detection time tested (HMF = 4.22 min) (Fig S8A). We also found that the fastest detection time for gRNA-N1 was after 30 min of amplification (HMF = 13.25 min) (Fig S8B). To confirm that 30 min is the best amplification time regardless

of gRNA used, we further evaluated the fluorescence signal increase for each amplification group. We found no significant difference in fluorescence signal between all amplification times for gRNA-S3 (15m vs 30m: p=0.5576, 15m vs 45m: 0.9935, 30m vs 45m: p=0.2735), however, for gRNA-N1, fluorescence signal in the 15 min amplification group was significantly lower than 30 min and 45 min amplification times (15m vs 30m & 15m vs 45m: $p<0.0001$) (Fig S8C-D). We therefore selected 30 min amplification as the standard time of amplification for future fluorescence experiments.

3.3.4 Fluorescence limit of detection

Following optimization of SENSR fluorescence reaction parameters, we moved to determine the limit of detection (LoD) for gRNA-S3 and gRNA-N1 via serial dilution of the synthetic RNA templates from 10,000 to 0 copy/µL added directly into RT-RPA reactions. For both gRNA-S3 and gRNA-N1 we observed a significant increase in fluorescence signal above noise for 10,000, 1,000 and 100 copies compared to the NTC (Fig 2A-B). Interestingly, we also observed significantly higher fluorescence at 10 copies for gRNA-S3 (Fig 2A), however this signal increase was not comparable to the increase observed in higher copy groups (Dunnett's multiple comparison test; 10,000, 1,000, 100 copies vs 10: $p<0.0001$). We therefore determined the LOD for synthetic RNA to be 100 copy/µL for both gRNA-S3 and gRNA-N1, indicating SENSR exhibits attomolar analytical sensitivity comparable to other CRISPRDx systems (Gootenberg et al. 2017; Chen et al. 2018). These results demonstrate that CasRx can robustly detect and report the presence of synthetic SARS-CoV-2 RNA via fluorescence readout.

3.3.5 Lateral flow assay development

Collateral cleavage by CasRx can additionally be exploited to detect SARS-CoV-2 RNA by LF readout which facilitates detection by simple paper test strips and eliminates the need for expensive laboratory equipment (Fig 1A, Fig S3). Similar to previously developed assays (Kellner et al. 2019; Gootenberg et al. 2018), we generated a 6nt ssRNA probe modified with 5' 6-FAM and 3' biotin (Bio) compatible with Millenia HybriDetect LF strips (Table S4). In brief, collateral cleavage results in separation of 6-FAM from biotin, detectable following capillary action down a paper dipstick imprinted with streptavidin and anti-FAM secondary antibodies at distant ends (Fig 1A). An increase in saturation of the upper band therefore indicates a negative or positive result, respectively (Summarized in Fig 1A, Fig S3).

LF protocols only allow one time-point for assessing detection and we therefore optimized the SENSR protocol by varying the amplification and cleavage times for our LF assay. We amplified the synthetic SARS-CoV-2 template for 15, 30, or 45 min and then proceeded to incubate the CCR for 15, 30, or 45 min. We observed cleavage of the LF probe at all amplification times in all cleavage times tested with clear evidence that extended cleavage time enhances detection, while extended amplification time does not have a dramatic effect on detection (Fig S9). In an effort to keep the total processing time within 1 hr we chose 30 min amplification and 30 min cleavage for our standard protocol. This decision was driven by fluorescence evidence suggesting gRNA-N1 requires at least 30 min of amplification to increase the rate of probe cleavage for sufficient detection at lower copy numbers (Fig S8B).

Using this optimized protocol, we then determined the LoD of SENSR in a LF assay format by performing a serial dilution of the synthetic RNA templates from 10,000 to 0 copy/μL added directly into RT-RPA reactions. We determined the LoD for each gRNA detecting

synthetic RNA to be 100 copy/µL, where gRNA-S3 exhibited significantly more probe cleavage compared to gRNA-N1 (Fig 2C-D, Supplementary Video 1). These results confirm that SENSR, like other CRISPRDx systems, can be adapted for readout by LF, indicating the potential for future point-of-care applications.

3.3.6 SENSR detection of SARS-CoV-2 from patient isolates

Previously, we analyzed a set of RT-qPCR verified patient samples using a preliminary protocol for SENSR and targeting a different segment of the N gene (Table S4, S6-S7). We observed only moderate detection of patient samples with 92% of samples detected for $C_t \leq 20$. The highest C_t value detected was 27.5 and we observed a 23% and 28% false-negative rate for $C_t \leq 27.5$ for fluorescence and lateral flow, respectively (Fig S10). Furthermore, this version of SENSR took nearly 2 hrs to obtain results from both readouts. Therefore, we further optimized SENSR and improved the sensitivity and time to detection as previously described (Figs S7-S9).

Following optimization, we set to determine the ability of SENSR to detect SARS-CoV-2 from clinically verified patient samples. Samples were analyzed by RT-qPCR targeting the N-, S-, and Orf1ab-genes (Table S7), and accordingly, we selected gRNA-S3 and gRNA-N1 to directly compare detection by SENSR to RT-qPCR. We performed fluorescence and LF analysis on 60 total samples: 30 positive (n = 30) (C_t range: 13-34) and 30 negative samples. The C_t values of the positive patient samples were distributed as follows: 13.3% (4) between <16, 23.3% (7) between 17-22, 16.7% (5) between 23-28, and 46.7% (14/30) between 29-34.

From the patient samples tested, we observed considerably different results from each gRNA applied. gRNA-N1 performed less effectively than gRNA-S3 detecting only 17/30

positive samples by fluorescence and 16/30 samples by LF (Fig S11, Table S7), whereas gRNA-S3 detected 21/30 positive samples by fluorescence, 20/30 samples by LF (Fig 3, Table S7). gRNA-S3 exhibited a 0% false-positive rate in both read outs, while gRNA-N1 resulted in a 3.3% false-positive rate by fluorescence read out (Fig 3, Figs S11-S12). Furthermore, each gRNA resulted in different detection limits based on copy number values extrapolated from the RT-qPCR standard curve (Fig S13). To elaborate, gRNA-S3 detected SARS-CoV-2 in 100% of samples (18/18) by both readouts for samples with a $C_t \leq 29.83$ (45 copy/µL), whereas gRNA-N1 detected vRNA in 94% (16/17) and 88% (15/17) of samples for fluorescence and LF, respectively, for samples with $C_t \leq 29.28$ (87 copy/µL) (Fig 3, Fig S11A, Fig S13, Table S7).

LF-based detection methods are limited to one detection point, however fluorescence data are acquired multiple times during incubation of the CCR. Therefore, we further evaluated the fluorescence data to find the time each sample crossed the Signal-to-Noise (S/N) threshold, thus determining a positive result. We found gRNA-S3 detected SARS-CoV-2 in patient samples between 30.58-40.98 min total reaction time (RT-RPA + CCR) and on average detected vRNA within 33.05 min. gRNA-N1 detected vRNA in patient samples between 30.76-59.86 min total reaction time 34.27 min on average (Table S8). Taken together these results demonstrate SENSR rapidly detects SARS-CoV-2 from patient isolates and exhibits sensitivity comparable to previously established CRISPRDx systems (Patchsung et al. 2020).

3.4 Discussion

Complementing the rapidly expanding CRISPRDx toolkit, here we outline the use of CasRx to detect SARS-CoV-2. The development of SENSR establishes proof-of-principle that

Cas13d ribonucleases can be adapted for nucleic acid diagnostics. Identifying targets conserved and specific to SARS-CoV-2, we applied SENSR and demonstrated detection with attomolar sensitivity by both fluorescence and lateral flow readouts comparable to previously developed systems (Zhang, Abudayyeh, and Jonathan 2020; Broughton, Deng, Yu, Fasching, Singh, et al. 2020; Patchsung et al. 2020). Furthermore, we determined the CasRx collateral cleavage sequence preference for uracil stretches over all other di-nucleotide combinations, demonstrating promise for multiplexing applications (Gootenberg et al. 2018; Kellner et al. 2019). Together these results demonstrate and provide evidence that SENSR can be adapted and further optimized for detection of a multitude of pathogens (Fig S14).

Comparatively, SENSR may present a worthy advancement for the CRISPRDx toolkit due to the fundamental properties of Cas13d effectors. Similar to LwaCas13a used in SHERLOCK systems, Cas13d effectors are more flexible than most Cas enzymes as they lack a protospacer flanking sequence (PFS) requirement (Konermann et al. 2018; Yan et al. 2018; Freije et al. 2019), permitting targeting of any sequence. Moreso, we demonstrate SENSR exhibits comparable sensitivity to SHERLOCK (Patchsung et al. 2020), suggesting CasRx as a viable alternative to LwaCas13a for nucleic acid detection. CasRx also presents great potential for multiplexing applications with other Cas13 effectors due to the uracil-specific collateral cleavage activity (Gootenberg et al. 2018). Due to the stringent criteria required for diagnostics, further optimization in advance of deployment is necessary. However, the flexible targeting, multiplexing capabilities, and sensitivity of SENSR present a potentially viable technology for disease diagnosis and establishes an alternative Cas effector subtype for nucleic acid detection.

SENSR presents great promise for future applications, however improvements to the system may be necessary. For instance, RT-RPA is less sensitive than RT-LAMP (Howson et al.

2017), and therefore, alternative isothermal amplification methods could further improve sensitivity. Although RT-LAMP is likely to improve sensitivity, the requirement of two separate reactions remains problematic, as this increases the likelihood of contamination due to sample transfers (Sefers and Schmitz 2018). Direct detection of RNA with CasRx could reduce the need for multiple transfers, as has been demonstrated with LbuCas13a (Fozouni et al. 2021). However, CasRx direct detection only demonstrated picomolar sensitivity (Fig S15) whereas SENSR exhibits attomolar sensitivity; a significant increase in sensitivity largely due to the preamplification via RT-RPA. Furthermore, considerable variation in sensitivities have been reported for different Cas13 enzymes and detection accuracy of LbuCas13a at 100 copy/μL drops 20% using direct detection in a POC format (Fozouni et al. 2021). Lastly, differences in gRNA sensitivities were observed in our analysis of infected patient samples and thus improvements to gRNA screening and selection could expedite the response time to future disease outbreaks.

Optimizing workflow, deployment, and distribution, while taking steps to reduce the risk of contamination, is imperative to develop SENSR to its full potential. We demonstrate that detection with CasRx is a viable alternative for nucleic acid detection, and could be further improved to become a powerful molecular diagnostic with numerous applications. Establishing these tools and frameworks now could expedite response times and help prevent future outbreaks, avoiding the economic and health consequences which have resulted from poor preparedness to the current pandemic.

3.5 Methods

3.5.1 *In vitro* gRNA cleavage assays

Preliminary *in vitro* cleavage cleavage assays were prepared with RNA templates for S-gene (1000ng) or N-gene (1200ng), followed by addition of CasRx (110.8ng) and 10ng of each gRNA in a 2:1 molar ratio. Reactions were prepared in 20mM HEPES pH 7.2 and 9mM $MgCl_2$, incubated at 37ºC for 1 hr, then denatured at 85ºC for 10 min in 2X RNA loading dye (New England Biolabs, #B0363) and loaded on 2% TBE agarose gel stained with SYBR gold nucleic acid stain (Invitrogen #S11494).

3.5.2 RT-RPA amplification of viral genomic sequences

RT-RPA primers were designed to amplify 30bp gRNA spacer regions flanked by 30bp priming regions from the synthetic vRNA template while also incorporating a T7 promoter sequence into the 5' end of the dsDNA gene fragment with +2 G's thus increasing transcription efficiency (Fig 1B) (Brieba, Padilla, and Sousa 2002). RT-RPA was performed at 42ºC for 30 min by combining M-MuLV-RT (NEB #M0253L) with TwistAmp Basic (TwistDx #TABAS03KIT). All RT-RPA primer sequences can be found in Table S1. Final conditions for optimized (28.5% sample input) RT-RPA protocol in 10µL reaction: 5.9µL rehydration buffer, 0.35µL primer mix (10µM each), 0.4µL RT (200U/µL), 0.5µL MgOAc (280mM), 2.85µL vRNA.

3.5.3 Fluorescence-based detection of SARS-CoV-2

We developed an IVT-coupled cleavage assay with a fluorescence readout using 6-Carboxyfluorescein (6-FAM) as our fluorescent molecule. We developed a 6nt poly-U probe conjugated to a 5' 6-FAM and a 3' IABlkFQ (FRU, Table S4) and custom ordered from IDT. Probes with variable di-nucleotide repeat sequences were ordered from IDT and can be found in

Table S4. Direct detection experiments were performed in 20µL reactions as follows: 13.32µL water, 0.4µL HEPES pH 7.2 (1M), 0.24µL MgCl2 (1M), 2µL CasRx (55.4ng/µL), 1µL RNase inhibitor (40U/µL), 1µL gRNA (10ng/µL), and 1µL FRU probe (2µM). Optimized fluorescence protocol (10µL): 2.82µL water, 0.4µL HEPES pH 7.2 (1M), 0.18µL MgCl2 (1M), 1µL rNTPs (25mM each), 2µL CasRx (55.4ng/µL), 1µL RNase inhibitor (40U/µL), 0.6µL T7 Polymerase (50U/µL), 1µL gRNA (10ng/µL), and 1µL FRU probe (2µM). This was followed by the addition of 10µL (100% RT-RPA reaction vol) of the RT-RPA pre-amplification reaction (described above), which initiates detection following incubation at 37ºC for 60 min. Experiments were immediately run on a LightCycler96 (Roche #05815916001) at 37ºC. Acquisition protocol: 5 sec acquisition followed by 5 sec incubation for the first 15 min, followed by 5 sec acquisition and 55 sec incubation for up to 45 min. Fluorescence readouts were analyzed by normalization to NTC at each respective time point or through background subtracted fluorescence by subtracting the initial fluorescence value from the final value.

3.5.4 Lateral flow detection of SARS-CoV-2

For lateral flow-based detection we modified the HybriDetect system to detect the presence of SARS-CoV-2 sequences using SENSR (Zhang, Abudayyeh, and Jonathan 2020). In brief, we designed a ssRNA probe composed of a 6nt poly-U probe conjugated on opposite ends with a 5’ 6-FAM and a 3’ biotin which was custom ordered from IDT (LFRU, Table S4). Following incubation of 2.82µL water, 0.4µL HEPES pH 7.2 (1M), 0.18µL $MgCl_2$ (1M), 2µL CasRx (55.4ng/µL), 1µL gRNA (10ng/µL), 10µL RT-RPA reaction mix, 0.6µL T7 RNA polymerase (50U/µL), 1µL rNTPs (25mM each), 1µL LFRU probe (20µM), at 37ºC for 30 min. 80µL of HybriDetect Assay buffer was added to each reaction. Next, the lateral flow dipstick

was placed into the reaction and allowed to flow upwards by capillary action for a maximum of 3 min. The presence or absence of upper or lower bands was analyzed to detect evidence of SARS-CoV-2 by collateral cleavage. The presence of a solitary upper band or both an upper and lower band was interpreted as a positive result, a solitary lower band with a faint upper band was interpreted as a negative result.

3.5.5 LoD analysis

For SENSR LoD analysis against synthetic RNA, synthetic templates were serially diluted on a logarithmic scale. Template stock concentrations were analyzed via nanodrop prior to dilutions. Dilution scales were calculated using NEBioCalculator for each respective template. For fluorescence analysis, the LoD was determined by statistical significance of the lowest copy number experimental group compared to the NTC. For lateral flow analysis, the LoD was determined by noticeable saturation of the upper test band compared to the NTC.

LoD for RT-qPCR was performed by spiking viral particles into verified SARS-CoV-2 negative samples contained in viral transport media (VTM) at specified copy numbers. 2µL of extracted sample were run for the assay therefore all copy/µL values must be multiplied by 2 (Fig S13A).

SENSR detection limit in patient samples was calculated by performing nonlinear regression analysis and fitting a semi-log function ($y = b + m*log(x)$, where b is the y-intercept and m is the slope) to the data where x values are treated as log and y values linear (Fig S13A). We then solved for x to generate a function for calculating the copy/µL based on C_t value ($x = 10^{(y-b)/m}$). Copy/µL values were then calculated by entering the C_t value of each sample for y and multiplying x by 2 (for each µL added).

3.5.6 SENSR detection of patient samples

Fluorescence and LF assays were run against SARS-CoV-2 positive (N = 30) and negative (N = 30) samples. Samples were amplified for 30 min RT-RPA reaction at 28.5% v/v. 10μL of RT-RPA reaction was then incubated in an CCR reaction as previously described. Fluorescence data for analysis were acquired on LightCycler96 (Roche #05815916001) following the protocol previously described. Data were processed by generating background subtracted fluorescence for each replicate by subtracting the final (30 min) fluorescence value from the initial (0 min) fluorescence value. Noise was set as the average of the four NTC background subtracted values. S/N was then calculated by dividing the background subtracted value for each sample by the noise. To account for 99.9% of expected negative sample S/N measurements, the S/N threshold was set at $\mu+3.5\sigma$ of the negative samples. Samples were determined to be positive if S/N > 2.22 for gRNA-S3 and positive if S/N > 3.51 for gRNA-N1. Time to detection was calculated by calculating S/N at each time-point where N is equal to the average of NTC after 30 min.

Lateral flow analysis was run in parallel to fluorescence analysis. The samples were amplified via RT-RPA for 30 min at 28.5% v/v. 10μL of RT-RPA then transferred to the cleavage reaction with LFRU and incubated at 37ºC for 30 min. LF analysis was then performed with 3 min incubation of strips. Images were taken using an iPhone 12 mini. Positives and negatives were determined in comparison to the NTC samples and using a positive control (synthetic template) as a standard.

3.6 Acknowledgements

This chapter, in full, is a reprint of the material as it appears in ACS Sensors. Daniel J Brogan[*], Duverney Chaverra-Rodriguez[*], Calvin P Lin[*], Andrea L Smidler, Ting Yang, Lenissa M. Alcantara, Igor Antoshechkin, Junru Liu, Robyn R Raban, Pedro Belda-Ferre, Rob Knight, Elizabeth A Komives, Omar S. Akbari. “Development of a Rapid and Sensitive CasRx-Based Diagnostic Assay for SARS-CoV-2”. ACS Sensors. (2021). [* = Equal contribution]. The dissertation/thesis author was the first-author and primary investigator of this paper.

3.7 Figures

Figure 3.1: Harnessing SENSR to detect evidence of SARS-CoV-2 viral transcripts. [A] Overview of assay workflow. Following extraction of viral RNA, the detection protocol requires two distinct reactions and the readout may be performed in two different methods: fluorescence or lateral flow. In the first reaction, specific target sequences within the viral RNA are reverse transcribed (RT) into cDNA and amplified by RPA at 42°C for 30 min. During amplification a T7 promoter is incorporated into the 5' terminus of the amplicons (T7, purple). In the next reaction, both *in vitro* transcription (IVT) and CasRx collateral cleavage occur simultaneously (pink). Recognition and cleavage of the target RNA sequence (purple) complementary to the gRNA (emerald) induces collateral cleavage of bystander RNA molecules. Collateral cleavage of a modified probe conjugated to 6-FAM and a fluorescence quencher facilitates readout by fluorescence (top right). Collateral cleavage of a modified probe conjugated to 6-FAM and biotin facilitates readout by lateral flow assay (bottom right). [B] Schematic of the SARS-CoV-2 genome, RPA reagents, gRNA target sites, and synthetic amplicon position. The S (*Spike*) and N (*Nucleocapsid*) genes are enlarged to depict the design schematic in more detail (sage and blue respectively). The schematic depicts the relative position of the RT-RPA primers and gRNAs used in the amplification and cleavage reactions, respectively. The gRNA target map outlines all six gRNAs tested through the course of this work, with their relative positions. The S-gene is 3,822nt in length while the N-gene is 1,260nt in length. Synthetic target amplicon denotes the length of the synthetic viral genome fragment used in SENSR assays and its relative position within the gene coding sequence. [C] Preliminary characterization of S-targeting gRNAs (gRNA-S1, -S2, and -S3) in a SENSR fluorescence assay. Synthetic *Spike* gene templates (10,000 copy/μL) were subjected to RT-RPA for each respective gRNA tested and then transferred into a CCR. Data were acquired for 60 min and plotted by normalizing individual experimental data points to averaged data of controls at each time point. Half-Maximum Fluorescence (HMF) was then calculated to determine which gRNA resulted in the fastest cleavage of the fluorescent probe. Intersecting lines represent the location of HMF for each group. [D] Preliminary characterization of N-targeting gRNAs (gRNA-N1, -N2, and -N3) in a SENSR fluorescence assay. Synthetic *Nucleocapsid* gene templates (10,000 copy/μL) were subjected to RT-RPA for each respective gRNA tested and then transferred into a CCR. Data were acquired, plotted, and analyzed identical to the method described for S-targeting guides. Intersecting lines represent the location of HMF for each group.

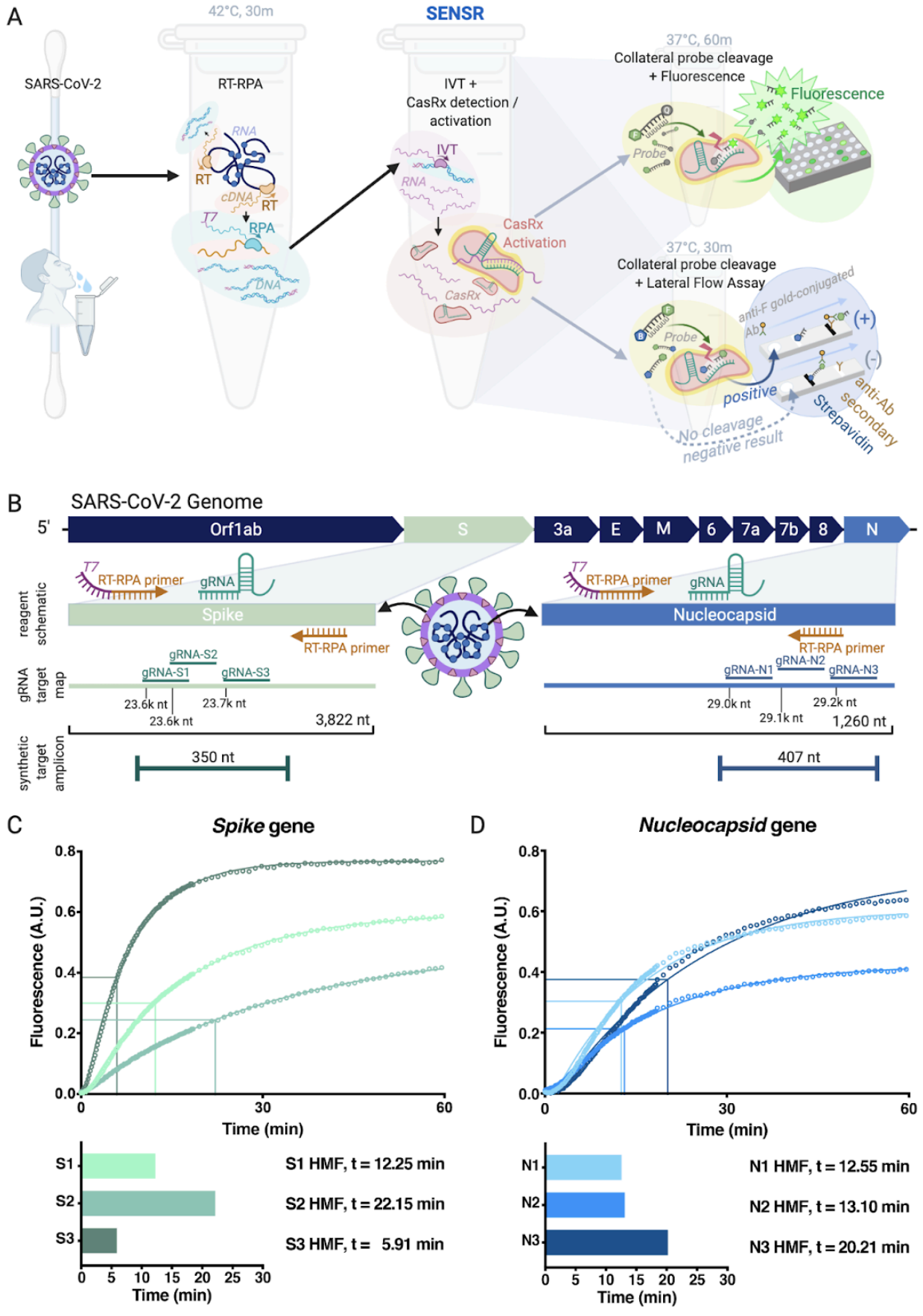
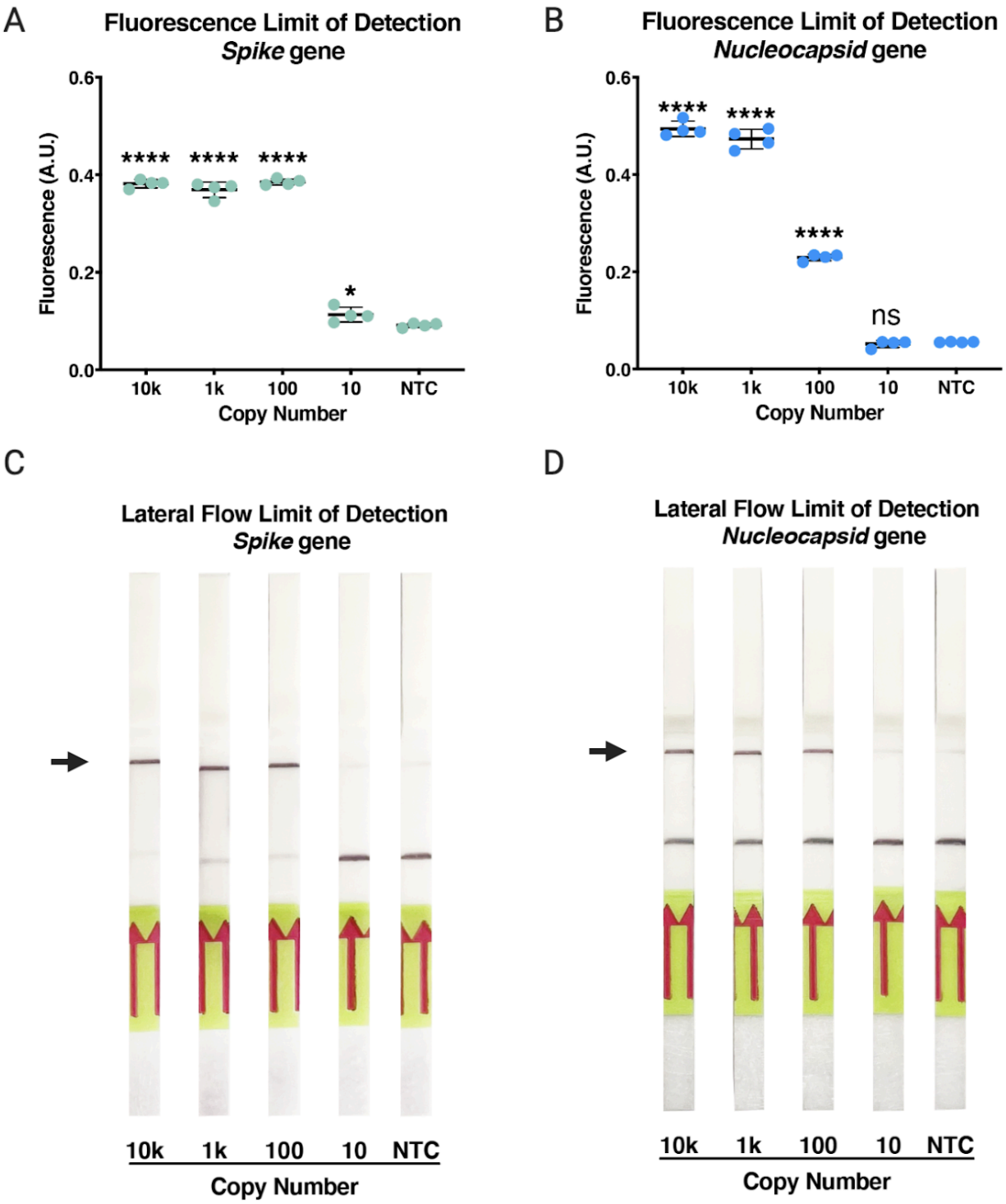

A
42°C, 30m
SENSR
SARS-CoV-2
RT-RPA
IVT +
CasRx detection /
activation
37°C, 60m
Collateral probe cleavage
+ Fluorescence
Fluorescence
RNA
RT
cDNA
RT
T7
RPA
DNA
IVT
RNA
CasRx
Activation
CasRx
Probe
37°C, 30m
Collateral probe cleavage
+ Lateral Flow Assay
anti-F gold-conjugated
Ab
(+)
(-)
Probe
positive
anti-Ab
secondary
Strepavidin
No cleavage
negative result
B
SARS-CoV-2 Genome
5'
Orf1ab
S
3a
E
M
6
7a
7b
8
N
reagent schematic
T7
RT-RPA primer
gRNA
Spike
RT-RPA primer
T7
RT-RPA primer
gRNA
Nucleocapsid
RT-RPA primer
gRNA target map
gRNA-S2
gRNA-S1
gRNA-S3
23.6k nt
23.7k nt
23.6k nt
3,822 nt
gRNA-N1
gRNA-N2
gRNA-N3
29.0k nt
29.2k nt
29.1k nt
1,260 nt
synthetic target amplicon
350 nt
407 nt
C
Spike gene
Fluorescence (A.U.)
Time (min)
S1
S2
S3
S1 HMF, t = 12.25 min
S2 HMF, t = 22.15 min
S3 HMF, t = 5.91 min
Time (min)
D
Nucleocapsid gene
Fluorescence (A.U.)
Time (min)
N1
N2
N3
N1 HMF, t = 12.55 min
N2 HMF, t = 13.10 min
N3 HMF, t = 20.21 min
Time (min)

Figure 3.2: SARS-CoV-2 detection by SENSR via Fluorescence and Lateral Flow assay. [A] Analysis of limit of detection (LoD) for gRNA-S3 targeting the *Spike* gene. Synthetic S-gene templates were diluted on a logarithmic scale from 10,000 copy/µL to 10 copy/µL. Fluorescence signal plotted represents the final data acquired at 60 min subtracted by the initial fluorescence signal recorded. A.U., Arbitrary Units. Data represents mean ± s.d. from quadruplicate measurements. Significance is representative of Dunnett's multiple comparisons test for experimentals compared to NTC; (10k, 1k, 100) $p<0.0001$, (10) $p=0.0461$). [B] Analysis of limit of detection (LoD) for gRNA-N1 targeting the *Nucleocapsid* gene. Synthetic N-gene templates were diluted on a logarithmic scale from 10,000 copy/µL to 10 copy/µL. Fluorescence signal plotted represents the final data acquired at 60 min subtracted by the initial fluorescence signal recorded. A.U., Arbitrary Units. Data represents mean ± s.d. from quadruplicate measurements. Significance is representative of Dunnett's multiple comparisons test for experimentals compared to NTC; (10k, 1k, 100) $p<0.0001$). [C] LoD analysis of lateral flow (LF) for gRNA-S3 targeting the *Spike* gene. Synthetic S-gene templates were diluted on a logarithmic scale from 10,000 copy/µL to 10 copy/µL. CasRx detection reaction incubated for 30 min prior to addition of dipstick and imaging results. Positive results were indicated by increased saturation of the top band (arrow). [D] LoD analysis of Lateral Flow (LF) for gRNA-N1 targeting the *Nucleocapsid* gene. Synthetic N-gene templates were diluted on a logarithmic scale from 10,000 copy/µL to 10 copy/µL. CasRx detection reaction incubated for 30 min prior to addition of dipstick and imaging results. Positive results were indicated by increased saturation of the top band (arrow).


A
Fluorescence Limit of Detection
Spike gene
Fluorescence (A.U.)
0.6
0.4
0.2
0.0
****
****
****
*
10k
1k
100
10
NTC
Copy Number
B
Fluorescence Limit of Detection
Nucleocapsid gene
Fluorescence (A.U.)
0.6
0.4
0.2
0.0
****
****
****
ns
10k
1k
100
10
NTC
Copy Number
C
Lateral Flow Limit of Detection
Spike gene
10k
1k
100
10
NTC
Copy Number
D
Lateral Flow Limit of Detection
Nucleocapsid gene
10k
1k
100
10
NTC
Copy Number

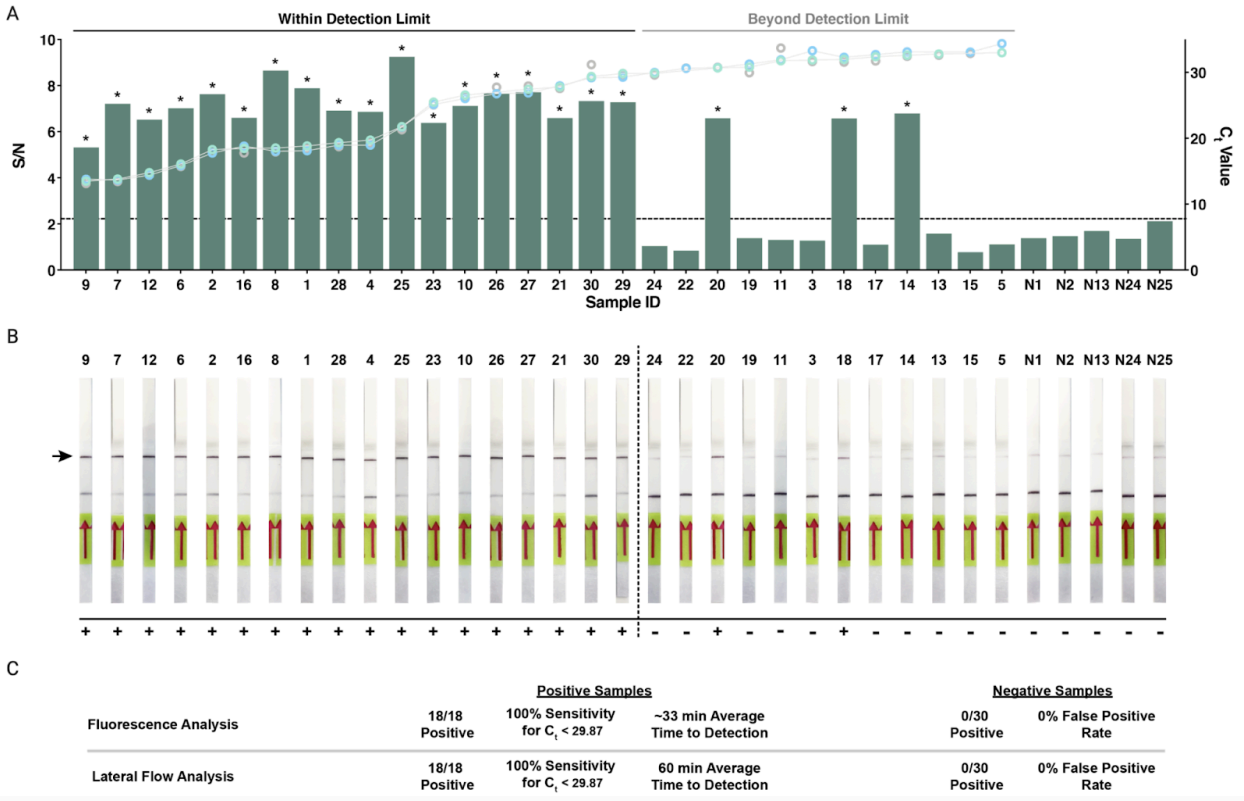
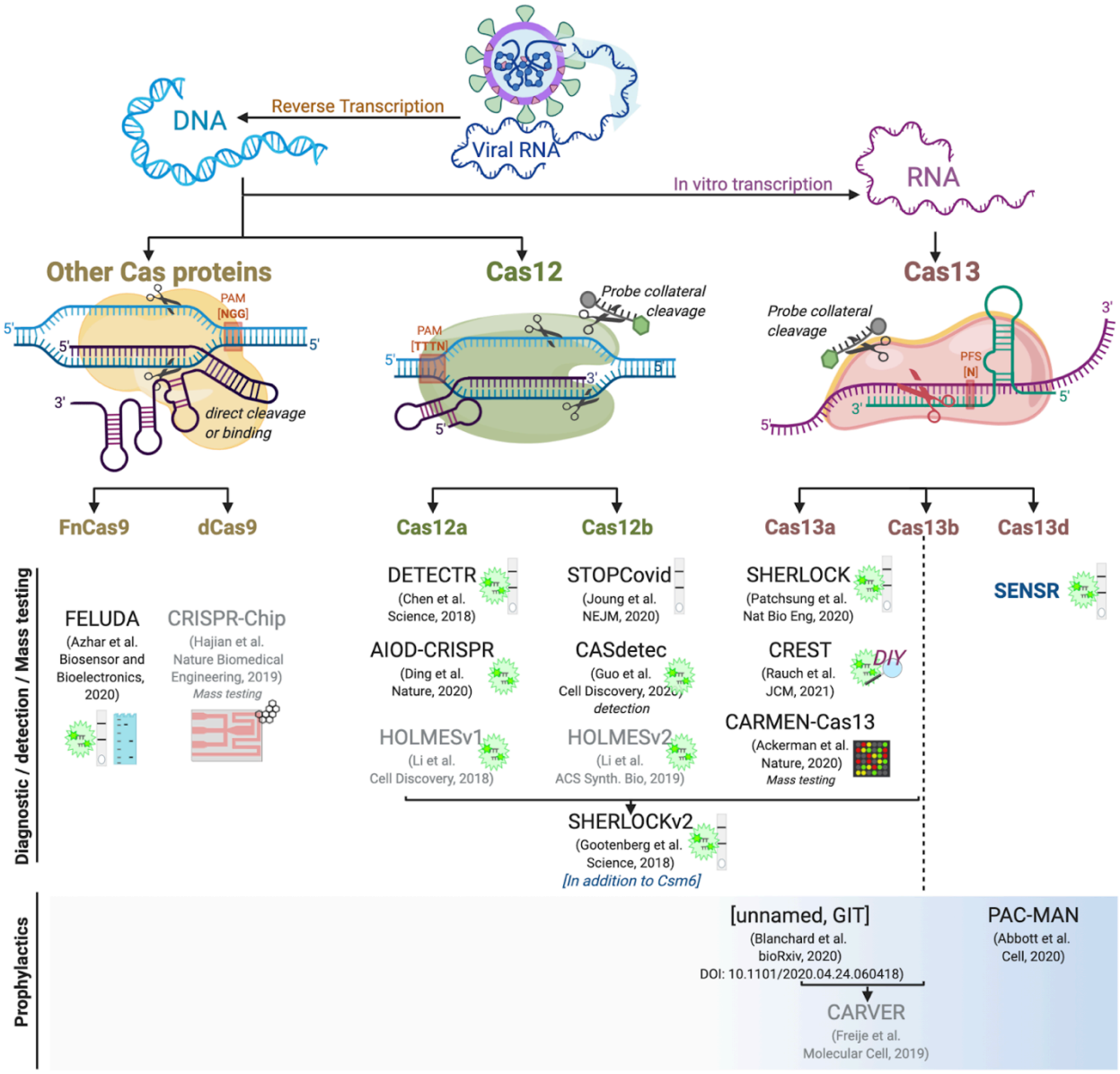


Figure 3.3: SENSR detection of positive SARS-CoV-2 validated patient samples. [A] SENSR fluorescence analysis of RT-qPCR validated patient samples using gRNA-S3 for detection. Signal-to-Noise (S/N) ratio of 30 positive and 5 negative samples are shown. Positive samples are listed in order from lowest to highest C_t value for S gene. Open circles represent C_t value of S- (light green), N- (blue), and Orf1ab (gray). The five negative samples are negative samples with the highest calculated S/N ratio among all negative samples analyzed. Dashed line represents the S/N threshold = 2.22. Asterisks represent a S/N > 2.22. [B] Lateral flow based detection of the 30 positive and 5 negative samples. The top band (arrow) represents the test band and the bottom the control band. An increase in saturation of the top band indicates a positive detection of SARS-CoV-2 in the sample. Vertical dashed line indicates the detection limit. Result of detection indicated below each test strip. [C] Summary of fluorescence and lateral flow detection results for the RT-qPCR validated positive and negative patient samples.

3.8 Supplementary Figures

Figure S3.1: Summary of current CRISPR-based anti-Covid technologies organized by Cas enzyme used and by role as diagnostic or detection tool, or as a putative prophylactic. Those technologies shown in black font have been demonstrated to have explicit activity against SARS-CoV-2, while those technologies shown in grey font have been publicly discussed or proposed to be technological candidates for SARS-CoV-2 detection or prophylaxis, however have not yet been fully demonstrated/optimized for said purpose as of the publication of this manuscript. Some technologies have not yet been given a formal name by their authors, and are therefore denoted as 'unnamed' followed by the acronym for the primary institutional affiliation behind the work. To better identify and distinguish these technologies, the DOI has been provided. [Diagnostics/detection systems/Mass testing] For all technologies viral RNA is extracted, reverse transcribed into cDNA, followed by template amplification by either PCR, RPA or LAMP, then input into subsequent reactions with the exception of CRISPR-Chip which does not require an amplification step. Cas12-based enzymes, as well as many other Cas proteins (including Cas9) recognize DNA species, while Cas13-based enzymes recognize ssRNA, and all can be used to detect evidence of specific sequences by fluorescence or lateral flow. The detection method for each technology is noted with the presence of an icon adjacent. The majority of the technologies summarized here use fluorescence or lateral flow or both. The lime glow icon indicates readout by fluorescence, while the grey bar indicates read-out by lateral flow. Some of the other technologies can be read out by different detection methods. Notably because FELUDA relies on direct sequence cleavage, and not the collateral cleavage property shared amongst the Cas12- and Cas13-based technologies, it can also be analyzed by gel electrophoresis as evidence of distinct band cleavage (aqua gel icon). Also, CRISPR-Chip has been discussed as a SARS-CoV-2 mass detection candidate, though these findings have not yet been made publicly available. Read-out by this technology is achieved via a graphene-based transistor (icon is red, grey with graphene structure adjacent). Other different technologies include CREST, which achieves detection via fluorescence using a distinctly do-it-yourself (DIY) BioHacking approach, by utilizing equipment easily procured and operated by novice scientists, as well as CARMEN-Cas13 which detects evidence of SARS-CoV-2 sequences by microarray. The technology outlined here is an addition to current SHERLOCK detection utilizing Cas13d (CasRx). The technical details of most technologies is summarized in Table S1.


DNA
Reverse Transcription
Viral RNA
In vitro transcription
RNA
Other Cas proteins
Cas12
Cas13
PAM
[NGG]
direct cleavage
or binding
PAM
[TTTN]
Probe collateral
cleavage
Probe collateral
cleavage
PFS
[N]
FnCas9
dCas9
Cas12a
Cas12b
Cas13a
Cas13b
Cas13d
Diagnostic / detection / Mass testing
FELUDA
(Azhar et al.
Biosensor and
Bioelectronics,
2020)
CRISPR-Chip
(Hajian et al.
Nature Biomedical
Engineering, 2019)
Mass testing
DETECTR
(Chen et al.
Science, 2018)
AIOD-CRISPR
(Ding et al.
Nature, 2020)
HOLMESv1
(Li et al.
Cell Discovery, 2018)
STOPCovid
(Joung et al.
NEJM, 2020)
CASdetec
(Guo et al.
Cell Discovery, 2020)
detection
HOLMESv2
(Li et al.
ACS Synth. Bio, 2019)
SHERLOCK
(Patchsung et al.
Nat Bio Eng, 2020)
CREST
(Rauch et al.
JCM, 2021)
DIY
CARMEN-Cas13
(Ackerman et al.
Nature, 2020)
Mass testing
SENSR
SHERLOCKv2
(Gootenberg et al.
Science, 2018)
[In addition to Csm6]
Prophylactics
[unnamed, GIT]
(Blanchard et al.
bioRxiv, 2020)
DOI: 10.1101/2020.04.24.060418)
CARVER
(Freije et al.
Molecular Cell, 2019)
PAC-MAN
(Abbott et al.
Cell, 2020)

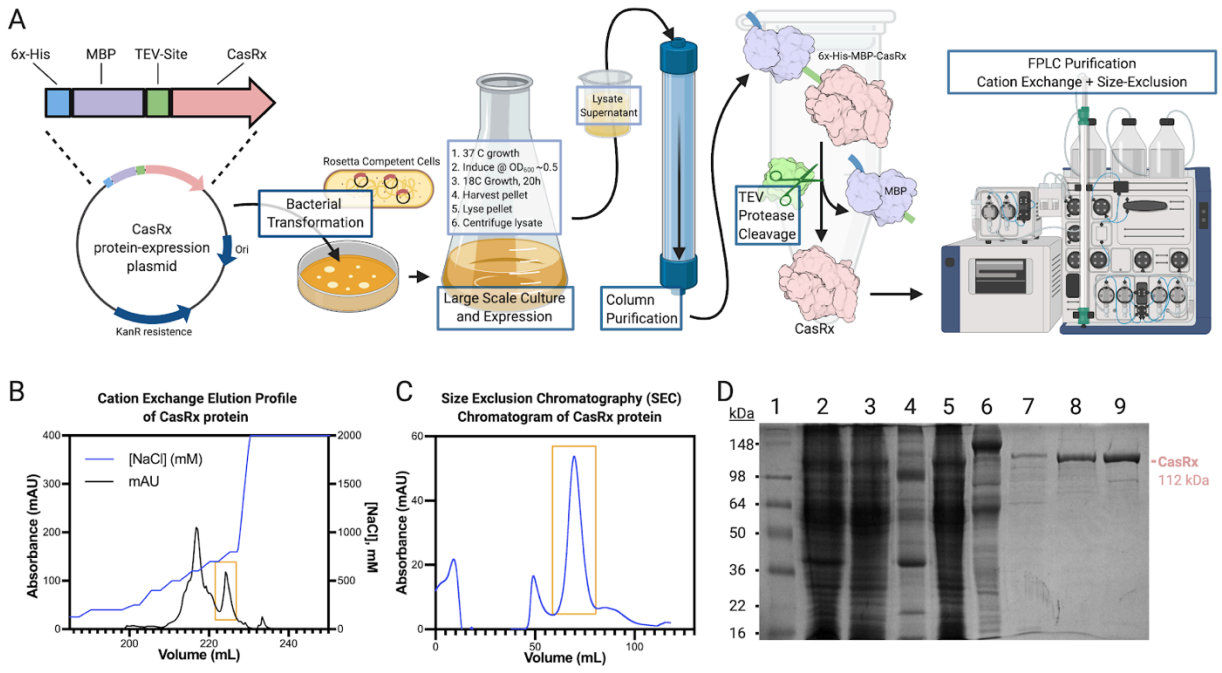
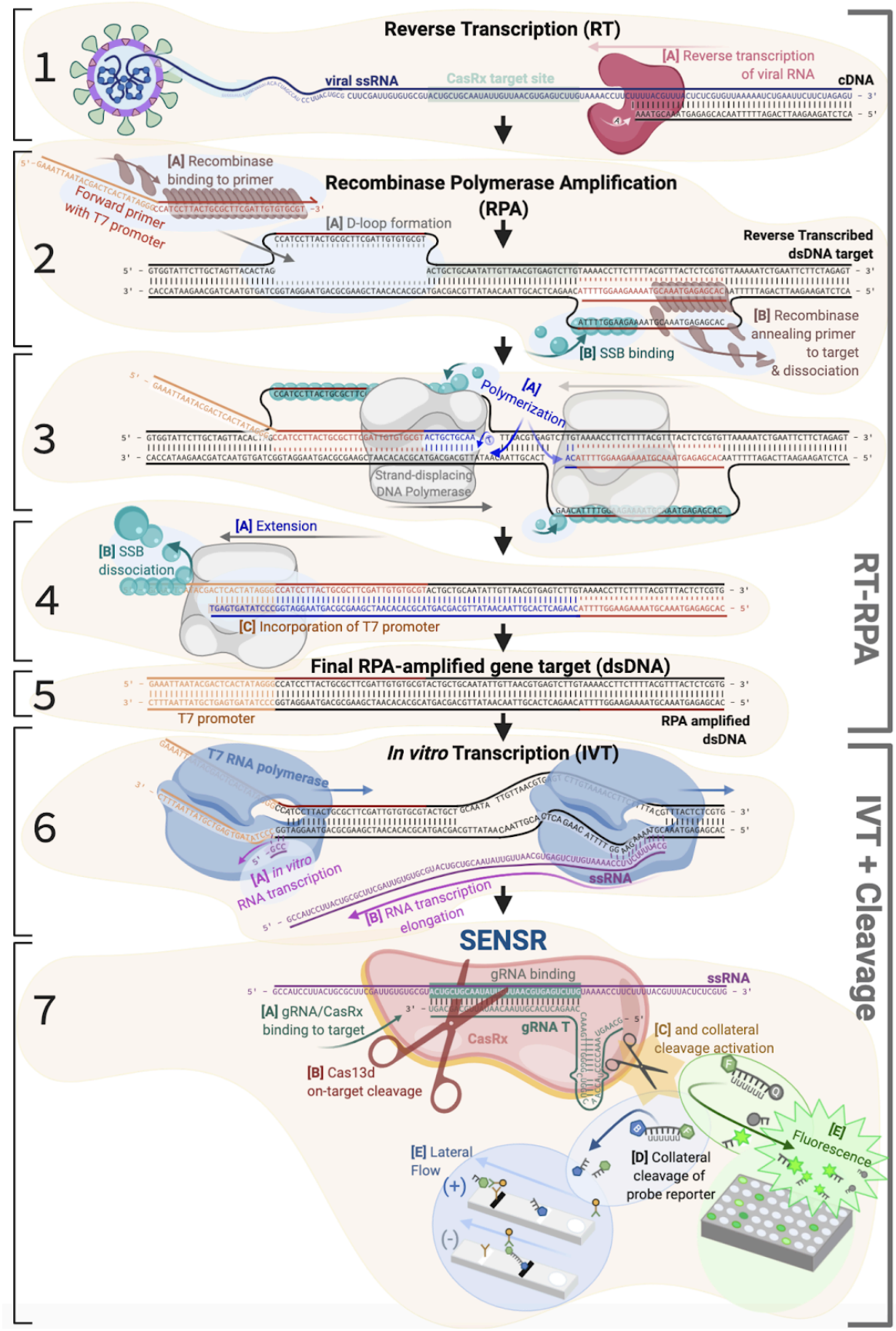


Figure S3.2: CasRx protein purification workflow and quality control. [A] CasRx protein was produced and purified essentially as described in (Konermann et al. 2018). A CasRx protein-expression plasmid was generated with CasRx (pink) downstream of a Maltose-binding protein (MBP, purple) domain with an N-terminal 6xHis tag (blue), connected by a TEV protease cleavage sequence (green). Origin of replication and KanR cassette shown in navy. Transformation of the plasmid into Rosetta2 (DE3) Competent *E. coli* was followed by large scale culture growth. Cell lysate supernatant containing 6xHis-MBP-CasRx soluble hybrid protein was run on a Ni-NTA column for purification by affinity chromatography. CasRx protein was then released from the 6xHis-MBP moiety by TEV protease cleavage during O/N dialysis, and was further purified by cation exchange and size exclusion using Fast Protein Liquid Chromatography (FPLC). [B] The cation exchange elution profile of CasRx with the concentration of NaCl (mM) shown in blue. Peak containing CasRx recombinant protein (boxed in gold) elutes at ~700 mM NaCl. [C] The SEC elution profile of CasRx recombinant protein following Size Exclusion Chromatography (SEC) (blue). Peak containing CasRx recombinant protein highlighted boxed in gold. [D] A 10% SDS-PAGE gel showing protein species present at different stages of the purification protocol. Lane 1 is SeeBlue Plus2 Pre-stained Protein Standard with the predicted final CasRx protein size marked at right in pink at approximately 112 kDa. Lane 2 is a resuspended cell pellet. Lane 3 is cell lysate supernatant. Lane 4 is cell lysate pellet. Lane 5 is Ni-NTA flow through. Lane 6 is Ni-NTA Elution. Lane 7 is the sample post O/N Dialysis. Lane 8 is the sample after IEC, and Lane 9 is the concentrated Final Sample post-SEC.

Figure S3.3: Molecular scale overview of CasRx detection protocol workflow. Bracketed into the two base component reactions, RT-RPA or CCR. [Panel 1A] Viral ssRNA is extracted and reverse transcribed into cDNA. A fragment of the viral RNA (navy), with a CasRx target site highlighted (mint). Reverse transcriptase (raspberry) actively reverse transcribes the viral template into cDNA. [Panel 2A] In the simultaneous RPA reaction, recombinase (brown) binds to the primer (red and orange) with D-loops formed at binding sites. [2B] Recombinase helps anneal the primer to the target site, and single strand binding protein (SSB, teal) begins annealing to ssDNA to stabilize the strand. [Panel 3A] Strand-displacing DNA polymerase (grey) amplifies the target DNA, with continued binding of SSB to ssDNA for stabilization. [Panel 4A] DNA amplicon extension is completed, [4B] while polymerase simultaneously dislodges SSB, and [4C] the T7 promoter region (orange) is added to the amplicon via primer extension. [Panel 5] The final product of the RT-RPA reaction is a small amplified fragment of target DNA encompassing the CasRx target site, with a T7 promoter added for subsequent IVT. [Panel 6A] T7 RNA polymerase-based *in vitro* transcription (blue) occurs, initiated from the T7 promoter (orange) in order to generate ssRNA (magenta) required as the activation substrate of CasRx. [6B] Elongation of the ssRNA template takes place. [Panel 7] CasRx detection of on-target sequence, activation of collateral cleavage activity, and detection. [7A] The CasRx/gRNA complex (pink and green) recognize, [7B] bind to, and cleave (red scissors) the ssRNA viral target template (magenta). [7C] This action activates the collateral cleavage property of CasRx (gold, black scissors). [7D] Following activation, collateral cleavage of an included small ssRNA probe can be analyzed via [7E] fluorescence or lateral flow assay.


RT-RPA
IVT + Cleavage
1
Reverse Transcription (RT)
viral ssRNA
CasRx target site
[A] Reverse transcription of viral RNA
cDNA
2
[A] Recombinase binding to primer
Forward primer with T7 promoter
Recombinase Polymerase Amplification (RPA)
[A] D-loop formation
Reverse Transcribed dsDNA target
[B] SSB binding
[B] Recombinase annealing primer to target & dissociation
3
[A] Polymerization
Strand-displacing DNA Polymerase
4
[A] Extension
[B] SSB dissociation
[C] Incorporation of T7 promoter
5
Final RPA-amplified gene target (dsDNA)
T7 promoter
RPA amplified dsDNA
6
In vitro Transcription (IVT)
T7 RNA polymerase
[A] in vitro RNA transcription
[B] RNA transcription elongation
ssRNA
7
SENSR
gRNA binding
ssRNA
[A] gRNA/CasRx binding to target
CasRx
gRNA T
[B] Cas13d on-target cleavage
[C] and collateral cleavage activation
[D] Collateral cleavage of probe reporter
[E] Fluorescence
[E] Lateral Flow
(+)
(-)

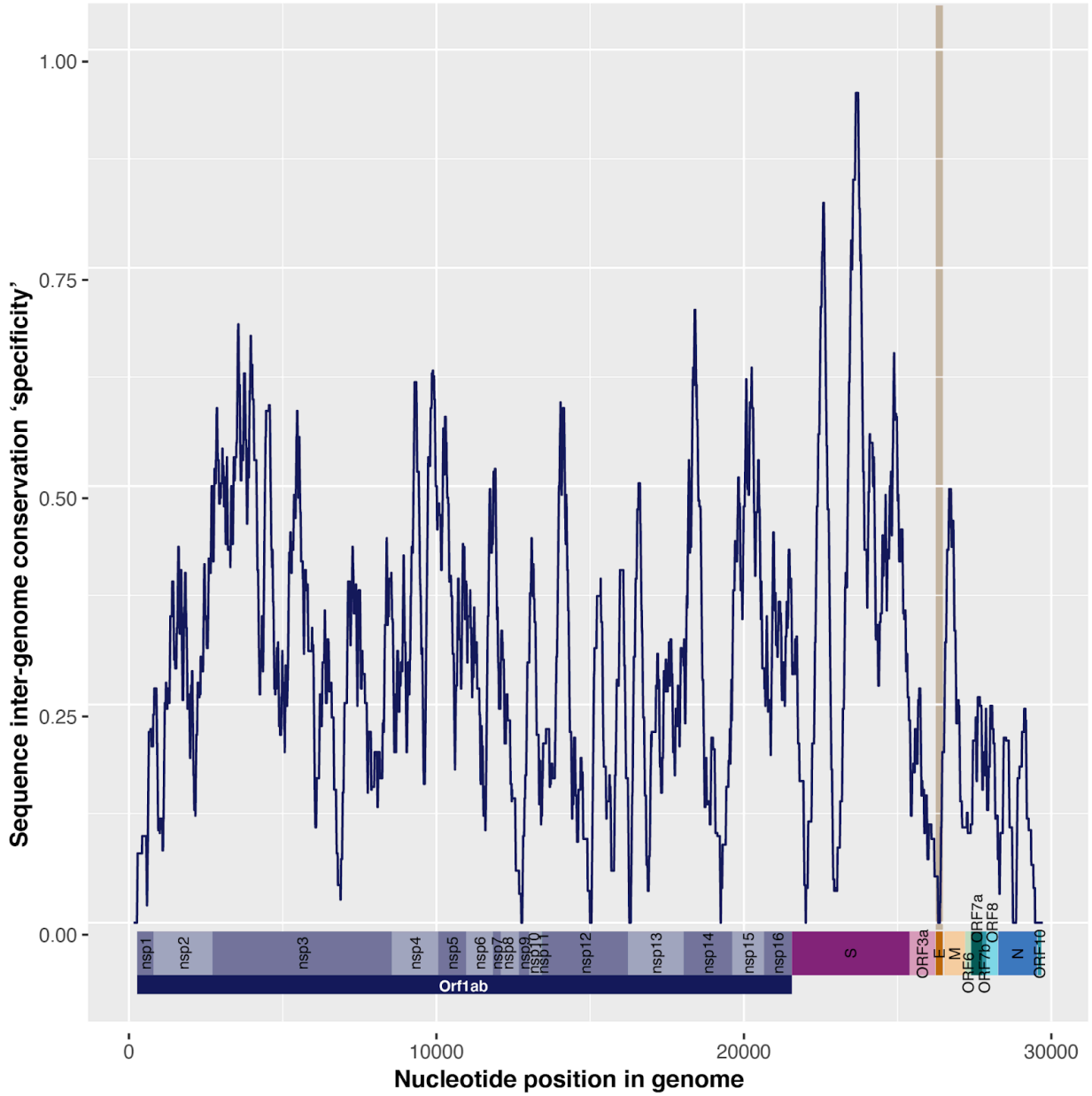


Figure S3.4: Depiction of unique gRNA target sequences across SAR-CoV-2 genome. Spread of unique and specific 30 nt putative gRNA target sequences (Table S3-S5) displayed across the SARS-CoV-2 genome, smoothened over a 301 nt window. Higher specificity score indicates higher density of unique and conserved targets. Representative genome map shown at bottom with ORFs and genes annotated. E-gene ORF highlighted in orange with beige vertical bar highlighting the zero specificity score of the gene.

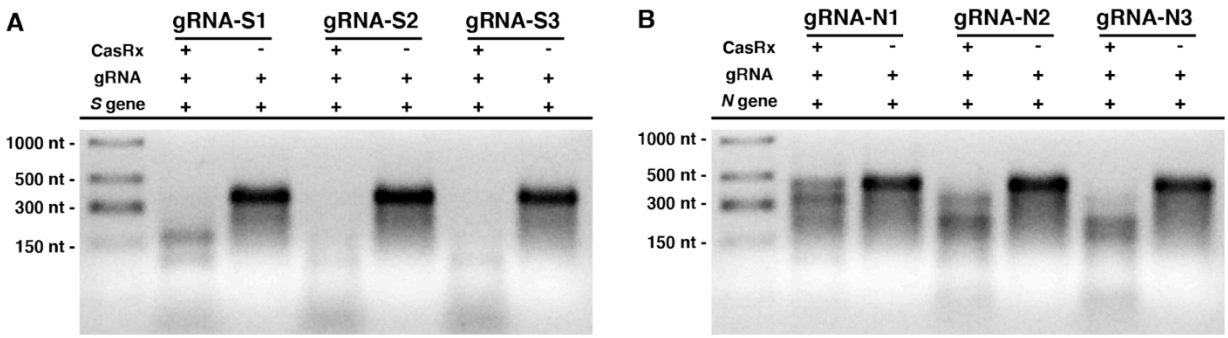


Figure S3.5: Preliminary assessment of candidate guides assayed in this study. [A] Characterization of *Spike* gene gRNA cleavage activity for the three selected candidate guides. Each gRNA was incubated with the S-gene target template (350 nt) in the presence or absence of CasRx for 60 minutes. Cleavage reactions were run on 2% agarose 1xTBE gel for 30 minutes to resolve products. [B] Characterization of *Nucleocapsid* gene gRNA cleavage activity for the three selected candidate guides. Each gRNA was incubated with the N-gene target template (407 nt) in the presence or absence of CasRx for 60 minutes. Cleavage reactions were run on 2% agarose 1xTBE gel for 30 minutes to resolve products.

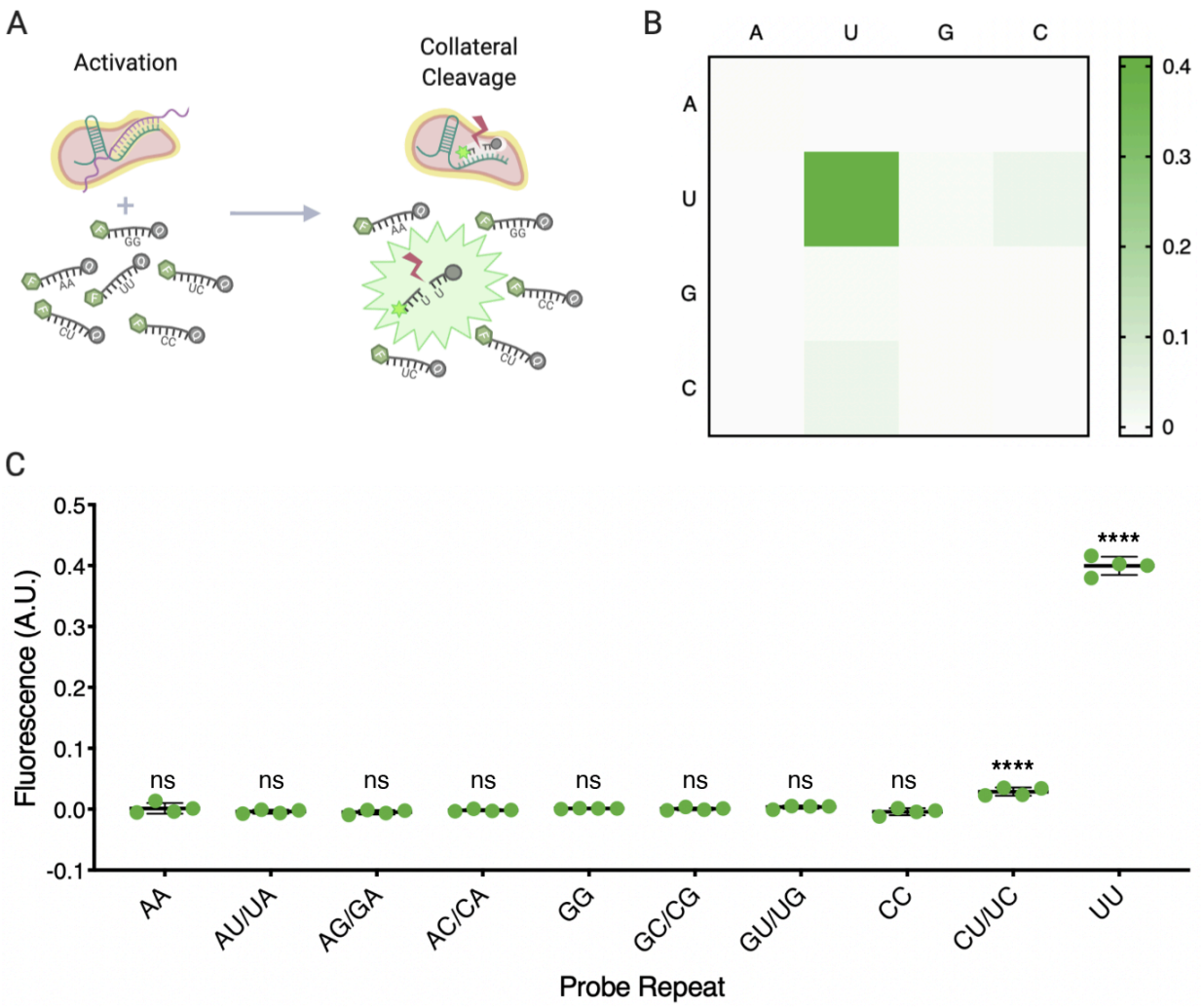


Figure S3.6: Probe di-nucleotide preference selection for SENSR fluorescence detection assay. [A] Schematic representation of experiment determining the collateral cleavage, di-nucleotide preference for CasRx. [B] Heat map layout of background subtracted fluorescence normalized to NTC controls from CasRx collateral cleavage for all di-nucleotide probe combinations. The more saturated the color, the larger the increase in fluorescence. [C] Scatter plot of background subtracted fluorescence of experimental group normalized to NTC controls for all di-nucleotide probes calculated after 60 minutes of acquisition. A.U., Arbitrary Units. Data represents mean ± s.d. from quadruplicate measurements. Significance is representative of Šídák's multiple comparisons test.

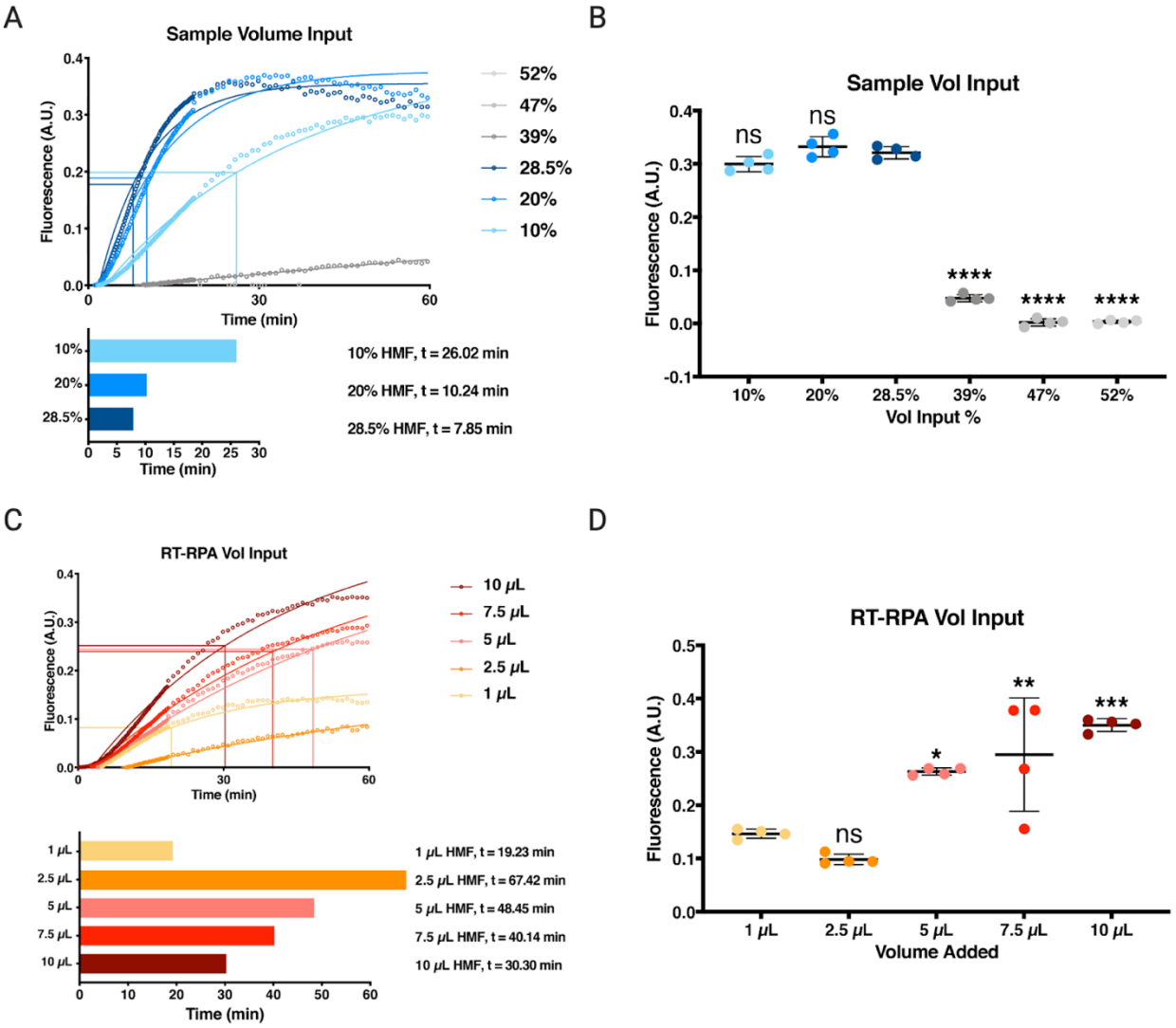


Figure S3.7: Optimization of SENSR volume input parameters. [A] Assessment of sample volume input into RT-RPA reaction by volume percentage. RT-RPA reactions were run for 45 minutes with gRNA-N1 primers targeting 1000 copy/µL of synthetic *Nucleocapsid* template. Data were acquired for 60 minutes and then plotted to perform HMF analysis. 39%, 47%, and 52% v/v sample input HMF not shown in plot: 39% = 238.76 min, 47% = N/A, 52% = N/A. Intersecting lines represent the location of HMF for each group. [B] Scatter plot of background subtracted fluorescence data normalized to NTCs after 60 minutes of acquisition. A.U., Arbitrary Units. Data represents mean ± s.d. from quadruplicate measurements. Significance is representative of Dunnett's multiple comparisons test comparing all groups to 28.5% sample volume input group. [C] Assessment of RT-RPA volume added to CCR. Analysis was performed on 1000 copy/µL of synthetic *Nucleocapsid* template with gRNA-N3, which exhibited the second slowest HMF time in the preliminary gRNA analysis. Data were acquired for 60 minutes and then plotted to perform HMF analysis. Intersecting lines represent the location of HMF for each group. [D] Scatter plot of background subtracted fluorescence data normalized to NTCs after 60 minutes of acquisition. A.U., Arbitrary Units. Data represents mean ± s.d. from quadruplicate measurements. Significance is representative of Dunnett's multiple comparisons test comparing all groups to 1 µL RT-RPA volume input group.

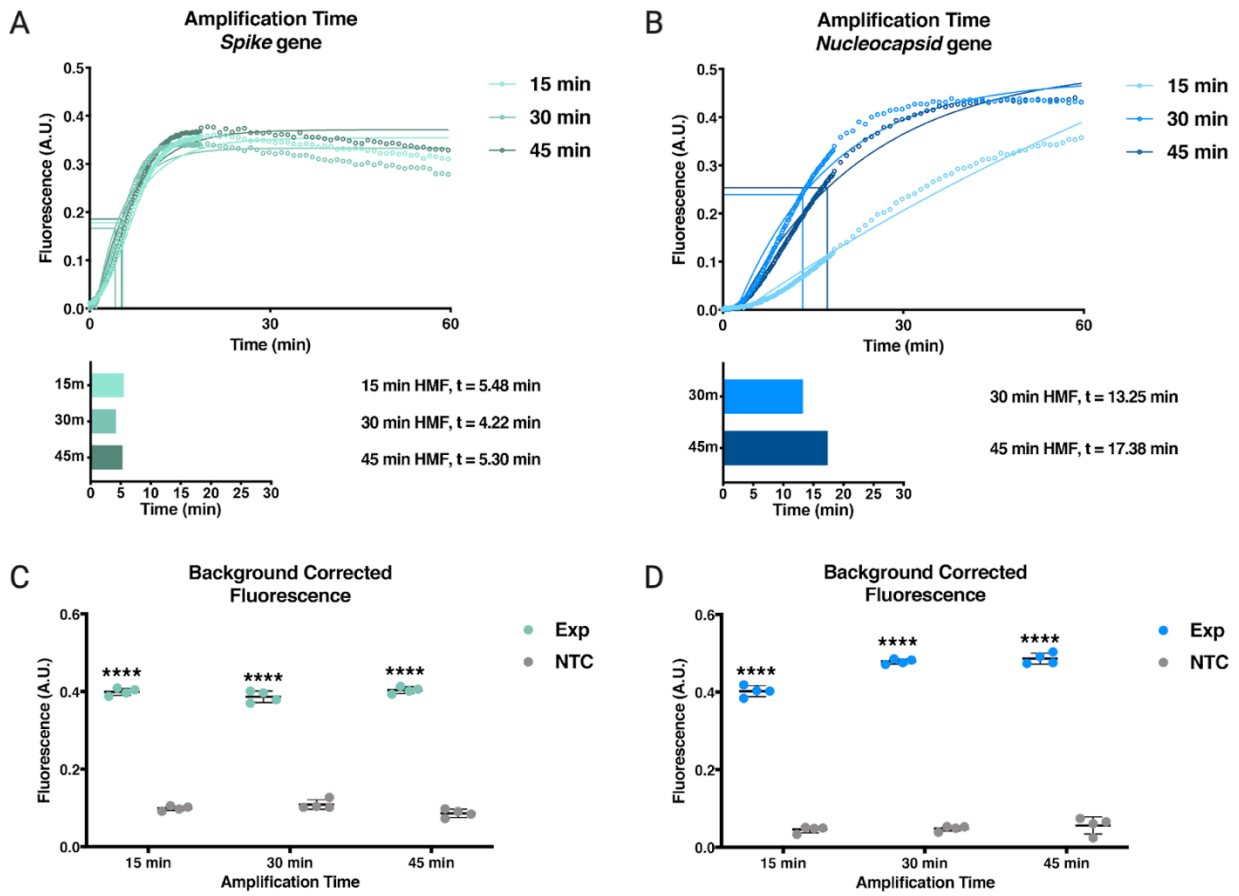


Figure S3.8: Assessment of optimal amplification time for detection of *Spike* and *Nucleocapsid* targets. [A] Fluorescence plot of experimentals targeting 100 copy/µL of synthetic *Spike* template with gRNA-S3 normalized to NTC. Data were acquired for 60 minutes and then plotted to perform HMF analysis. Intersecting lines represent the location of HMF for each group. [B] Fluorescence plot of experimentals targeting 100 copy/µL of synthetic *Spike* template with gRNA-N1 normalized to NTC. Data were acquired for 60 minutes and then plotted to perform HMF analysis. 15 minute amplification HMF not shown in plot: 15m = 89.24 min. Intersecting lines represent the location of HMF for each group. [C] Scatter dot plot of background subtracted fluorescence for *Spike* targeting with gRNA-S3. Fluorescence signal plotted represents the final data acquired at 60 minutes subtracted by the initial fluorescence signal recorded. A.U., Arbitrary Units. Data represents mean ± s.d. from quadruplicate measurements. Significance is representative of Tukey's multiple comparisons test for experimentals compared to respective NTCs. [D] Scatter dot plot of background subtracted fluorescence for *Nucleocapsid* targeting with gRNA-N1. Fluorescence signal plotted represents the final data acquired at 60 minutes subtracted by the initial fluorescence signal recorded. A.U., Arbitrary Units. Data represents mean ± s.d. from quadruplicate measurements. Significance is representative of Tukey's multiple comparisons test for experimentals compared to respective NTCs.

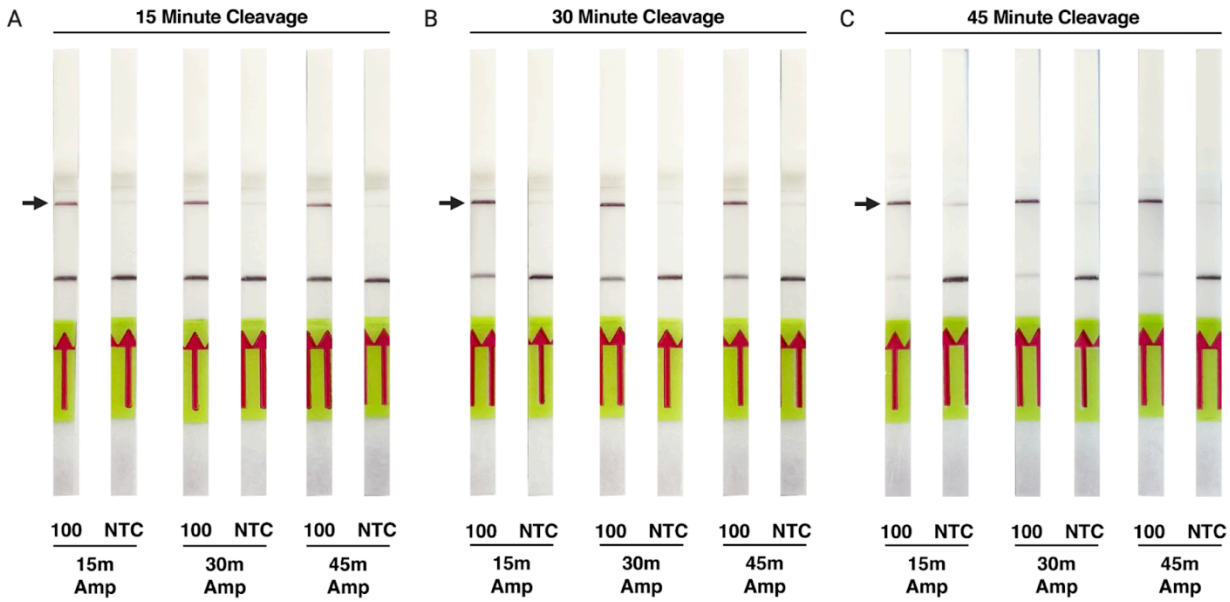
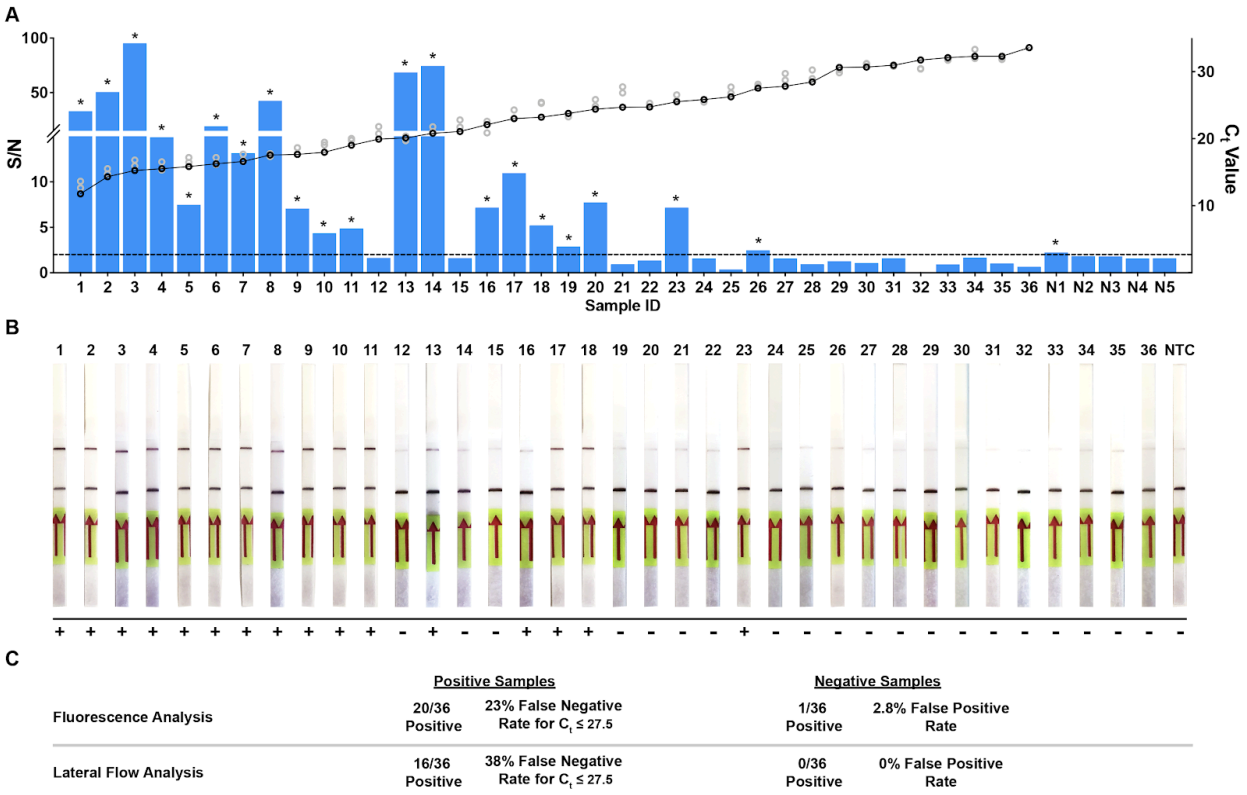


Figure S3.9: Optimization of amplification and cleavage time for lateral flow readout of SENSR. [A] Lateral flow readout detecting SARS-CoV-2 *Spike* gene with gRNA-S3 after 15 minutes cleavage. The synthetic S gene was amplified via RT-RPA for 15, 30, or 45 minutes and then incubated in a CCR for 15 minutes. Lateral flow readout was then performed for 3 minutes. Detection determined increased saturation of test band (arrow). [B] Lateral flow readout after 30 minutes cleavage. [C] Lateral flow readout after 45 minutes cleavage.



C

	Positive Samples		Negative Samples	
Fluorescence Analysis	20/36 Positive	23% False Negative Rate for $C_t \leq 27.5$	1/36 Positive	2.8% False Positive Rate
Lateral Flow Analysis	16/36 Positive	38% False Negative Rate for $C_t \leq 27.5$	0/36 Positive	0% False Positive Rate

Figure S3.10: SENSR detection of positive SARS-CoV-2 validated patient samples with unoptimized protocol. [A] SENSR fluorescence analysis of RT-qPCR validated patient samples using an N-targeting gRNA for detection. Signal-to-Noise (S/N) ratio of 36 positive and 5 negative samples are shown. Positive samples are listed in order from lowest to highest C_t value. Open circles represent C_t value of N- (black), S- (gray), and Orf1ab (gray). The five negative samples are negative samples with the highest S/N ratio. Dashed line represents the S/N threshold = 2. Asterisks represent a S/N > 2. [B] Lateral flow based detection of the 36 positive and NTC for comparison. The top band (arrow) represents the test band and the bottom the control band. An increase in saturation of the top band indicates a positive detection of SARS-CoV-2 in the sample. Result of detection indicated below each test strip. [C] Summary of fluorescence and lateral flow detection results for the RT-qPCR validated positive and negative patient samples.

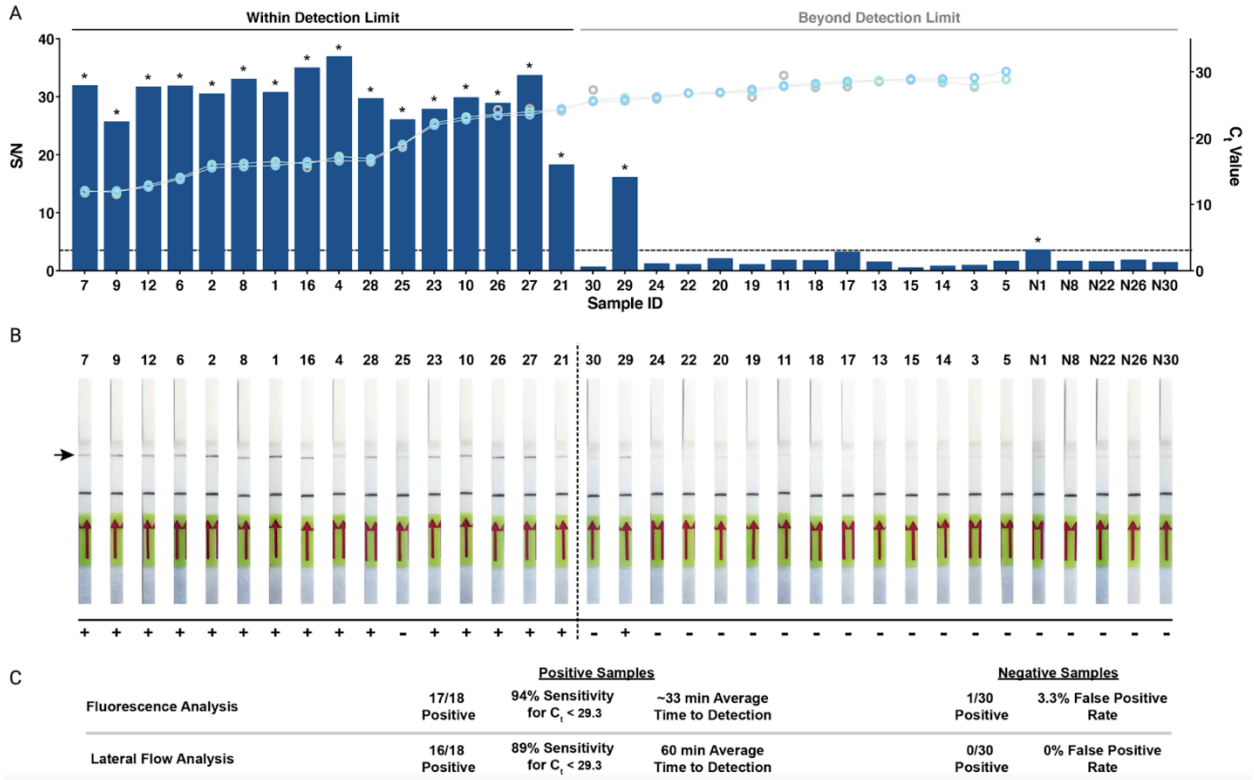


Figure S3.11: SENSR detection of positive SARS-CoV-2 validated patient samples. [A] SENSR fluorescence analysis of RT-qPCR validated patient samples using gRNA-N1 for detection. Signal-to-Noise (S/N) ratio of 30 positive and 5 negative samples are shown. Positive samples are listed in order from lowest to highest C_t value. Open circles represent C_t value of S- (spindrift), N- (ice), and Orf1ab (gray). The five negative samples are negative samples with the highest S/N ratio. Dashed line represents the S/N threshold = 3.506. Asterisks represent a S/N > 3.506. [B] Lateral flow based detection of the 30 positive and 5 negative samples. The top band (arrow) represents the test band and the bottom the control band. An increase in saturation of the top band indicates a positive detection of SARS-CoV-2 in the sample. Result of detection indicated below each test strip. [C] Summary of fluorescence and lateral flow detection results for the RT-qPCR validated positive and negative patient samples.

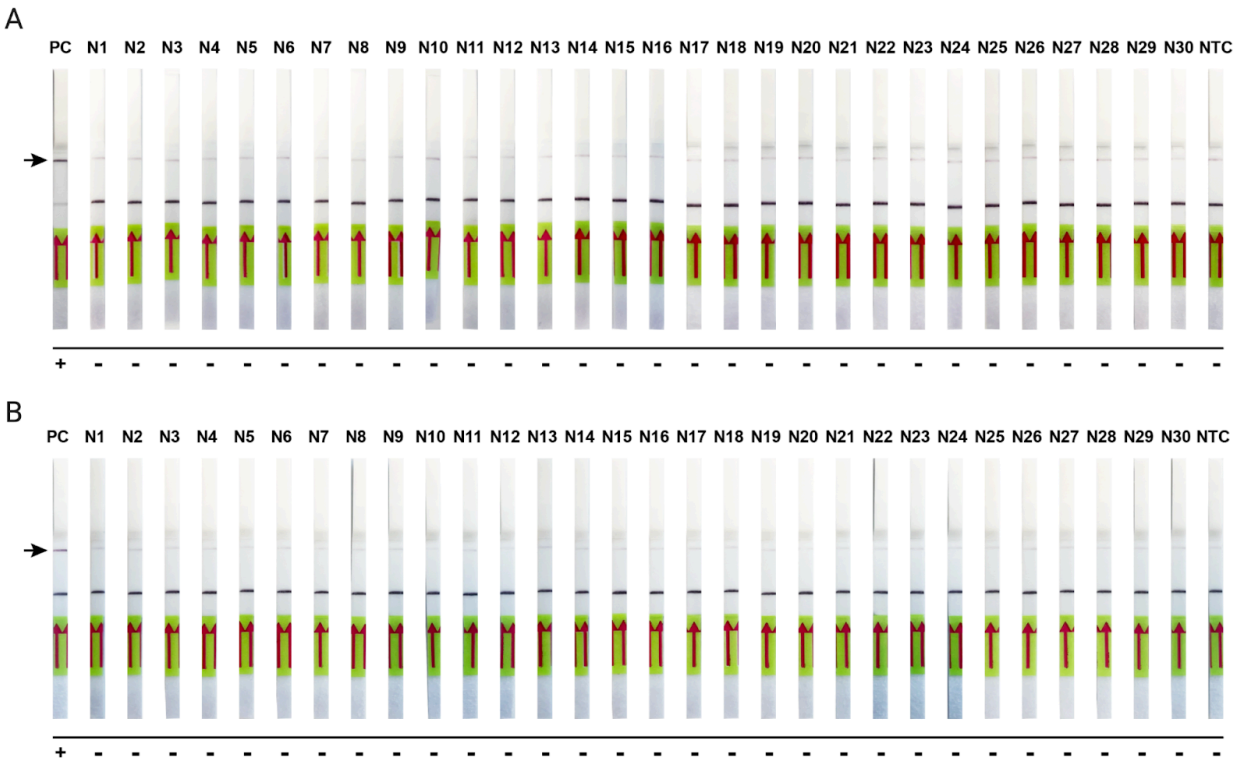


Figure S3.12: SENSR lateral flow analysis of RT-qPCR validated SARS-CoV-2 negative patient samples. [A] SARS-CoV-2 negative patient samples challenged with gRNA-S3 (*Spike*) in LF detection assay (30 minutes RT-RPA/ 30 minutes CCR). Positive Control (PC) and No-Template Controls (NTC) were run in parallel. Positive result determined by increase in saturation in top band (arrow). [B] SARS-CoV-2 negative patient samples challenged with gRNA-N1 (*Nucleocapsid*) in LF detection assay (30 minutes RT-RPA/ 30 minutes CCR). Positive result determined by increase in saturation in top band (arrow).

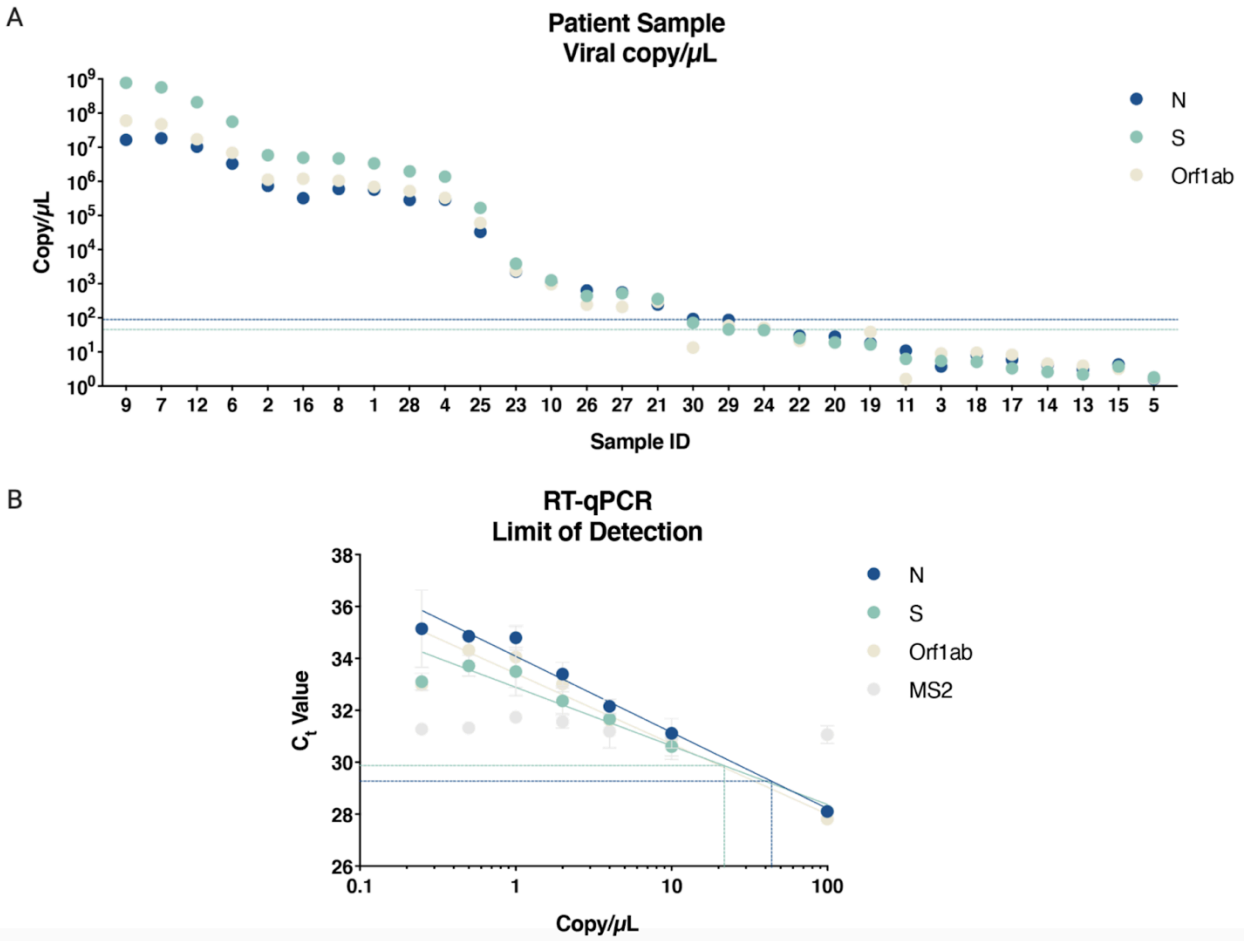


Figure S3.13: Viral copy number per µL for RT-qPCR verified positive samples. [A] Conversion of C_t value to copy/µL using nonlinear regression obtained from [B] for each positive patient sample analyzed. Horizontal dashed lines represent the detection limit for SENSR corrected for 2 µL input: S-gene (green; 45 copy/µL) and N-gene (blue; 89 copy/µL). [B] RT-qPCR LoD experiment for the three targets analyzed via RT-qPCR (N- blue, S-green, and Orf1ab-beige). Semi-log nonlinear regression analysis was performed to generate a line of best fit. RT-qPCR analysis was performed with 2 µL of sample input. Intersecting lines represent the copy/µL and C_t value detection limit for SENSR analysis of S-gene (green; C_t = 29.87, copy/µL = 21.78) and N-gene (blue; C_t = 29.27, copy/µL = 43.76).

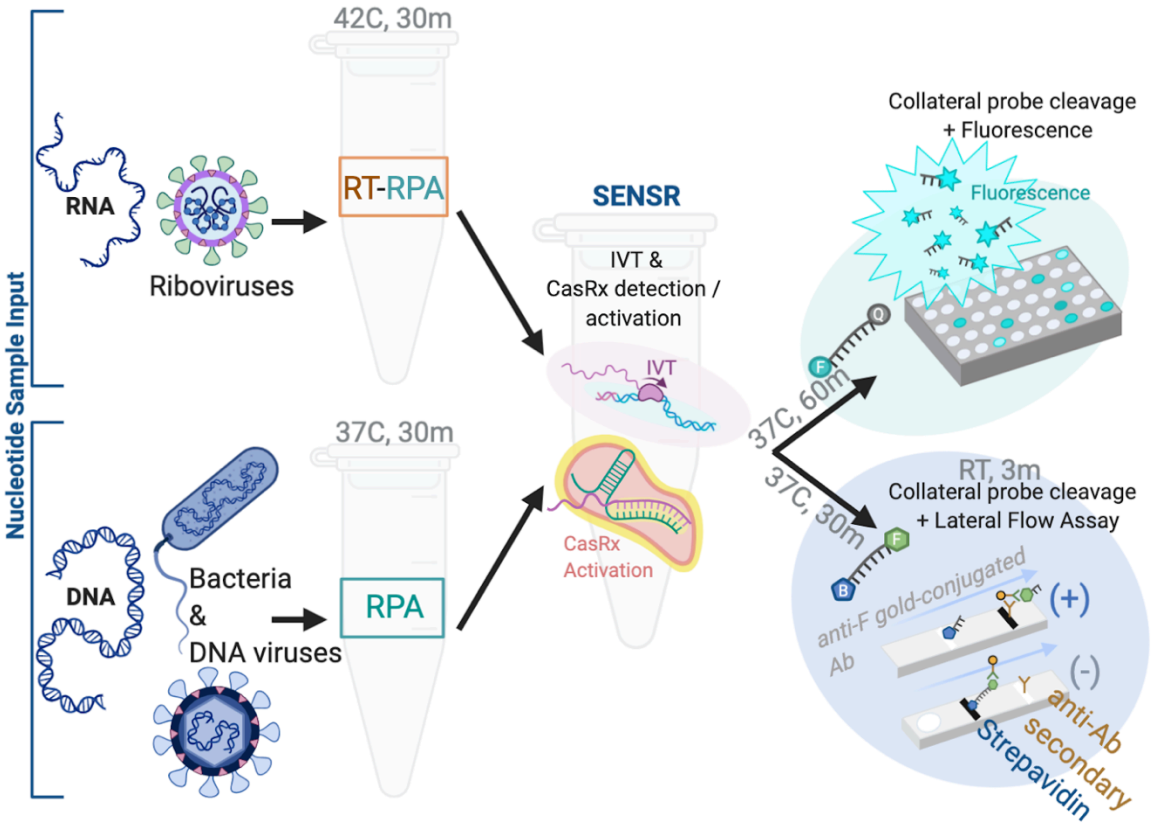


Figure S3.14: Schematic representation of SENSR application for detection of other pathogens. In addition to RNA-based templates (top row), SENSR technology could also be used for DNA-based templates, including, but not limited to, bacterial samples (shown) as well as DNA viruses such as herpes (shown, bottom, navy). DNA-based sample templates can be input directly into the RPA amplification reaction, negating the need for a simultaneous reverse transcription (RT) reaction as is required for RNA-based templates. Like SENSR on RNA-based templates, the reaction occurs in two steps. The first amplification step, RT-RPA or RPA, is followed by the SENSR reaction which is identical regardless of initial input molecular species. This reaction is composed of simultaneous *in vitro* transcription (IVT) as well as CasRx cleavage and activation of collateral cleavage activity. The assay can be read out by fluorescence or lateral flow in the same fashion regardless of DNA or RNA sample input.

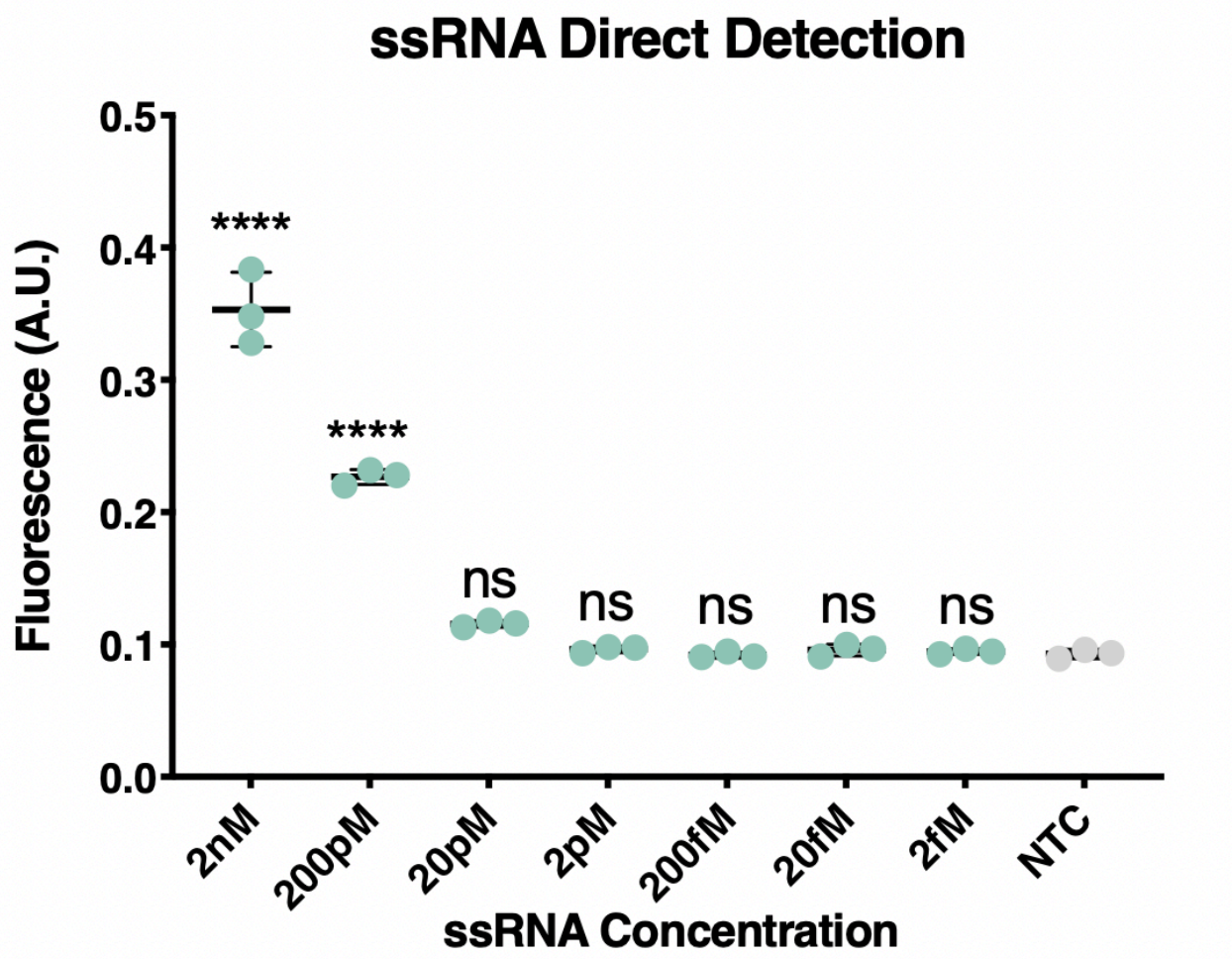


Figure S3.15: CasRx direct detection of synthetic Spike ssRNA. Varied concentrations of synthetic ssRNA Spike template incubated directly with CasRx and gRNA-S3 without preamplification. Fluorescence signal plotted represents the final data acquired at 30min subtracted by the initial fluorescence signal recorded. A.U., Arbitrary Units. Data represents mean ± s.d. from triplicate measurements. Significance is representative of Dunnett's multiple comparisons test for experimentals compared to NTC.

3.9 Supplementary Tables

Table S3.1: Summary Table of current CRISPR-based anti-Covid technologies.

Table S1. Summary Table of current CRISPR-based anti-Covid technologies.

Tool name	What does acronym stand for?	Diagnostic or Treatment?	Time	Enzyme	RNA or DNA Target?	SARS-CoV-2 Target Gene(s)	Mode of Detection	Mode of Amplification	Total Reactions	Readout	Reference: Covid Application	FDA Approved?	Reference : Original Tech
Cas12-based													
DETECTR	DNA Endonuclease-Targeted CRISPR Trans Reporter	Diagnostic	~40 m	Cas12a	DNA	E and N	Collateral cleavage	Yes, RT-RPA		Fluorescence & 2 Lateral flow	https://www.nature.com/articles/s41587-020-0513-4	Yes	https://science.sciencemag.org/content/early/2018/02/14/science.aar6245?srmessenel=true https://www.nature.com/articles/s41421-018-0028-z
AIOD-CRISPR	All-in-One Dual CRISPR-Cas12a	Diagnostic	~90 m	Cas12a	DNA	N	Collateral cleavage	Yes, RPA		1 Fluorescence	https://www.nature.com/articles/s41467-020-18575-6	No	
CASdetec	CRISPR-assisted detection	Diagnostic	~40-60 m	Cas12b	DNA	RdRp		Yes, RT-RAA		1 Fluorescence	https://www.nature.com/articles/s	No	
STOPCovid	SHERLOCK Testing in One Pot	Diagnostic	~15-45 m	Cas12b	DNA	N		Yes, RT-LAMP		1 Lateral Flow	https://www.nejm.org/doi/10.1056	No	
Cas13-based													
SHERLOCK	Sensitive Enzymatic Nucleic acid Sequence Reporter	Diagnostic	~35-70 m	Cas13a, Cas13b	RNA	N, S, and Orf1ab	Collateral Cleavage	Yes, RT-RPA		Fluorescence & 2 Lateral flow	https://www.nature.com/articles/s41551-020-00603-x https://www.fda.gov/media/137745/download https://www.fda.gov/media/137746/download	No	https://science.sciencemag.org/content/356/6336/438 https://science.sciencemag.org/content/360/6387/439
SherlockTM CRISPR SARS-CoV-2 kit	See above	Diagnostic	~1 h	LwaCas13a	RNA	N and Orf1ab	Collateral Cleavage	Yes, RT-LAMP		Fluorescence & 2 Lateral flow	https://www.fda.gov/media/137745/download	Yes	See above
SENSR	Sensitive Enzymatic Nucleic acid Sequence Reporter	Diagnostic	~1h	RfxCas13d	RNA	S and N	Collateral Cleavage	Yes, RT-RPA		Fluorescence & 2 Lateral flow	This paper	No	This paper
CREST	Cas13-based, Rugged, Equitable, Scalable Testing	Diagnostic	~1-2h	Cas13a	RNA	N	Collateral cleavage	Yes, RPA or PCR		Fluorescence & 3+ Lateral flow	https://jcm.asm.org/content/59/4/e02402-20	No	
CARMEN - Cas13	Combinatorial Arrayed Reactions for Multiplexed Evaluation of Nucleic acids	Diagnostic	N/A	Cas13a	RNA	N/A	Collateral cleavage	Yes, PCR or RPA.		2 Fluorescence	https://www.nature.com/articles/s	No	
PAC-MAN	Prophylactic Antiviral CRISPR in huMAN cells	Prophylactic	N/A	Cas13d	RNA	RdRp and N	N/A	No	N/A	N/A	https://www.cell.com/cell/pdf/S00	No	
[no name] Georgia Institute of Tech, Blanchard et al	N/A	Prophylactic	N/A	Cas13a	RNA	N and Orf1ab	N/A		N/A	N/A	https://www.biorxiv.org/content/1	No	
Other Cas's													
FELUDA	FnCas9 Editor Linked Uniform Detection Assay	Diagnostic	~1-2h	FnCas9 (dCas9)	DNA	N	Binding or Cleavage	Yes, PCR or RPA		Agarose Capillary electrophoresis (cleavage) Lateral flow Fluorescence 2 readout	https://www.sciencedirect.com/science/article/pii/S0956566321000244X	No	

Table S3.2: Analysis identifying putative 30 nt CasRx gRNA target sites conserved across, and specific to, the SARS-CoV-2 genome.

ORF id	total_targets (30nt)	specific&conserved_targets	percent_specific&conserved
	Total number putative target segments (30nt) identified per gene.	Total number of putative target sequences conserved within 433 available SARS-CoV-2 genomes, and uniquely specific to SARS-CoV-2 when compared to the 3164 publicly available Coronavirus genomes.	= [specific&conserved_targets] / [total_targets] (%)
endoRNAse	1009	483	47.869
S	3793	1568	41.339
nsp7	220	90	40.909
nsp4	1471	574	39.021
3C-like proteinase	889	334	37.57
nsp3	5806	2117	36.462
M	640	231	36.094
nsp6	841	261	31.034
2'-O-ribose methyltransferase	865	268	30.983
nsp10	388	112	28.866
3'-to-5' exonuclease	1552	432	27.835
nsp2	1885	495	26.26
RNA-dependent RNA polymerase	2766	672	24.295
helicase	1774	428	24.126
ORF7a	337	79	23.442
ORF8	337	75	22.255
nsp8	565	105	18.584
ORF3a	799	127	15.895
leader protein	511	64	12.524
N	1231	150	12.185
no_gene	1335	146	10.936
nsp9	310	30	9.677
ORF7b	103	5	4.854
E	199	0	0
ORF10	88	0	0
ORF6	157	0	0
Totals	29714	8846	NA

Table S3.5: RT-RPA protocols for sample volume input variation experiments.

RT-RPA (10μL)	Old Protocol	10% Sample	20% Sample	28.5% Sample	39% Sample	47% Sample	52% Sample
Component	**Volume (μL)**	**Volume (μL)**	**Volume (μL)**	**Volume (μL)**	**Volume (μL)**	**Volume (μL)**	**Volume (μL)**
Rehydration Buffer	5.9	5.9	5.9	5.9	5.9	5.9	5.9
Primer Mix (10uM ea)	0.9	0.35	0.35	0.35	0.42	0.49	0.56
RT (200U/μL)	0.2	0.4	0.4	0.4	0.4	0.4	0.4
Nuclease-Free Water	0	1.85	0.85	0	0	0	0
Sample	2	1	2	2.85	4.68	6.51	8.34
MgOAc (280mM)	(140mM) 1	0.5	0.5	0.5	0.6	0.7	0.8
Total	**10**	**10**	**10**	**10**	**12**	**14**	**16**
Paper Figure	Fig. S10	Fig. S7	Fig. 1, Fig. S7	Fig. 2, Fig. 3, Fig. S6, Fig. S7	Fig. S7	Fig. S7	Fig. S7

Table S3.6: CCR protocols for RT-RPA volume variation experiments.

CCR (20μL)	Old Protocol	1μL RT-RPA	2.5μL RT-RPA	5μL RT-RPA	7.5μL RT-RPA	10μL RT-RPA
Component	**Volume (uL)**	**Volume (μL)**	**Volume (μL)**	**Volume (μL)**	**Volume (μL)**	**Volume (μL)**
Nuclease-Free Water	5.32	11.82	10.32	7.82	5.32	2.82
HEPES (1M, pH 7.2)	0.4	0.4	0.4	0.4	0.4	0.4
MgCl2 (1M)	0.18	0.18	0.18	0.18	0.18	0.18
rNTP Mix (25mM ea)	3.2	1	1	1	1	1
CasRx (55.4ng/μL)	2	2	2	2	2	2
RNase Inhibitor (40U/μL)	1	1	1	1	1	1
T7 Polymerase (50U/μL)	0.6	0.6	0.6	0.6	0.6	0.6
gRNA (10ng/μL)	1	1	1	1	1	1
FRU (2uM)	1	1	1	1	1	1
RT-RPA	5	1	2.5	5	7.5	10
Total	**20**	**20**	**20**	**20**	**20**	**20**
Paper Figure	Fig. S10	Fig. S7	Fig. S7	Fig. 1, Fig. S7	Fig. S7	Fig. 2, Fig. 3, Fig S6, Fig. S7, Fig S8

Table S3.8: Time-to-detection from fluorescence analysis of all positive patient samples.

	Time-to-Detection (min)	
	gRNA-S3	**gRNA-N1**
Positive Sample 1	30.97	30.97
Positive Sample 2	31.15	31.35
Positive Sample 3	N/A	N/A
Positive Sample 4	35.4	30.77
Positive Sample 5	N/A	N/A
Positive Sample 6	31.55	30.97
Positive Sample 7	31.55	30.97
Positive Sample 8	30.97	31.15
Positive Sample 9	31.73	30.77
Positive Sample 10	31.55	32.12
Positive Sample 11	N/A	N/A
Positive Sample 12	31.35	30.77
Positive Sample 13	N/A	N/A
Positive Sample 14	39.05	N/A
Positive Sample 15	N/A	N/A
Positive Sample 16	31.55	30.97
Positive Sample 17	N/A	N/A
Positive Sample 18	40.98	N/A
Positive Sample 19	N/A	N/A
Positive Sample 20	38.08	N/A
Positive Sample 21	33.08	38.88
Positive Sample 22	N/A	N/A
Positive Sample 23	31.73	35.4
Positive Sample 24	N/A	N/A
Positive Sample 25	30.58	35.78
Positive Sample 26	34.23	35.4
Positive Sample 27	31.35	32.12
Positive Sample 28	31.15	30.97
Positive Sample 29	32.5	39.27
Positive Sample 30	33.47	N/A

3.10 References

Abbott, Timothy R., Girija Dhamdhere, Yanxia Liu, Xueqiu Lin, Laine Goudy, Leiping Zeng, Augustine Chemparathy, et al. 2020. “Development of CRISPR as an Antiviral Strategy to Combat SARS-CoV-2 and Influenza.” *Cell* 181 (4): 865–76.e12. https://doi.org/10.1016/j.cell.2020.04.020.

Abudayyeh, Omar O., Jonathan S. Gootenberg, Silvana Konermann, Julia Joung, Ian M. Slaymaker, David B. T. Cox, Sergey Shmakov, et al. 2016. “C2c2 Is a Single-Component Programmable RNA-Guided RNA-Targeting CRISPR Effector.” *Science* 353 (6299): aaf5573. https://doi.org/10.1126/science.aaf5573.

Ackerman, Cheri M., Cameron Myhrvold, Sri Gowtham Thakku, Catherine A. Freije, Hayden C. Metsky, David K. Yang, Simon H. Ye, et al. 2020. “Massively Multiplexed Nucleic Acid Detection Using Cas13.” *Nature*, April. https://doi.org/10.1038/s41586-020-2279-8.

Andersen, Kristian G., Andrew Rambaut, W. Ian Lipkin, Edward C. Holmes, and Robert F. Garry. 2020. “The Proximal Origin of SARS-CoV-2.” *Nature Medicine* 26 (4): 450–52. https://doi.org/10.1038/s41591-020-0820-9.

Azhar, Mohd, Rhythm Phutela, Manoj Kumar, Asgar Hussain Ansari, Riya Rauthan, Sneha Gulati, Namrata Sharma, et al. 2021. “Rapid and Accurate Nucleobase Detection Using FnCas9 and Its Application in COVID-19 Diagnosis.” *Biosensors & Bioelectronics* 183 (July): 113207. https://doi.org/10.1016/j.bios.2021.113207.

Brieba, Luis G., Robert Padilla, and Rui Sousa. 2002. “Role of T7 RNA Polymerase His784 in Start Site Selection and Initial Transcription.” *Biochemistry* 41 (16): 5144–49. https://doi.org/10.1021/bi016057v.

Broughton, James P., Xianding Deng, Guixia Yu, Clare L. Fasching, Venice Servellita, Jasmeet Singh, Xin Miao, et al. 2020. “CRISPR-Cas12-Based Detection of SARS-CoV-2.” *Nature Biotechnology*, April. https://doi.org/10.1038/s41587-020-0513-4.

Broughton, James P., Xianding Deng, Guixia Yu, Clare L. Fasching, Jasmeet Singh, Jessica Streithorst, Andrea Granados, et al. 2020. “Rapid Detection of 2019 Novel Coronavirus SARS-CoV-2 Using a CRISPR-Based DETECTR Lateral Flow Assay.” *Infectious Diseases (except HIV/AIDS)*. medRxiv. https://doi.org/10.1101/2020.03.06.20032334.

Buchman, Anna B., Dan J. Brogan, Ruichen Sun, Ting Yang, Patrick D. Hsu, and Omar S. Akbari. 2020. “Programmable RNA Targeting Using CasRx in Flies.” *The CRISPR Journal* 3 (3): 164–76. https://doi.org/10.1089/crispr.2020.0018.

Chen, Janice S., Enbo Ma, Lucas B. Harrington, Maria Da Costa, Xinran Tian, Joel M. Palefsky, and Jennifer A. Doudna. 2018. “CRISPR-Cas12a Target Binding Unleashes Indiscriminate Single-Stranded DNase Activity.” *Science* 360 (6387): 436–39. https://doi.org/10.1126/science.aar6245.

Corman, Victor M., Olfert Landt, Marco Kaiser, Richard Molenkamp, Adam Meijer, Daniel K. W. Chu, Tobias Bleicker, et al. 2020. "Detection of 2019 Novel Coronavirus (2019-nCoV) by Real-Time RT-PCR." *Euro Surveillance: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin* 25 (3). https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045.

Ding, Xiong, Kun Yin, Ziyue Li, Rajesh V. Lalla, Enrique Ballesteros, Maroun M. Sfeir, and Changchun Liu. 2020. "Ultrasensitive and Visual Detection of SARS-CoV-2 Using All-in-One Dual CRISPR-Cas12a Assay." *Nature Communications* 11 (1): 4711. https://doi.org/10.1038/s41467-020-18575-6.

Fozouni, Parinaz, Sungmin Son, María Díaz de León Derby, Gavin J. Knott, Carley N. Gray, Michael V. D'Ambrosio, Chunyu Zhao, et al. 2021. "Amplification-Free Detection of SARS-CoV-2 with CRISPR-Cas13a and Mobile Phone Microscopy." *Cell* 184 (2): 323–33.e9. https://doi.org/10.1016/j.cell.2020.12.001.

Freije, Catherine A., Cameron Myhrvold, Chloe K. Boehm, Aaron E. Lin, Nicole L. Welch, Amber Carter, Hayden C. Metsky, et al. 2019. "Programmable Inhibition and Detection of RNA Viruses Using Cas13." *Molecular Cell* 76 (5): 826–37.e11. https://doi.org/10.1016/j.molcel.2019.09.013.

Gootenberg, Jonathan S., Omar O. Abudayyeh, Max J. Kellner, Julia Joung, James J. Collins, and Feng Zhang. 2018. "Multiplexed and Portable Nucleic Acid Detection Platform with Cas13, Cas12a, and Csm6." *Science* 360 (6387): 439–44. https://doi.org/10.1126/science.aaq0179.

Gootenberg, Jonathan S., Omar O. Abudayyeh, Jeong Wook Lee, Patrick Essletzbichler, Aaron J. Dy, Julia Joung, Vanessa Verdine, et al. 2017. "Nucleic Acid Detection with CRISPR-Cas13a/C2c2." *Science*. https://doi.org/10.1126/science.aam9321.

Hajian, Reza, Sarah Balderston, Thanhtra Tran, Tara deBoer, Jessy Etienne, Mandeep Sandhu, Noreen A. Wauford, et al. 2019. "Detection of Unamplified Target Genes via CRISPR-Cas9 Immobilized on a Graphene Field-Effect Transistor." *Nature Biomedical Engineering* 3 (6): 427–37. https://doi.org/10.1038/s41551-019-0371-x.

Harrington, Lucas B., David Burstein, Janice S. Chen, David Paez-Espino, Enbo Ma, Isaac P. Witte, Joshua C. Cofsky, Nikos C. Kyrpides, Jillian F. Banfield, and Jennifer A. Doudna. 2018. "Programmed DNA Destruction by Miniature CRISPR-Cas14 Enzymes." *Science* 362 (6416): 839–42. https://doi.org/10.1126/science.aav4294.

Howson, Emma L. A., Yohei Kurosaki, Jiro Yasuda, Masayoshi Takahashi, Hiroaki Goto, Ashley R. Gray, Valerie Mioulet, Donald P. King, and Veronica L. Fowler. 2017. "Defining the Relative Performance of Isothermal Assays That Can Be Used for Rapid and Sensitive Detection of Foot-and-Mouth Disease Virus." *Journal of Virological Methods* 249 (November): 102–10. https://doi.org/10.1016/j.jviromet.2017.08.013.

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012. “A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity.” *Science* 337 (6096): 816–21. https://doi.org/10.1126/science.1225829.

Joung, Julia, Alim Ladha, Makoto Saito, Nam-Gyun Kim, Ann E. Woolley, Michael Segel, Robert P. J. Barretto, et al. 2020. “Detection of SARS-CoV-2 with SHERLOCK One-Pot Testing.” *The New England Journal of Medicine*, September. https://doi.org/10.1056/NEJMc2026172.

Kellner, Max J., Jeremy G. Koob, Jonathan S. Gootenberg, Omar O. Abudayyeh, and Feng Zhang. 2019. “SHERLOCK: Nucleic Acid Detection with CRISPR Nucleases.” *Nature Protocols* 14 (10): 2986–3012. https://doi.org/10.1038/s41596-019-0210-2.

Konermann, Silvana, Peter Lotfy, Nicholas J. Brideau, Jennifer Oki, Maxim N. Shokhirev, and Patrick D. Hsu. 2018. “Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors.” *Cell* 173 (3): 665–76.e14. https://doi.org/10.1016/j.cell.2018.02.033.

Kushawah, Gopal, Luis Hernandez-Huertas, Joaquin Abugattas-Nuñez Del Prado, Juan R. Martinez-Morales, Michelle L. DeVore, Huzaifa Hassan, Ismael Moreno-Sanchez, et al. 2020. “CRISPR-Cas13d Induces Efficient mRNA Knockdown in Animal Embryos.” *Developmental Cell* 54 (6): 805–17.e7. https://doi.org/10.1016/j.devcel.2020.07.013.

Li, Shi-Yuan, Qiu-Xiang Cheng, Jing-Man Wang, Xiao-Yan Li, Zi-Long Zhang, Song Gao, Rui-Bing Cao, Guo-Ping Zhao, and Jin Wang. 2018. “CRISPR-Cas12a-Assisted Nucleic Acid Detection.” *Cell Discovery* 4 (April): 20. https://doi.org/10.1038/s41421-018-0028-z.

Lucia, Curti, Pereyra-Bonnet Federico, and Gimenez Carla Alejandra. 2020. “An Ultrasensitive, Rapid, and Portable Coronavirus SARS-CoV-2 Sequence Detection Method Based on CRISPR-Cas12.” *bioRxiv*, March. https://doi.org/10.1101/2020.02.29.971127.

Lu, Xiaoyan, Lijuan Wang, Senthilkumar K. Sakthivel, Brett Whitaker, Janna Murray, Shifaq Kamili, Brian Lynch, et al. 2020. “US CDC Real-Time Reverse Transcription PCR Panel for Detection of Severe Acute Respiratory Syndrome Coronavirus 2.” *Emerging Infectious Diseases* 26 (8). https://doi.org/10.3201/eid2608.201246.

Mallapaty, Smriti. 2020. “Animal Source of the Coronavirus Continues to Elude Scientists.” *Nature*, May. https://doi.org/10.1038/d41586-020-01449-8.

Metsky, Hayden C., Catherine A. Freije, Tinna-Solveig F. Kosoko-Thoroddsen, Pardis C. Sabeti, and Cameron Myhrvold. 2020. “CRISPR-Based Surveillance for COVID-19 Using Genomically-Comprehensive Machine Learning Design.” *bioRxiv*. https://doi.org/10.1101/2020.02.26.967026.

Mukama, Omar, Jinghua Wu, Zhiyuan Li, Qiongxin Liang, Zhijian Yi, Xuewen Lu, Yujie Liu, et al. 2020. “An Ultrasensitive and Specific Point-of-Care CRISPR/Cas12 Based Lateral Flow

Biosensor for the Rapid Detection of Nucleic Acids." *Biosensors and Bioelectronics*. https://doi.org/10.1016/j.bios.2020.112143.

Patchsung, Maturada, Krittapas Jantarug, Archiraya Pattama, Kanokpol Aphicho, Surased Suraritdechachai, Piyachat Meesawat, Khomkrit Sappakhaw, et al. 2020. "Clinical Validation of a Cas13-Based Assay for the Detection of SARS-CoV-2 RNA." *Nature Biomedical Engineering*, August. https://doi.org/10.1038/s41551-020-00603-x.

Piepenburg, Olaf, Colin H. Williams, Derek L. Stemple, and Niall A. Armes. 2006. "DNA Detection Using Recombination Proteins." *PLoS Biology* 4 (7): e204. https://doi.org/10.1371/journal.pbio.0040204.

Rauch, Jennifer N., Eric Valois, Sabrina C. Solley, Friederike Braig, Ryan S. Lach, Morgane Audouard, Jose Carlos Ponce-Rojas, et al. 2021. "A Scalable, Easy-to-Deploy Protocol for Cas13-Based Detection of SARS-CoV-2 Genetic Material." *Journal of Clinical Microbiology* 59 (4). https://doi.org/10.1128/JCM.02402-20.

Rezaei, Meysam, Sajad Razavi Bazaz, Sareh Zhand, Nima Sayyadi, Dayong Jin, Martin P. Stewart, and Majid Ebrahimi Warkiani. 2020. "Point of Care Diagnostics in the Age of COVID-19." *Diagnostics (Basel, Switzerland)* 11 (1). https://doi.org/10.3390/diagnostics11010009.

Sefers, Susan, and Jonathan E. Schmitz. 2018. "Molecular Contamination and Amplification Product Inactivation." In *Advanced Techniques in Diagnostic Microbiology: Volume 1: Techniques*, edited by Yi-Wei Tang and Charles W. Stratton, 505–26. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-33900-9_24.

Sun, Ruichen, Daniel Brogan, Anna Buchman, Ting Yang, and Omar S. Akbari. 2021. "Ubiquitous and Tissue-Specific RNA Targeting in Drosophila Melanogaster Using CRISPR/CasRx." *Journal of Visualized Experiments: JoVE*, no. 168 (February). https://doi.org/10.3791/62154.

Vogels, Chantal B. F., Anderson F. Brito, Anne L. Wyllie, Joseph R. Fauver, Isabel M. Ott, Chaney C. Kalinich, Mary E. Petrone, et al. 2020. "Analytical Sensitivity and Efficiency Comparisons of SARS-CoV-2 RT-qPCR Primer-Probe Sets." *Nature Microbiology* 5 (10): 1299–1305. https://doi.org/10.1038/s41564-020-0761-6.

Wang, Chen, Peter W. Horby, Frederick G. Hayden, and George F. Gao. 2020. "A Novel Coronavirus Outbreak of Global Health Concern." *The Lancet*. https://doi.org/10.1016/S0140-6736(20)30185-9.

WHO. 2020. "Coronavirus Disease 2019 (COVID-19) Situation Report - 93." https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200422-sitrep-93-covid-19.pdf?sfvrsn=35cf80d7_4.

Yan, Winston X., Shaorong Chong, Huaibin Zhang, Kira S. Makarova, Eugene V. Koonin, David

R. Cheng, and David A. Scott. 2018. “Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein.” *Molecular Cell* 70 (2): 327–39.e5. https://doi.org/10.1016/j.molcel.2018.02.028.

Zetsche, Bernd, Jonathan S. Gootenberg, Omar O. Abudayyeh, Ian M. Slaymaker, Kira S. Makarova, Patrick Essletzbichler, Sara E. Volz, et al. 2015. “Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System.” *Cell* 163 (3): 759–71. https://doi.org/10.1016/j.cell.2015.09.038.

Zhang, Feng, Omar O. Abudayyeh, and S. Gootenberg Jonathan. 2020. “A Protocol for Detection of COVID-19 Using CRISPR Diagnostics.” https://www.broadinstitute.org/files/publications/special/COVID-19%20detection%20(updated).pdf.

Zhu, Na, Dingyu Zhang, Wenling Wang, Xingwang Li, Bo Yang, Jingdong Song, Xiang Zhao, et al. 2020. “A Novel Coronavirus from Patients with Pneumonia in China, 2019.” *The New England Journal of Medicine* 382 (8): 727–33. https://doi.org/10.1056/NEJMoa2001017.

Chapter 4: Advancement of CRISPR diagnostics toward the point of care

4.1 Abstract

CRISPR diagnostics have recently emerged as powerful diagnostic tools for the rapid detection of infections. The ultimate goal is to develop these diagnostics for the point of care, where patients quickly receive and easily interpret results. Although they are in their infancy, the COVID-19 pandemic has accelerated innovation of CRISPR diagnostics and led to an explosion of improvements to these systems. Challenges that have impeded the implementation at the point of care have been addressed, and CRISPR diagnostics have been dramatically simplified. Here we outline recent developments and advancements in CRISPR diagnostics that have pushed these technologies to the point of care.

4.2 Introduction

Rapid diagnostics are crucial in the battle against highly transmissible diseases such as COVID-19. Point of care (POC) diagnostics ideally require intuitive laboratory equipment and rapidly produce easily interpretable results(Valera et al. 2021). For the detection of infectious diseases, POC diagnostics have primarily relied on the detection of antigens to identify a pathogen in an infected individual(S. Pavia and M. Plummer 2021). Antigen tests produce rapid results, however accuracy can vary and successful detection is dependent on virus maturation and high viral titres(Pilarowski et al. 2021), whereas nucleic acid tests can detect an infection prior to a patient becoming transmissive(Pan et al. 2020; Byrne et al. 2020). Recently, CRISPR-based diagnostics (CRISPR-Dx) have emerged as revolutionary nucleic acid tests with the potential to operate in POC settings.

CRISPR-Dx generally combine amplification technologies with CRISPR-Cas (Cas) enzymes and exploit their sequence-specific recognition properties to detect pathogens. CRISPR-Dx have primarily utilized two main CRISPR effector types to accomplish nucleic acid detection: Type-V (Cas12) DNA-targeting systems(Zetsche et al. 2015) and Type-VI (Cas13) RNA-targeting systems(Abudayyeh et al. 2016; East-Seletsky et al. 2016) (Table 1). Both Cas12 and Cas13 effectors exhibit collateral cleavage activity, or subsequent cleavage of bystander nucleic acid sequences, where Cas12 cleaves secondary ssDNA molecules(Chen et al. 2018; Li et al. 2018) and Cas13 cleaves secondary ssRNA molecules(Abudayyeh et al. 2016; East-Seletsky et al. 2016; Smargon et al. 2017; Konermann et al. 2018). For CRISPR-Dx, the collateral cleavage activity has been harnessed as the primary mode of nucleic acid detection(Gootenberg et al. 2017; Chen et al. 2018; Gootenberg et al. 2018; Ackerman et al. 2020; Broughton et al. 2020; Brogan et al. 2021) (Figure 1). This is accomplished by introducing a modified secondary probe that, upon cleavage, provides readouts by fluorescence or visually by lateral flow. Further investigation into this phenomenon identified distinct nucleic acid sequence preferences of collateral probes for various Cas13 orthologs and subtypes, ultimately demonstrating multiplexing capabilities of CRISPR-Dx(Gootenberg et al. 2018; Brogan et al. 2021). Cas effectors have improved clinical diagnostic systems by providing highly specific sequence recognition, however the sensitivity of CRISPR-Dx is primarily dependent on the amplification method applied.

The concept of harnessing collateral cleavage for nucleic acid detection was first established in 2016(East-Seletsky et al. 2016). Subsequently, the addition of an amplification method in the SHERLOCK system dramatically improved the sensitivity of CRISPR-Dx(Gootenberg et al. 2017). To simplify these diagnostic tools, the first iterations of

CRISPR-Dx utilized an isothermal amplification method for targeted sequence amplification termed Recombinase Polymerase Amplification (RPA)(Piepenburg et al. 2006). Successive versions of these diagnostics swapped RPA for an alternative isothermal amplification method with greater specificity and sensitivity termed Loop-mediated isothermal AMPlification (LAMP)(Notomi et al. 2000; Howson et al. 2017; Broughton et al. 2020; Agrawal et al. 2020). The use of isothermal amplification methods simplifies the technical handling and equipment requirements of CRISPR-Dx, and thus smooths the transition to the POC setting.

CRISPR-Dx have the potential to be developed as effective POC diagnostics, however notable incompatibilities between isothermal amplification methods and Cas enzymes have limited this progress. For example, the viscosity of RPA has been reported to decrease the sensitivity and diminish reliability in single-reaction SHERLOCK diagnostics(Kellner et al. 2019), and LAMP reactions operate at high temperatures (60-65ºC) beyond the optimal operation temperature of most Cas effectors (37-42ºC)(Zetsche et al. 2015; East-Seletsky et al. 2016; Abudayyeh et al. 2016; Konermann et al. 2018; Yan et al. 2018). These incompatibilities have limited the development of single-reaction CRISPR-Dx and led researchers to develop two-reaction diagnostics that are technically complex and prone to contamination. However, through discovery and innovation, significant progress has been made to further the development of CRISPR-Dx for POC settings.

4.3 Single-reaction CRISPR-Dx

The transfer steps required in earlier versions of CRISPR-Dx is the main drawback for implementing these tests in POC settings. Recently, significant progress has been made in the development of single-reaction CRISPR-Dx technologies. For example, the SHERLOCK group

recently demonstrated a system for detection of SARS-CoV-2 termed STOP (SHERLOCK Testing in One Pot)(Joung et al. 2020). To develop this assay, the group identified a thermostable Cas12 effector (AapCas12b) that demonstrates robust cleavage activity at higher temperatures(Teng et al. 2018). The thermostable property of this Cas12 enzyme enabled the integration of this unique effector with LAMP into a single reaction. Furthermore, this test utilized a simplified extraction protocol that allowed for the entire sample to be used as input leading to improved detection compared with the CDC standard RT-qPCR protocol.

STOP is a significant advance for CRISPR-Dx and demonstrates the robustness of these CRISPR-Dx tools, however, this development also highlights significant obstacles these technologies face. For instance, the use of AapCas12b underscores the need to discover more thermostable Cas effectors to enable multiplexing capabilities previously demonstrated in two-reaction systems(Gootenberg et al. 2018). Moreso, although STOP only requires a single temperature for detection, the heat requirement of this reaction increases the technological requirements for this tool, whereas POC diagnostics ideally operate at room temperature.

While much effort has been focused on integrating CRISPR-mediated detection with isothermal amplification into a single reaction, other researchers have focused solely on exploiting the robust cleavage activity of Cas effectors for direct detection of nucleic acids. By removing the amplification methods, CRISPR-Dx systems can be simplified into a single reaction with a focus on improving the effect of collateral cleavage activity. Recently, groups have developed amplification-free CRISPR-Dx for the detection of SARS-CoV-2, by exploiting the robust collateral cleavage activity of Cas13a effectors(Fozouni et al. 2021; Shinoda et al. 2021; Liu et al. 2021). To achieve sensitive detection comparable to amplification-based CRISPR-Dx, the groups screened multiple targets and combined the best targets together to

improve detection. The use of multiple targets in one reaction dramatically improved the sensitivity of these systems, however considerable variation was observed between orthologs. Using LwaCas13a, detection of 3,400 copy/μL was achieved(Shinoda et al. 2021), whereas detection of 100 copy/μL was achieved using LbuCas13a(Fozouni et al. 2021). In a later study, direct detection with LbuCas13a was enhanced by leveraging a Type-III CRISPR effector, Csm6(Jiang, Samai, and Marraffini 2016), and modifying the collateral probe to contain a Csm6 activator resulting in improved sensitivity (31 copy/μL). Although the sensitivity was significantly improved, this method required 8 targets as opposed to 3 targets used in the prior mentioned studies(Liu et al. 2021).

Amplification-free CRISPR-Dx that exploit the targeting activities of Cas effectors are very attractive for their simplicity, however notable drawbacks hinder these applications. One considerable drawback is the requirement for multiple target sites which prevents detection of single-nucleotide polymorphisms and subjects the accuracy of these diagnostics to sequence variations. Amplification-free methods also require RNA extraction to achieve detection and thus until direct-detection is achieved without extraction, application of these tests in the POC setting remains impractical. Furthermore, recent evidence has suggested a disagreement between published results and theoretical limits of direct-detection with Cas effectors presenting a clear need to evaluate these technologies further(Ramachandran and Santiago 2021). Lastly, Cas effectors preferentially target one nucleic acid species and therefore the flexibility of amplification-free CRISPR-Dx is limited to one nucleic acid species that depends on the Cas effector chosen.

As groups work to innovate the CRISPR-Dx reactions, others have decided to engineer novel devices and take advantage of the mass use of smart-phones. Two innovative methods

were developed based on the established SHERLOCK methods using Cas13 or Cas12, respectively: SHINE and miSHERLOCK(Arizti-Sanz et al. 2020; de Puig et al. 2021). SHINE created a streamlined detection method by improving HUSDON(Myhrvold et al. 2018), a previously developed lysis method, and developing an app-based readout of fluorescence to minimize sample contamination and user bias. miSHERLOCK engineered a unique lysis method and device to enable simplified lysis of viral particles and interpretation of results. The device is built to allow users to capture fluorescence results on a smartphone app that automatically quantifies and determines results, similar to the app used in SHINE. Alternatively, another group engineered a smartphone-compatible device where numerous Cas12 detection reactions run on a chip in parallel(Ning et al. 2021). This diagnostic, termed CRISPR-FDS, enables user-friendly assay set-up and interpretation, does not require an extraction step, and provides rapid turnover of results with demonstrated sensitivity comparable to RT-PCR. These innovations provide promise for engineering novel devices or developing smartphone applications to simplify CRISPR-Dx reactions and interpretation that will further streamline these diagnostic tools.

4.4 Conclusions

CRISPR-Dx remains in their nascency, yet the pace of innovation is progressing rapidly. Although still early in development, advances towards simplified CRISPR-Dx are happening rapidly and, due to the COVID-19 pandemic, has been dramatically accelerated. Future developments of CRISPR-Dx are likely to address concerns related to technical handling, sensitivity, and multiplexing capabilities necessary for POC diagnostics. With these rapid advances, CRISPR-Dx presents promising nucleic acid detection technologies that are likely to shape the future of POC diagnostics.

4.5 Acknowledgements

This chapter, in full, is a reprint of the material as it appears in Biochemistry. Daniel J Brogan and Omar S. Akbari. “CRISPR Diagnostics: Advances toward the point of care”. Biochemistry. (2022). The dissertation/thesis author was the first-author and primary investigator of this paper.

4.6 Figures

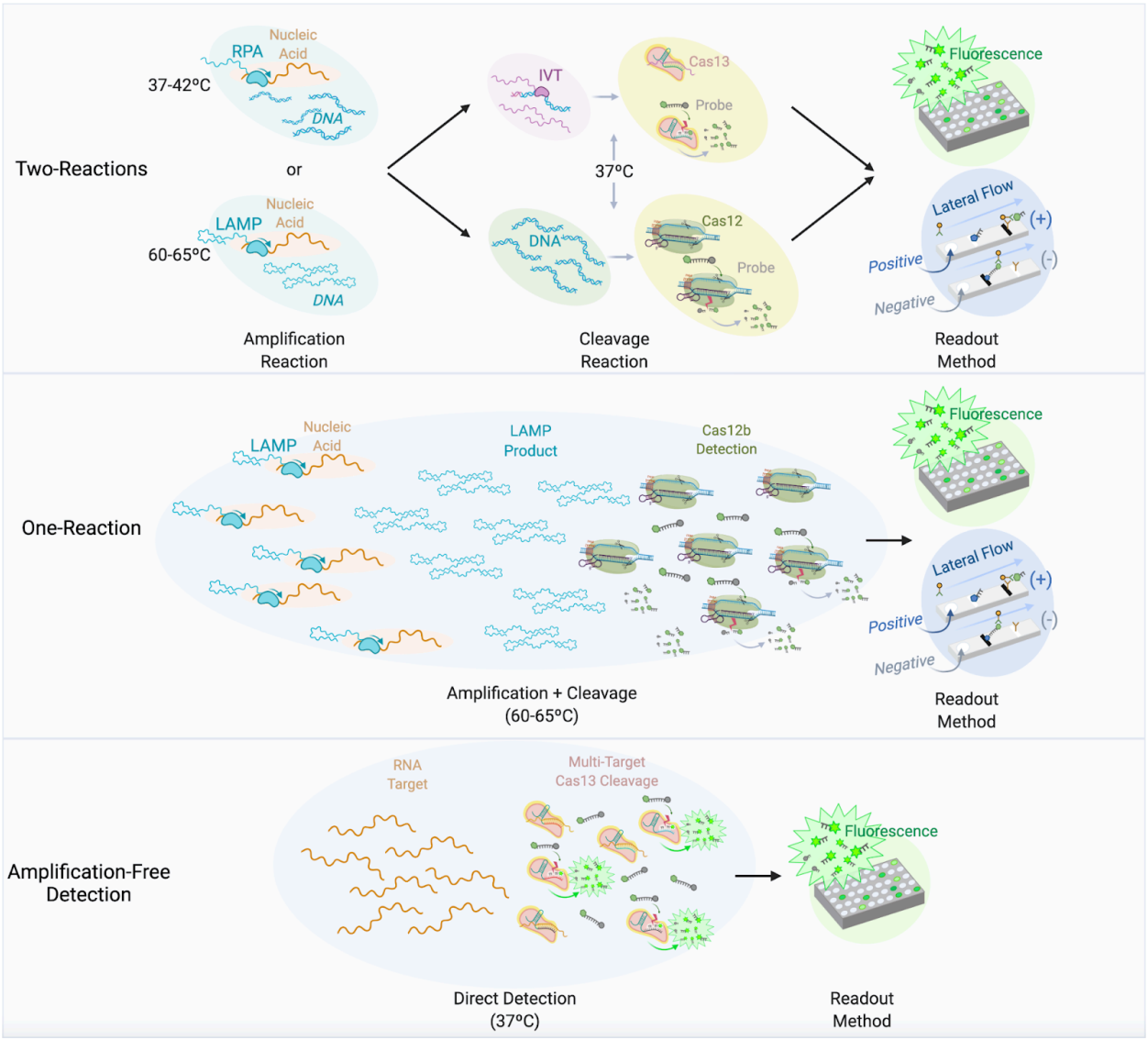


Figure 4.1: CRISPR-Dx Formats For Detection of Nucleic Acids. Schematic representation of different formats of CRISPR-Dx systems. Two-reaction diagnostics represent the original formats established for CRISPR-Dx. One-reaction diagnostics involve the combination of isothermal amplification methods and Cas detection into one-reaction, most notably represented by STOP, which uses LAMP and Cas12b for detection. Amplification-free diagnostics exploit direct detection of nucleic acids using Cas enzymes and utilize multiple targets to amplify detection signal. Only amplification-based detection methods have been established for both fluorescence and lateral flow detection, while amplification-free systems have only been established for detection via fluorescence.

4.7 Supplementary tables

Table S4.1: Summary of CRISPR-Dx Technologies Using Collateral Cleavage.

Diagnostic Tool	Cas Effector	Amplification Method	Readout Method	Reference
SHERLOCK	Cas13a, Cas13b, Cas12a, Csm6	RPA	Fluorescence/Lateral Flow	DOI: 10.1126/science.aam9321 DOI: 10.1126/science.aaq0179
DETECTR	Cas12a	RPA, LAMP	Fluorescence/Lateral Flow	DOI: 10.1126/science.aar6245 DOI: 10.1038/s41587-020-0513-4
STOP	Cas12b	LAMP	Fluorescence/Lateral Flow	DOI: 10.1056/NEJMc2026172
SENSR	Cas13d	RPA	Fluorescence/Lateral Flow	DOI: 10.1021/acssensors.1c01088
Amplification-Free Detection	Cas13a, Csm6	N/A	Fluorescence	DOI: 10.1038/nature19802 DOI: 10.1016/j.cell.2020.12.001 DOI: 10.1038/s42003-021-02001-8 DOI: 10.1038/s41589-021-00842-2

4.8 References

Abudayyeh, Omar O., Jonathan S. Gootenberg, Silvana Konermann, Julia Joung, Ian M. Slaymaker, David B. T. Cox, Sergey Shmakov, et al. 2016. "C2c2 Is a Single-Component Programmable RNA-Guided RNA-Targeting CRISPR Effector." *Science* 353 (6299): aaf5573. https://doi.org/10.1126/science.aaf5573.

Ackerman, Cheri M., Cameron Myhrvold, Sri Gowtham Thakku, Catherine A. Freije, Hayden C. Metsky, David K. Yang, Simon H. Ye, et al. 2020. "Massively Multiplexed Nucleic Acid Detection with Cas13." *Nature* 582 (7811): 277–82. https://doi.org/10.1038/s41586-020-2279-8.

Agrawal, Shreeya, Alison Fanton, Sita S. Chandrasekaran, Noam Prywes, Maria Lukarska, Scott B. Biering, Dylan C. J. Smock, et al. 2020. "Rapid Detection of SARS-CoV-2 with Cas13." *bioRxiv*. medRxiv. https://doi.org/10.1101/2020.12.14.20247874.

Arizti-Sanz, Jon, Catherine A. Freije, Alexandra C. Stanton, Brittany A. Petros, Chloe K. Boehm, Sameed Siddiqui, Bennett M. Shaw, et al. 2020. "Streamlined Inactivation, Amplification, and Cas13-Based Detection of SARS-CoV-2." *Nature Communications* 11 (1): 5921. https://doi.org/10.1038/s41467-020-19097-x.

Brogan, Daniel J., Duverney Chaverra-Rodriguez, Calvin P. Lin, Andrea L. Smidler, Ting Yang, Lenissa M. Alcantara, Igor Antoshechkin, et al. 2021. "Development of a Rapid and Sensitive CasRx-Based Diagnostic Assay for SARS-CoV-2." *ACS Sensors*, October. https://doi.org/10.1021/acssensors.1c01088.

Broughton, James P., Xianding Deng, Guixia Yu, Clare L. Fasching, Venice Servellita, Jasmeet Singh, Xin Miao, et al. 2020. "CRISPR-Cas12-Based Detection of SARS-CoV-2." *Nature Biotechnology*, April. https://doi.org/10.1038/s41587-020-0513-4.

Byrne, Andrew William, David McEvoy, Aine B. Collins, Kevin Hunt, Miriam Casey, Ann Barber, Francis Butler, et al. 2020. "Inferred Duration of Infectious Period of SARS-CoV-2: Rapid Scoping Review and Analysis of Available Evidence for Asymptomatic and Symptomatic COVID-19 Cases." *BMJ Open* 10 (8): e039856. https://doi.org/10.1136/bmjopen-2020-039856.

Chen, Janice S., Enbo Ma, Lucas B. Harrington, Maria Da Costa, Xinran Tian, Joel M. Palefsky, and Jennifer A. Doudna. 2018. "CRISPR-Cas12a Target Binding Unleashes Indiscriminate Single-Stranded DNase Activity." *Science* 360 (6387): 436–39. https://doi.org/10.1126/science.aar6245.

East-Seletsky, Alexandra, Mitchell R. O'Connell, Spencer C. Knight, David Burstein, Jamie H. D. Cate, Robert Tjian, and Jennifer A. Doudna. 2016. "Two Distinct RNase Activities of CRISPR-C2c2 Enable Guide-RNA Processing and RNA Detection." *Nature* 538 (7624): 270–73. https://doi.org/10.1038/nature19802.

Fozouni, Parinaz, Sungmin Son, María Díaz de León Derby, Gavin J. Knott, Carley N. Gray, Michael V. D'Ambrosio, Chunyu Zhao, et al. 2021. "Amplification-Free Detection of

SARS-CoV-2 with CRISPR-Cas13a and Mobile Phone Microscopy." *Cell* 184 (2): 323–33.e9. https://doi.org/10.1016/j.cell.2020.12.001.

Gootenberg, Jonathan S., Omar O. Abudayyeh, Max J. Kellner, Julia Joung, James J. Collins, and Feng Zhang. 2018. "Multiplexed and Portable Nucleic Acid Detection Platform with Cas13, Cas12a, and Csm6." *Science* 360 (6387): 439–44. https://doi.org/10.1126/science.aaq0179.

Gootenberg, Jonathan S., Omar O. Abudayyeh, Jeong Wook Lee, Patrick Essletzbichler, Aaron J. Dy, Julia Joung, Vanessa Verdine, et al. 2017. "Nucleic Acid Detection with CRISPR-Cas13a/C2c2." *Science*. https://doi.org/10.1126/science.aam9321.

Howson, Emma L. A., Yohei Kurosaki, Jiro Yasuda, Masayoshi Takahashi, Hiroaki Goto, Ashley R. Gray, Valerie Mioulet, Donald P. King, and Veronica L. Fowler. 2017. "Defining the Relative Performance of Isothermal Assays That Can Be Used for Rapid and Sensitive Detection of Foot-and-Mouth Disease Virus." *Journal of Virological Methods* 249 (November): 102–10. https://doi.org/10.1016/j.jviromet.2017.08.013.

Jiang, Wenyan, Poulami Samai, and Luciano A. Marraffini. 2016. "Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity." *Cell* 164 (4): 710–21. https://doi.org/10.1016/j.cell.2015.12.053.

Joung, Julia, Alim Ladha, Makoto Saito, Nam-Gyun Kim, Ann E. Woolley, Michael Segel, Robert P. J. Barretto, et al. 2020. "Detection of SARS-CoV-2 with SHERLOCK One-Pot Testing." *The New England Journal of Medicine*, September. https://doi.org/10.1056/NEJMc2026172.

Kellner, Max J., Jeremy G. Koob, Jonathan S. Gootenberg, Omar O. Abudayyeh, and Feng Zhang. 2019. "SHERLOCK: Nucleic Acid Detection with CRISPR Nucleases." *Nature Protocols* 14 (10): 2986–3012. https://doi.org/10.1038/s41596-019-0210-2.

Konermann, Silvana, Peter Lotfy, Nicholas J. Brideau, Jennifer Oki, Maxim N. Shokhirev, and Patrick D. Hsu. 2018. "Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors." *Cell* 173 (3): 665–76.e14. https://doi.org/10.1016/j.cell.2018.02.033.

Li, Shi-Yuan, Qiu-Xiang Cheng, Jia-Kun Liu, Xiao-Qun Nie, Guo-Ping Zhao, and Jin Wang. 2018. "CRISPR-Cas12a Has Both Cis- and Trans-Cleavage Activities on Single-Stranded DNA." *Cell Research* 28 (4): 491–93. https://doi.org/10.1038/s41422-018-0022-x.

Liu, Tina Y., Gavin J. Knott, Dylan C. J. Smock, John J. Desmarais, Sungmin Son, Abdul Bhuiya, Shrutee Jakhanwal, et al. 2021. "Accelerated RNA Detection Using Tandem CRISPR Nucleases." *Nature Chemical Biology* 17 (9): 982–88. https://doi.org/10.1038/s41589-021-00842-2.

Myhrvold, Cameron, Catherine A. Freije, Jonathan S. Gootenberg, Omar O. Abudayyeh, Hayden C. Metsky, Ann F. Durbin, Max J. Kellner, et al. 2018. "Field-Deployable Viral Diagnostics Using CRISPR-Cas13." *Science* 360 (6387): 444–48. https://doi.org/10.1126/science.aas8836.

Ning, Bo, Tao Yu, Shengwei Zhang, Zhen Huang, Di Tian, Zhen Lin, Alex Niu, et al. 2021. "A Smartphone-Read Ultrasensitive and Quantitative Saliva Test for COVID-19." *Science Advances* 7 (2). https://doi.org/10.1126/sciadv.abe3703.

Notomi, T., H. Okayama, H. Masubuchi, T. Yonekawa, K. Watanabe, N. Amino, and T. Hase. 2000. "Loop-Mediated Isothermal Amplification of DNA." *Nucleic Acids Research* 28 (12): E63. https://doi.org/10.1093/nar/28.12.e63.

Pan, Yang, Daitao Zhang, Peng Yang, Leo L. M. Poon, and Quanyi Wang. 2020. "Viral Load of SARS-CoV-2 in Clinical Samples." *The Lancet Infectious Diseases* 20 (4): 411–12. https://doi.org/10.1016/S1473-3099(20)30113-4.

Piepenburg, Olaf, Colin H. Williams, Derek L. Stemple, and Niall A. Armes. 2006. "DNA Detection Using Recombination Proteins." *PLoS Biology* 4 (7): e204. https://doi.org/10.1371/journal.pbio.0040204.

Pilarowski, Genay, Paul Lebel, Sara Sunshine, Jamin Liu, Emily Crawford, Carina Marquez, Luis Rubio, et al. 2021. "Performance Characteristics of a Rapid Severe Acute Respiratory Syndrome Coronavirus 2 Antigen Detection Assay at a Public Plaza Testing Site in San Francisco." *The Journal of Infectious Diseases* 223 (7): 1139–44. https://doi.org/10.1093/infdis/jiaa802.

Puig, Helena de, Rose A. Lee, Devora Najjar, Xiao Tan, Luis R. Soeknsen, Nicolaas M. Angenent-Mari, Nina M. Donghia, et al. 2021. "Minimally Instrumented SHERLOCK (miSHERLOCK) for CRISPR-Based Point-of-Care Diagnosis of SARS-CoV-2 and Emerging Variants." *Science Advances* 7 (32). https://doi.org/10.1126/sciadv.abh2944.

Ramachandran, Ashwin, and Juan G. Santiago. 2021. "CRISPR Enzyme Kinetics for Molecular Diagnostics." *Analytical Chemistry* 93 (20): 7456–64. https://doi.org/10.1021/acs.analchem.1c00525.

Shinoda, Hajime, Yuya Taguchi, Ryoya Nakagawa, Asami Makino, Sae Okazaki, Masahiro Nakano, Yukiko Muramoto, et al. 2021. "Amplification-Free RNA Detection with CRISPR-Cas13." *Communications Biology* 4 (1): 476. https://doi.org/10.1038/s42003-021-02001-8.

Smargon, Aaron A., David B. T. Cox, Neena K. Pyzocha, Kaijie Zheng, Ian M. Slaymaker, Jonathan S. Gootenberg, Omar A. Abudayyeh, et al. 2017. "Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28." *Molecular Cell*. https://doi.org/10.1016/j.molcel.2016.12.023.

S. Pavia, Charles, and Maria M. Plummer. 2021. "The Evolution of Rapid Antigen Detection Systems and Their Application for COVID-19 and Other Serious Respiratory Infectious Diseases." *Journal of Microbiology, Immunology, and Infection = Wei Mian Yu Gan Ran Za Zhi* 54 (5): 776–86. https://doi.org/10.1016/j.jmii.2021.06.003.

Teng, Fei, Tongtong Cui, Guihai Feng, Lu Guo, Kai Xu, Qingqin Gao, Tianda Li, Jing Li, Qi Zhou, and Wei Li. 2018. “Repurposing CRISPR-Cas12b for Mammalian Genome Engineering.” *Cell Discovery* 4 (November): 63. https://doi.org/10.1038/s41421-018-0069-3.

Valera, Enrique, Aaron Jankelow, Jongwon Lim, Victoria Kindratenko, Anurup Ganguli, Karen White, James Kumar, and Rashid Bashir. 2021. “COVID-19 Point-of-Care Diagnostics: Present and Future.” *ACS Nano* 15 (5): 7899–7906. https://doi.org/10.1021/acsnano.1c02981.

Yan, Winston X., Shaorong Chong, Huaibin Zhang, Kira S. Makarova, Eugene V. Koonin, David R. Cheng, and David A. Scott. 2018. “Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein.” *Molecular Cell* 70 (2): 327–39.e5. https://doi.org/10.1016/j.molcel.2018.02.028.

Zetsche, Bernd, Jonathan S. Gootenberg, Omar O. Abudayyeh, Ian M. Slaymaker, Kira S. Makarova, Patrick Essletzbichler, Sara E. Volz, et al. 2015. “Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System.” *Cell* 163 (3): 759–71. https://doi.org/10.1016/j.cell.2015.09.038.

Chapter 5: Engineering type III-E CRISPR-Cas systems to develop highly effective and specific RNA-therapeutic technologies

5.1 Abstract

The recent discovery of the type III-E CRISPR-Cas effector class has reshaped our fundamental understanding of CRISPR-Cas evolution and classification. Type III-E effectors are composed of several Cas7-like domains and a single Cas11-like domain naturally fused together to create a single polypeptide capable of programmably targeting and degrading RNA. Here we identify a novel composition of a type III-E-like effector where the function of the Cas11 domain is replaced by a Cas1 domain, presenting a new function of Cas1 in RNA-targeting. Furthermore, we demonstrate a unique modularity of type III-E effectors by methodically substituting domains between orthogonal type III-E proteins to engineer compact Synthetic Cas (SynCas) effectors. We refine our methods to generate several SynCas effectors for programmable RNA-targeting in mammalian cells. SynCas represents a new understanding of type III-E architecture and modularity, and provides a platform for engineering powerful and compact genome engineering technologies from the blueprint of nature.

5.2 Introduction

The continuous evolution of CRISPR-Cas systems has resulted in the discovery and engineering of revolutionary genome engineering technologies. Recently, the type III-E CRISPR-Cas system was discovered[1] and characterized [2,3] and has changed how we understand CRISPR-Cas protein evolution. Type III-E effectors are categorized as class 1 CRISPR systems, which typically utilize a multisubunit complex formation to achieve DNA- and/or RNA-targeting[1]. However, type III-E effectors accomplish programmable RNA-targeting by

employing a single, large protein composed of four Cas7-like domains, one Cas11 domain, and a large insertion (INS) domain, hence being named Cas7-11[4,5] (or gRAMP for giant Repeat Associated Mysterious Protein[6]). Recent studies have demonstrated that type III-E effectors are capable of programmable RNA degradation in prokaryotic and eukaryotic cells, and lack RNA collateral cleavage activity that consistently limits applications of Cas13 systems[7–10]. Using prokaryotic expression, it was demonstrated that Cas7-11 target binding activates a downstream caspase activity that results in cell dormancy, or death, and appears to serve as the primary immune response for type III-E systems[6,11,12]. Therefore when type III-E effectors are expressed independently, they operate as highly specific ribonucleases and their unique architecture provides a pathway for further engineering.

We envisioned type III-E systems as a group of modular proteins with interchangeable orthologous domains. We reasoned that the unique assembly of, essentially, individual proteins into a single polypeptide provided evidence of a natural fusion system. We hypothesized that interchanging domains between orthologs may enhance certain functions, such as improving catalytic activity. Here we identify a novel architecture of a type III-E-like protein composed of Cas7-like domains and a Cas1-like domain. We use this novel composition as inspiration to engineer recently characterized type III-E effectors into Synthetic Cas (SynCas) effectors. We uncover a novel function for the Cas1 domain in RNA-targeting, and provide a method for designing synthetic RNA-targeting Cas effectors. Together this study provides a new understanding of CRISPR biology, as well as an engineering method to produce compact and optimized synthetic Cas effectors capable of targeting RNA.

5.3 Results

5.3.1 Identification of novel type III-E-like architecture containing a Cas1 domain

To begin engineering type III-E effectors, we initially conducted a bioinformatic search and alignment to create a domain library (Table S1) and identify any variants with possible novel architectures. In our initial search we obtained an unusual blast hit for a protein denoted as a Cas1 nuclease (MBU1487208.1). We analyzed the domain architecture of this protein using HHPred and found the protein to be composed of three Cas7-like domains and one Cas1 nuclease domain (Figure S1A-E). For simplicity, we term this protein Cas7-1. Cas7-1 aligns well to all known type III-E effectors and through further analysis we found interesting features of the protein. Cas7-1 appears to lack an INS domain – a domain found in all known type III-E effectors[3] (Figure 1A). We also found Cas7-1 contains a conserved catalytic residue in the Cas7.3 domain and zinc-finger motifs unique to type III-E effectors[4] (Figure 1B-E). Lastly, the Cas1 domain aligns to a Reverse Transcriptase-Cas1 (RT-Cas1) fusion protein commonly associated with type III systems of all characterized subtypes (III-A, III-B, III-C, and III-D) (Figure 1F, File S1), where an active Cas1 domain is required for adaptation of new spacers from interfering RNA[13]. However, from the alignments and structural prediction (Figure 1G), it appears Cas7-1 is missing a portion of the N-terminus, including a CRISPR RNA (crRNA) processing domain and a secondary catalytic residue (File S2).

5.3.2 Engineering synthetic Cas effectors with varying architectures

Based on prior characterization of type III-E effectors, we suspected the missing domains of Cas7-1 would abolish crRNA recognition and processing and, possibly, cleavage activity. Therefore, we reasoned we could rescue activity of the Cas7-1 protein through the incorporation

of Cas7.1 and Cas7.2 domains from *D. ishimotonii* Cas7-11 (DiCas7-11). We also hypothesized that these SynCas effectors could be generated using a method we term Orthologous Domain Substitution (ODS), where aligned domains in the type III-E family are substituted into a synthetic protein sequence. To rescue Cas7-1 activity and test ODS we designed 16 initial SynCas effectors with different domain and linker substitutions, while simultaneously varying the overall architecture based on Cas7-1 or typical type III-E effectors (Figure 2A, Figure S2, Figure S3A, Table S2). We also analyzed all protein sequences with AlphaFold[14–16] to obtain predicted structures for later interrogation (File S3-25).

Since no previous study had utilized a Cas1 domain in programmable RNA-targeting, we focused on SynCas variants that were either Cas1-based or Cas11-based (Figure S2). We divided the groups by architecture, aiming to determine if the Cas7-1 or typical type III-E architecture reigned supreme. We tested our SynCas variants in HEK293T cells using an EGFP reporter assay and analyzed samples via qPCR. crRNAs used in assessment contained direct repeats (DRs) respective to DiCas7-11 and spacer sequences complementary to EGFP (Table S3). Although we observed significant RNA knockdown with the Cas7-1 architecture (SynCas.v1-v9), most variants exhibited high variability and the best variant (SynCas.v7) included the large INS domain (Figure 2B). In the type III-E architecture (SynCas.v10-v16), we observed consistent RNA knockdown comparable to DiCas7-11 from multiple variants including one Cas1-based variant (SynCas.v14) (Figure 2B). We identified SynCas.v10-12 (where SynCas.v12 is nearly identical to Cas7-11S[4]) as the best Cas11-based effectors and SynCas.v14 as the best Cas1-based effector.

Due to the natural orientation of the Cas1 domain on the C-terminus of Cas7-1, we suspected the N-terminal region of the domain would be unstructured and highly flexible in order

to help orient the Cas1 domain underneath target substrates. We identified a 29AA stretch that aligned specifically to the N-terminus of the RT portion of RT-Cas1 fusion proteins (Figure S4, File S26). Using the predicted structure of SynCas.v14, we generated two truncated versions of the Cas1 domain (Figure 2C) and tested these truncations with SynCas.v14, SynCas.v15, and SynCas.v16. Aside from SynCas.v15B, we found both truncations generally improved knockdown activity for the SynCas variants and observed that complete deletion of the linker region consistently improved activity likely due to increased stabilization of the Cas1 domain (Figure 2D). Together these findings demonstrate RNA-targeting type III-E effectors are highly modular and Cas1 can be used in RNA-targeting applications.

5.3.3 Programmable *in vitro* cleavage of ssRNA with SynCas effectors

With evidence of RNA knockdown in cell culture, we next wanted to characterize the mechanism of cleavage of the SynCas effectors. We focused on SynCas.v10 to demonstrate that substitution of enzymatic domains permits retention of enzymatic activity and that RNA-targeting is unperturbed. Since SynCas.v10 contains the Cas7.1 domain deriving from DiCas7-11, all crRNAs were designed with DRs respective to DiCas7-11 (Table S3). We ran an initial assay truncating the spacer of the crRNA and observed cleavage for spacers between 20-30nt in length. Interestingly, we noticed a substantial increase in both crRNA processing and target cleavage as spacers were shortened, with the best cleavage activity observed for 22-24nt spacers (Figure S5A). This increase in cleavage activity due to spacer truncation was further replicated with SynCas.v10 in cell culture via qPCR analysis (Figure 3A). We also incubated the cleavage reaction for various times and observed near-complete cleavage after 1.5 hours and

complete cleavage (cleavage by both Cas7.2 and Cas7.3) of the target RNA after 6 hours of incubation (Figure S5B).

To visualize both crRNA processing and target cleavage, we designed several crRNAs across a 40nt ssRNA target and incubated the SynCas effectors with unprocessed crRNAs (Figure 3B). Tiling the crRNAs across the ssRNA target and incubating for 6 hours to obtain complete cleavage, we observed stepwise cleavage products demonstrating activity of both enzymatic domains (Figure 3C). Cleavage activity was completely abolished through mutagenesis of both active sites for DiCas7-11, SynCas.v10, and SynCas.v12 (D429A/D654A), while crRNA processing was unaffected (Figure 3D).

We investigated the crRNA processing activity further by introducing a full array structure containing the same spacer as crRNA-5 (Figure 3E, Table S3). Differing from other reports[2,3], we found that neither SynCas.v12, nor wild-type (WT) DiCas7-11, could completely process an array structure into a mature 37nt crRNA, but instead yielded an incompletely processed 57nt product (Figure 3F). We found only the Cas7.1 domain processed the DR producing a product where a 15nt mature DR (mDR) is upstream of the spacer and a 20nt DR fragment remains at the 3' end (Figure 3E). This suggests that type III-E effectors do not cleave off the 3' end downstream of the spacer when processing arrays. However, by modifying the array structure to incorporate mDRs surrounding the spacer sequence we were able to obtain a crRNA without excess nucleotides on the 3' end (Figure 3G). This array design is extremely important for targeting and multiplexing applications with SynCas effectors, as we found SynCas.v10, which contains Cas7.3/Cas7.4/CTE domains different than DiCas7-11, appears to require the mDR array structure to accomplish efficient RNA cleavage (Figure 3H). These

results suggest that SynCas effectors accomplish RNA cleavage like other type III-E proteins and type III-E systems use a unique crRNA structure and processing mechanism.

5.3.4 Generation of compact SynCas effectors for RNA-targeting

Type III-E CRISPR systems are the largest single-effector CRISPR protein family discovered[2,3]. In therapeutic applications, this large size is a major obstacle for delivery modalities and efforts have quickly been made to reduce the size of these enzymes[4]. We selected three effectors (DiCas7-11, HvsCas7-11, and Cas7-1) to generate a set of 36 compact SynCas effectors and further refine our ODS design method for systematic interrogation of domain arrangement (Figure 4A, Figure S6, Figure S3B, File S27-S62). The compact effectors varied in size from 1233-1292AA in length and far smaller than the type III-E average size (>1600AA) (Table S2).

To assess the RNA knockdown capabilities of the compact SynCas effectors, we analyzed these effectors using the previously described reporter assay in HEK293T cells, and quantified knockdown using qPCR or flow cytometry (Figure 4A). We found that SynCas effectors containing the DiCas7.1 domain resulted in consistent EGFP knockdown comparable to DiCas7-11, while effectors containing the HvsCas7.1 domain demonstrated little-to-no knockdown activity (Figure 4B). Within the DiCas7.1 groups, we also found that either DiCas11- or HvsCas11-based variants greatly outperformed SynCas effectors with a Cas1 domain, contrary to prior analysis of the first 16 variants, where SynCas.v14A demonstrated substantial knockdown capabilities. Of the Cas11-based effectors, SynCas.v35, SynCas.v39, and SynCas.v41 demonstrated knockdown comparable to DiCas7-11, where SynCas.v41 is the most compact (1241AA) and represents the smallest type III-E effector engineered (Figure S3C).

Interestingly, composing the SynCas effector of HvsCas7.2 and HvsCas7.3 resulted in high RNA knockdown regardless of the Cas11/Cas1 domain used, which was not observed for any other Cas7.2-Cas7.3 combination (Figure 4B, Figure S3D).

We further analyzed the SynCas effectors to gauge their ability to reduce the expression of target proteins. Using flow cytometry, we observed knockdown of EGFP and mCherry with SynCas.41 and SynCas.v47 at comparable levels to DiCas7-11 for both targets (Figure 4C). This demonstrates that refined SynCas effectors can elicit phenotypic effects on a target of interest regardless of domain composition. Furthermore, we demonstrate SynCas effectors are capable of knocking down endogenous genes as effectively as DiCas7-11 (Figure 4D). We also show SynCas effectors do not possess the collateral cleavage activity found in other CRISPR-based RNA-targeting systems, such as Cas13. We accomplished this by running SENSR reactions[17] with SynCas effectors and comparing them to RfxCas13d. From the fluorescence assays, we found no evidence of collateral cleavage activity with any effectors composed of type III-E domains (Figure 4E), consistent with previous results[3]. Here we demonstrate that the type III-E proteins are modular and can be engineered into compact, efficient, and highly specific SynCas effectors for RNA-targeting applications.

5.4 Discussion

As new discoveries of CRISPR-Cas systems emerge, the continuous evolution of these enzymes reveals new possibilities for generating new genome engineering technologies. Here we identify a novel polypeptide composition similar to that of type III-E effectors, termed Cas7-1, where the Cas11 domain is replaced by a Cas1 domain. Cas7-1 is also the first example of a type III-E-like effector naturally lacking an INS domain and demonstrating the natural evolution of a

compact effector. Using Cas7-1 and type III-E architectures, we developed an engineering method (ODS) to generate synthetic type III-E effectors capable of RNA-targeting. The ODS method reveals a unique modularity of the type III-E effector class, where orthologous domains are interchangeable and minimally impact fold or activity of SynCas effectors in mammalian cells. Furthermore, we demonstrate the applicability of this design method for generating the most compact type III-E effector to date with further improvements and applications possible.

Along with presenting an engineering method that manipulates the modularity of type III-E effectors, we also uncover a unique crRNA processing activity where the mature crRNA contains a spacer surrounded by fragments of the DR on both the 5' and 3' end. This mode of crRNA processing has not been observed in any other single-effector CRISPR system capable of array processing[18,19]. For SynCas effectors, we found this crRNA processing activity to directly impact both array processing and targeting activity and we provide a solution for crRNA array design when applying SynCas effectors for RNA-targeting applications using a mDR array structure. These findings demonstrate the need for further exploration and characterization of type III-E systems, especially in the context of downstream engineering.

The ODS method and SynCas effectors are useful for protein and transcriptome engineering, respectively. Although we observed clear evidence of RNA-targeting *in vitro* and in mammalian cells, the level of activity exhibited by SynCas effectors in mammalian cells did not generally outperform WT DiCas7-11. This activity gap suggests further refinement is required to improve engineering of SynCas effectors. This refinement can be accomplished with larger libraries of SynCas effectors, incorporating machine learning to expedite library design, or employing directed evolution methods[20]. Type III-E RNA cleavage activity is inherently slow due to the evolution of a caspase activity for antiviral defense, which operates more effectively

with longer target binding[6,11,12]. Therefore, future studies should focus on increasing the turnover rate for these enzymes to improve their activity, while maintaining the high target specificity they already possess. Taken together, this study represents a significant leap in engineering type III-E CRISPR enzymes to generate improved research tools and effective therapeutics.

5.5 Methods

5.5.1 Design and cloning of constructs

Plasmids for expression of effectors in HEK293T cells were designed with a CMV promoter and bovine growth hormone (bGH) terminator. Effector plasmids were generated using standard gibson assembly methods. crRNA plasmids were designed for expression driven by a U6 promoter and terminator. crRNA expression plasmids were generated using standard golden gate assembly methods. Select plasmids are available at https://www.addgene.org/ (Table S4).

5.5.2 Mammalian cell culture

HEK293T cell lines were obtained from the American Type Culture Collection (ATCC CRL-3216) and used in all mammalian cell experiments. HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Thermo Fisher Scientific 11995073) supplemented with 10% Fetal Bovine Serum (FBS) (Corning 35-011-CV) and 1% penicillin-streptomycin (Thermo Fisher Scientific 15070063). Transient transfections were carried out with Lipofectamine 3000 (Thermo Fisher Scientific L3000001) according to the manufacturer's protocol.

5.5.3 RNA-knockdown Assay

To assess RNA knockdown, HEK293T cells were co-transfected with 20ng of either the EGFP or mCherry reporter plasmid, 600ng of a respective Cas effector plasmid, and 300ng of crRNA plasmid. Cells were seeded in 48-wells 18-20 hours prior to transfection at seeding densities of 32,000 cells per well. Following transfection, cells were incubated for 48 hours at 37ºC, 5% CO_2. After 48 hours, cells were removed from the plate with 250µL of RNAprotect Cell Reagent (Qiagen 76104) and stored at -20ºC for at least 24 hours prior to extraction. For each effector tested, four biological replicates were performed and a non-target crRNA was used as a control.

5.5.4 Total RNA collection and qPCR

To measure the reduction of endogenous or reporter genes, the transfected cells, preserved in RNAprotect, were lysed using QIAshredder (Qiagen 79656). Total RNA was then extracted using the RNeasy Mini Kit (Qiagen 74106) according to the manufacturer's protocol. Following extraction, the total RNA was treated with DNase (Thermo Fisher Scientific AM1907) to remove any remaining DNA. Sample RNA concentration was analyzed using a Nanodrop OneC UV-vis spectrophotometer (Thermo Fisher Scientific NDONEC-W). Approximately 0.5µg of total RNA was used to synthesize cDNA with a RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific K1622). cDNA was diluted (1:1000) for qPCR analysis. qPCR was performed with SYBR green (qPCRBIO SyGreen Blue Mix Separate-ROX 17-507B, Genesee Scientific), using 4µl of dilute cDNA in each 20µl reaction containing a final primer concentration of 200nM and 10µl of SYBR green buffer solution. Three technical replicates were performed for each reaction. Pipetting was performed using EpMotion® 5075

(Eppendorf). The following qPCR protocol was used on the LightCycler® 96 (Roche): 3 min of activation phase at 95°C, 40 cycles of 5 sec at 95°C, and 30 sec at 60°C. The relative expression levels of EGFP were calculated using the manufacturer's software and the delta-delta Ct method ($2^{-\Delta\Delta Ct}$), with the GAPDH gene serving as a reference[21]. To assess differences in EGFP expression between the non-target and targeting crRNA conditions for each effector, statistical analysis was performed using GraphPad Prism 10. Specifically, data were analyzed using an unpaired t-test.

5.5.5 Flow Cytometry

Fluorescence knockdown was assessed using flow cytometry on the BioRad S3e cell sorter system. To prepare cells for analysis, transfected HEK293T cells were washed with PBS pH 7.4 (Gibco 10010031), all media was removed, then 40µL of Accumax (VWR AM105) was added to the well. 230µL of cold PBS pH 7.4 was added to the wells and cells were resuspended and transferred to a cell strainer (Corning 352235). Cells were then assessed on the S3e cell sorter with a count of 80,000 cells per sample. Flow cytometry data were analyzed using Floreada.io (https://floreada.io/) (Figure S7).

5.5.6 Recombinant Protein Expression and Purification

Plasmids used for protein expression in *E. coli* were generated by subcloning ORFs from plasmids used for mammalian cell culture expression, along with an N-terminal 6xHis-SUMO into a pET28a vector backbone. Plasmids were then transformed into Rosetta™ 2(DE3)pLysS Competent Cells (Novagen). Single colonies were used to inoculate a 40mL LB overnight starter culture. The starter cultures were then transferred to a 1L LB culture in 4L baffled flasks and

grown at 37°C with shaking (200 rpm) until OD600 reached ~1.0. Cultures were then brought to 4°C before protein expression was induced with 0.2 mM isopropyl β-D-thiogalactopyranoside (IPTG) and grown for 18-20 hours at 18°C. Cells were then pelleted and resuspended in a lysis buffer containing 50mM Tris-HCl pH 8.0, 300mM NaCl, 10mM imidazole, 5% glycerol (v/v), and 3 mM β-mercaptoethanol supplemented with Protease Inhibitor Cocktail (Sigma P2714), 5mM PMSF, and 2.5 U/mL salt active nuclease (Sigma SRE0015). Cells were lysed via sonication and clarified by centrifugation. His-tagged protein was bound to 2mL HisPur™ Ni-NTA Resin (Thermo 88221) equilibrated with lysis buffer in a glass Econo-Column® Chromatography Columns (Bio-Rad) by flowing through the clarified lysate. A wash step was then performed by adding 12 column volume (CV) of a wash buffer containing 50mM Tris-HCl pH 8.0, 500mM NaCl, 30mM imidazole, and 5% glycerol followed by elution with 7 CV elution buffer containing 50mM Tris-HCl pH 8.0, 300 mM NaCl, 300mM imidazole, and 5% glycerol. The 6xHis-SUMO tag was cleaved off by dialyzing the eluate overnight at 4°C with ~0.6 mg of in-house produced 6xHis-tagged Ulp1 (Yeast SUMO Protease) in a dialysis buffer containing 20mM HEPES-NaOH pH 7.5, 250mM NaCl, 5% glycerol, and 1mM DTT. Dialyzed sample was then flowed over the same Ni-NTA column equilibrated in a buffer containing 20mM HEPES-NaOH pH 7.5, 250mM NaCl, 25mM imidazole, 5% glycerol, and 1mM DTT to remove additional *E. coli* impurities, Ulp1, and the 6xHis-SUMO tag. Cation exchange chromatography was performed with 2x1mL HiTrap Heparin HP Column (Cytiva) with a NaCl gradient from 200-1000mM NaCl followed by gel filtration chromatography with a HiLoad 16/600 Superdex 200 column equilibrated in 20mM HEPES-NaOH pH 7.5, 600mM NaCl, 5% glycerol (v/v), 2mM DTT on an ÄKTA Pure (Cytiva). Proteins were then concentrated to ~10μM and stored in small aliquots at -80°C for future use.

5.5.7 Nucleic acid target and crRNA preparation

Synthetic ssRNA templates and the full DiCas7-11 crRNA array were ordered custom from IDT. All other crRNAs were produced in house. To generate crRNAs, a dsDNA template with a T7 promoter incorporated was produced through templateless PCR and purified with MinElute PCR Purification Kit (Qiagen 28004). dsDNA templates were converted to ssRNA through *in vitro* transcription using MEGAScript™ T7 Transcription Kit (Invitrogen AM1334) and purified with MEGAClear™ Transcription Clean-Up Kit (Invitrogen AM1908).

5.5.8 Nuclease assays

Nuclease assays were run in 10µL reactions with the following final concentrations: 500nM protein, 200nM crRNA, 240nM probe, 40U RNase Inhibitor, and 10mM $MgCl_2$. Reactions were supplemented with a custom 10X reaction buffer (200mM HEPES (pH 7.5), 600mM NaCl) and incubated at 37ºC. Times were varied based on experiment. Upon completion, reactions were denatured with 2X RNA dye (NEB B0363S) at 95ºC for 10 min. 10µl of dyed samples were loaded onto a prerun (200V for 60 min) 15% TBE-Urea PAGE gel (BioRad 4566055) and resolved at 200V for 35 min. Gels were then stained with SYBR Gold (Invitrogen S11494) and incubated for 10 min at room temperature on a shaker and washed in 1X TBE for 10 min on a shaker. Stained gels were imaged using Enduro™ GDS (Labnet).

SENSR assays were conducted following a previously described protocol[17] with slight alterations: 100nM of RfxCas13d incubated in reaction and 200nM of DiCas7-11 or SynCas effector incubated in reaction.

5.6 Acknowledgements

This chapter, in full, is a copy of the material currently in preparation for publication and submitted to Nature Biotechnology. Daniel J Brogan, Elena Dalla Benetta, Tianqi Wang, Calvin P. Lin, Fangying Chen, Harry Li, Claire Lin, Elizabeth A. Komives, Omar S. Akbari. “Engineered type III-E CRISPR-Cas effectors”. (2023). The dissertation/thesis author is the first-author and primary investigator of this paper.

5.7 Figures

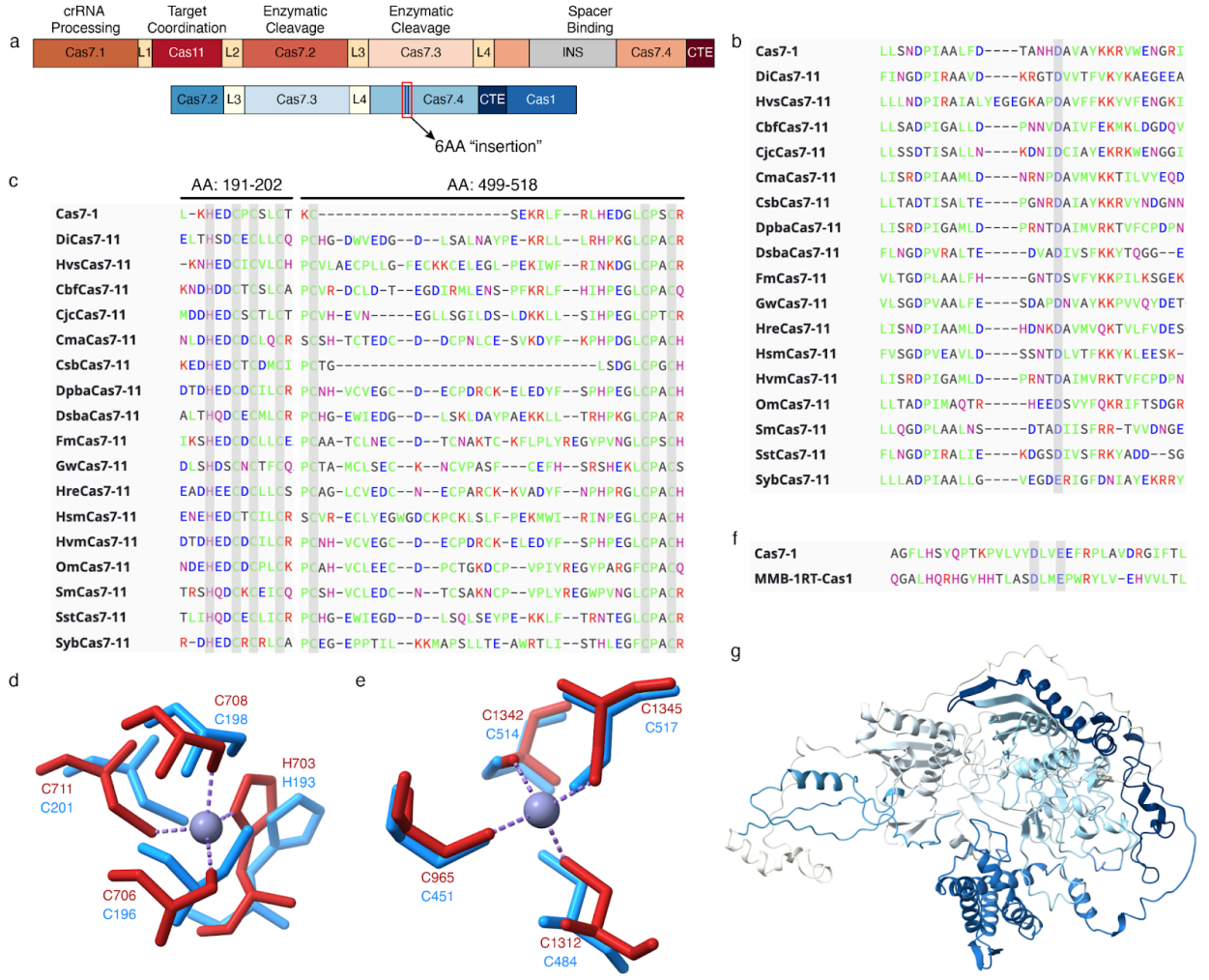


Figure 5.1: Cas7-1 sequence and predictive structure in relation to type III-E effectors. (A) Schematic representation of a typical type III-E architecture (red) and the architecture of Cas7-1 (blue). (B) Alignment of catalytic residue in the Cas7.3 domain of Cas7-1 against published type III-E effectors[3]. (C) Alignment of zinc finger residues in the Cas7.3 and Cas7.4 domains of Cas7-1 and type III-E effectors. Amino Acid (AA) position represents position in Cas7-1 protein sequence. (D) Structural alignment of Cas7.3 zinc finger residues of Cas7-1 (blue; H193/C196/C198/C201) with aligned residues of DiCas7-11 (red; H703/C706/C708/C711). (E) Structural alignment of Cas7.4 zinc finger residues of Cas7-1 (blue; C451/C484/C514/C517) with aligned residues of DiCas7-11 (red; C965/C1312/C1342/C1345). (F) Alignment of Cas1 domain of Cas7-1 to MMB1 RT-Cas1 fusion protein (NCBI Ascension here). Highlighted sections represent catalytic domains of Cas1, where E847 of Cas7-1 aligns to E870 of MMB1 RT-Cas1. (G) Predicted structure of Cas7-1 obtained from AlphaFold 2. Coloring of domains is as depicted in (A).

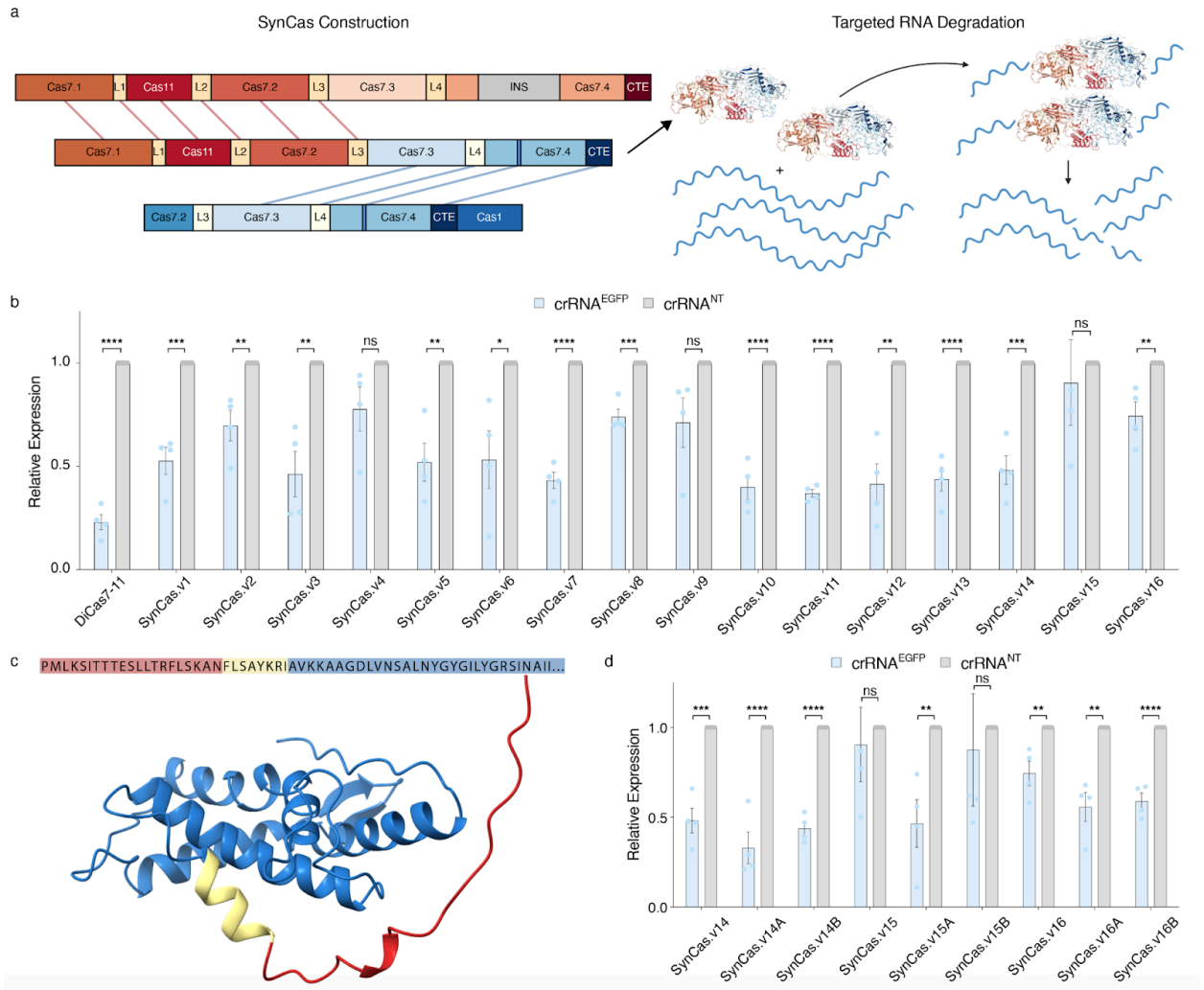


Figure 5.2: SynCas construction and analysis of initial variants. (A) Schematic representing the construction of SynCas.v10 and a depiction of RNA degradation by SynCas.v10. Red lines represent domains from DiCas7-11 and blue lines represent domains from Cas7-1. (B) qPCR analysis of EGFP targeted knockdown with SynCas variants that have a C-terminal orientation of the Cas1 domain and SynCas variants that orient the Cas11 or Cas1 domain internally in the typical type III-E architecture. (C) Predicted structure and partial amino acid sequence of Cas1 domain with highlighted areas representing truncation of the domain. Red and yellow combined represent truncation for "A" variants, while red alone represents truncation for "B" variants. (D) qPCR analysis quantifying EGFP knockdown by SynCas variants with different truncations of Cas1. In all qPCR plots, significance is calculated and determined using unpaired t-test between $crRNA^{EGFP}$ and $crRNA^{NT}$. Error bars represent sem (n = 4).

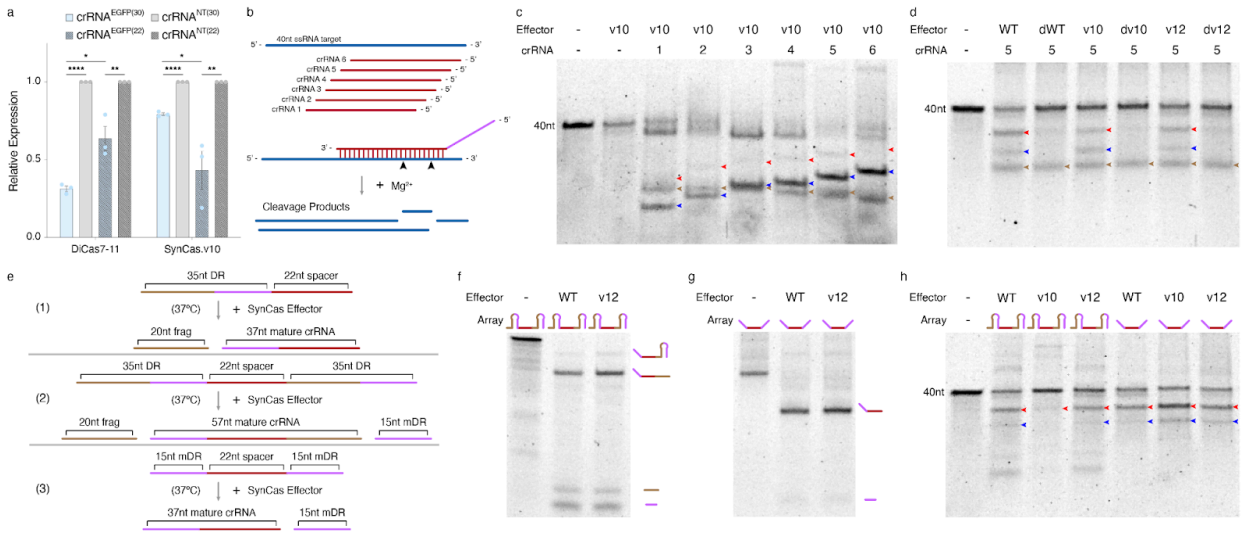


Figure 5.3: Biochemical assessment of SynCas cleavage activity. (A) qPCR analysis of EGFP knockdown comparing 30nt vs 22nt spacers for DiCas7-11 and SynCas.v10. Significance is calculated and determined using unpaired t-test between $crRNA^{EGFP}$ and $crRNA^{NT}$. Error bars represent sem (n = 3). (B) Schematic depicting the crRNAs used for assessment of *in vitro* cleavage activity and depiction of cleavage and expected cleavage products. Black arrowheads represent cut sites of crRNA-5. (C) Experiment tiling six 22nt crRNAs across the 40nt ssRNA target to obtain stepwise cleavage products. (D) Assessment of cleavage pattern and catalytic inactivation of DiCas7-11, SynCas.v10, and SynCas.v12 through site specific mutagenesis (D429A/D654A). (E) Schematic representation of different crRNA processing outcomes. (1) "single-guide" crRNA processing. (2) native array crRNA processing. (3) modified array crRNA processing. (F-G) CRISPR array processing activity by DiCas7-11 (WT) and ΔINS mutant, SynCas.v12. Array structures represented by schematics above each lane. crRNA processing products and approximate location on gel depicted by schematics on the side of gel. (H) Comparison cleavage assay between WT DiCas7-11, SynCas.v10, and SynCas.v12 (ΔINS) with varying array structures. Red arrows represent cleavage by Cas7.2 only. Blue arrows represent cleavage products by Cas7.3 or from both Cas7.2/Cas7.3 active sites cleaving the target. Brown arrow represents 20nt processed DR fragment. Purple represents 15nt mature DR (3' only).

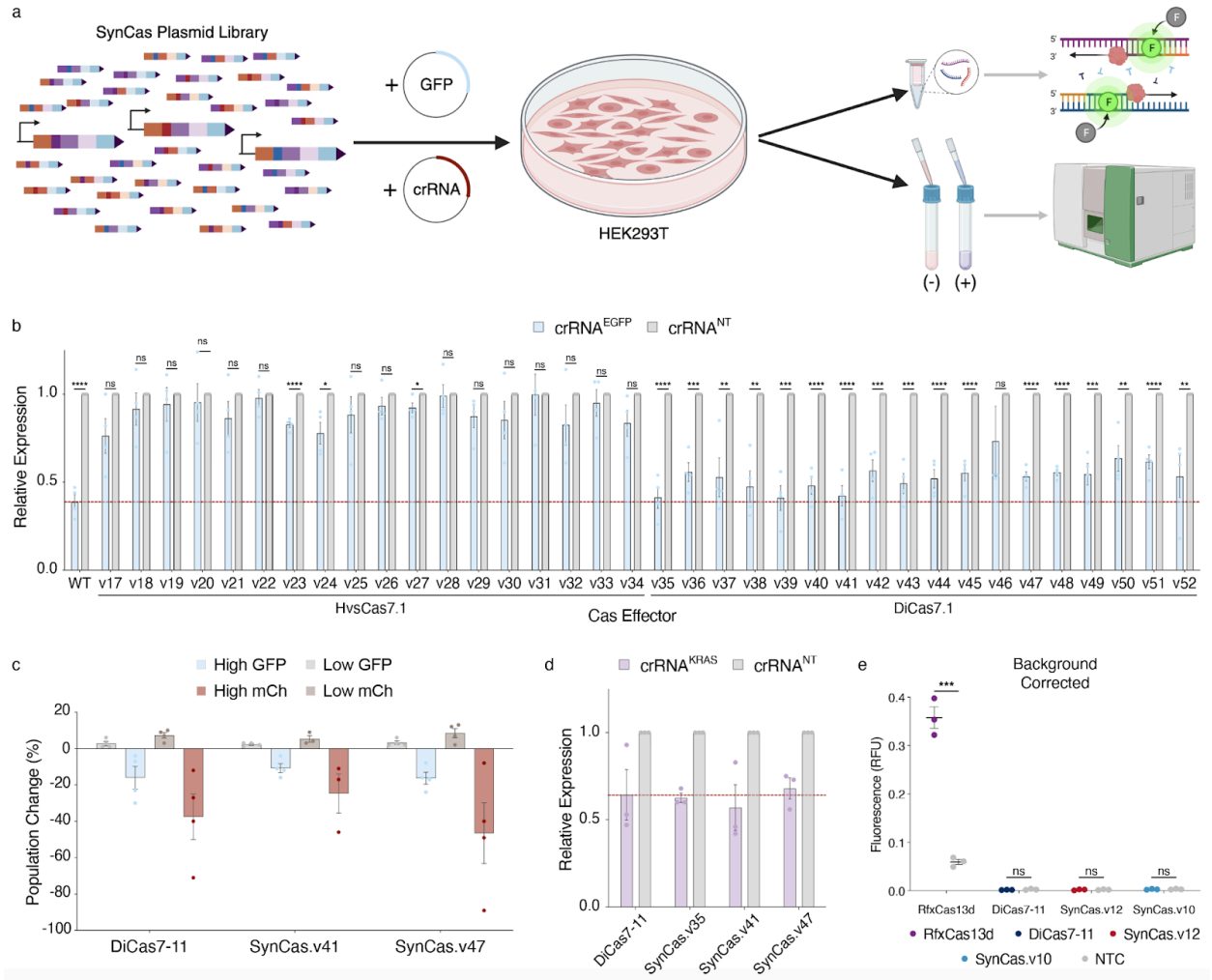


Figure 5.4: Generation and validation of compact SynCas effectors. (A) Schematic depicting workflow for analysis of SynCas library with refined ODS design method. (B) qPCR analysis of EGFP knockdown with all 36 compact SynCas effectors. Data is broken into two groups HvsCas7.1-based (v17-v34) and DiCas7.1-based (v35-v52) SynCas effectors. Dashed red line represents average knockdown by DiCas7-11 (WT). Significance is calculated and determined using unpaired t-test between crRNA^{EGFP} and crRNA^{NT}. WT = wild type DiCas7-11. Error bars represent sem (n = 4). (C) Flow cytometry analysis of select SynCas effectors demonstrating effect of RNA knockdown on translation of EGFP and mCherry. Data depicts population change in High or Low GFP, or mCherry, expression groups (see methods for gating). Error bars represent sem (n = 4). (D) RNA knockdown of KRAS gene with DiCas7-11 and select SynCas effectors analyzed using qPCR. Error bars represent sem (n = 3). (E) SENSR assay determining ssRNA collateral activity of SynCas effectors depicted by background corrected fluorescence levels. RfxCas13d was used as a positive control. Error bars represent sem (n = 3). RFU = relative fluorescence unit.

5.8 Supplementary Figures

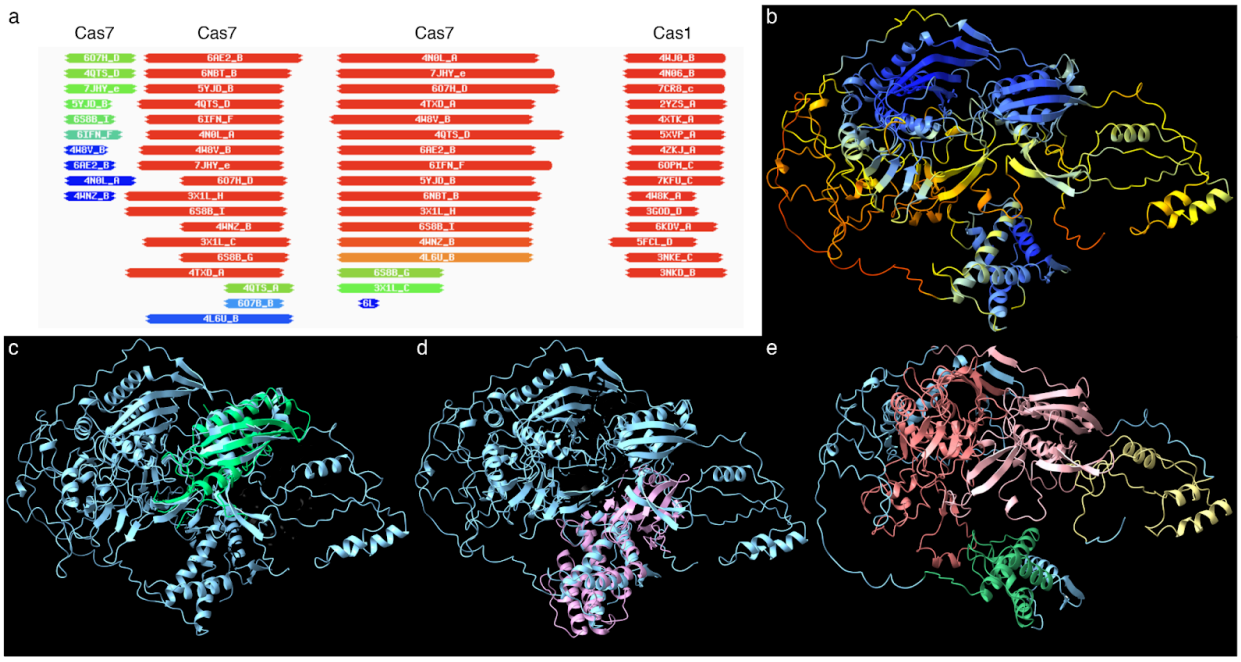


Figure S5.1: Sequence and structure prediction of Cas7-1. (A) HHpred domain assignment results. Domains are segmented based on alignments to most likely results. Results are depicted from red to blue (most likely to least likely). (B) AlphaFold predicted structure and heat map representation of Cas7-1. Coloring represents high confidence (blue) down to low confidence (red) of prediction. (C) Spatial alignment of Csm3 (green; PDB: 6nbt) over Cas7-1. (D) Spatial alignment of Cas1 (purple; PDB: 4w8k) over Cas7-1. (E) Recoloring of predicted domains from HHPred analysis. Khaki represents Csm3 domain 1 (mid confidence), light pink represents Csm3 domain 2 (high confidence), light coral represents Csm3 domain 3 (high confidence), and sea green represents Cas1 (high confidence).

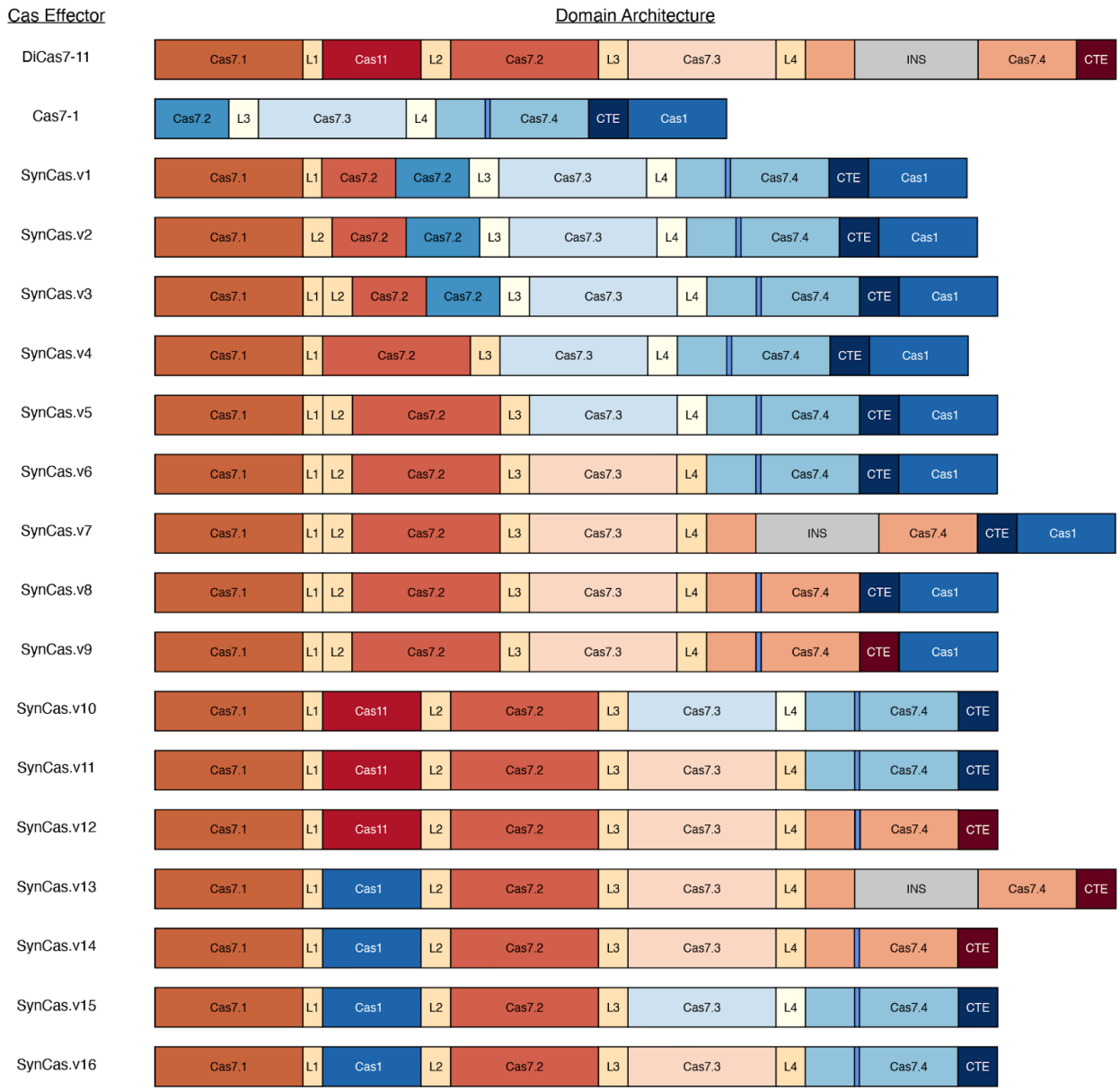


Figure S5.2: Domain breakdown of initial SynCas variants. Original proteins to build the initial SynCas variants are DiCas7-11 (red) and Cas7-1 (blue). Color scheme represents which native protein individual domains originate. Domains are not scaled.

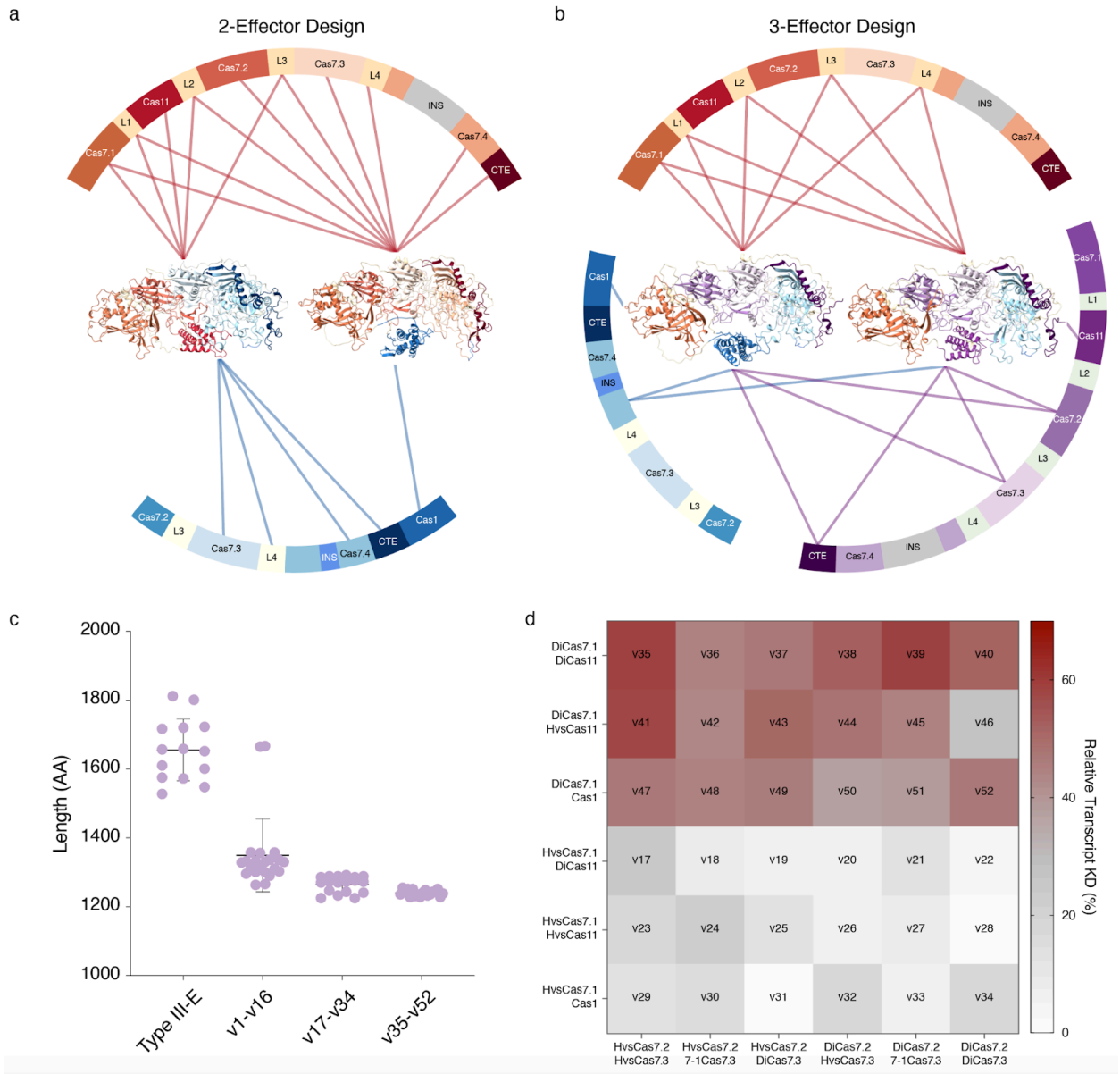


Figure S5.3: Design and assessment of SynCas effectors. (A) Schematic representation of 2-effector design strategy with SynCas.v10 (left) and SynCas.v14 (right) depicted. (B) Schematic representation of 3-effector design strategy with SynCas.v41 (right) and SynCas.v47 (left) depicted. (C) Size comparison between type III-E effectors and the various SynCas effectors generated in this study. (D) Heat map representation of qPCR data from SynCas.v17-v52 analysis. The more intense the red saturation the stronger the knockdown level. 4 replicates in each cell.



Figure S5.4: Alignment of 29AA N-terminus of Cas1 domain to RT-Cas1 fusion proteins. The Cas7-1 Cas1 N-terminal linker domain aligns well with multiple N-terminal regions of RT-Cas1 fusion proteins. NCBI ascension identities listed for each RT-Cas1 protein.

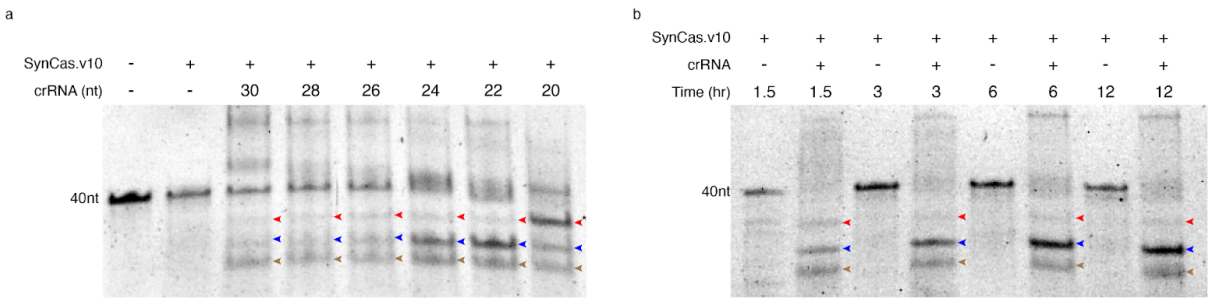
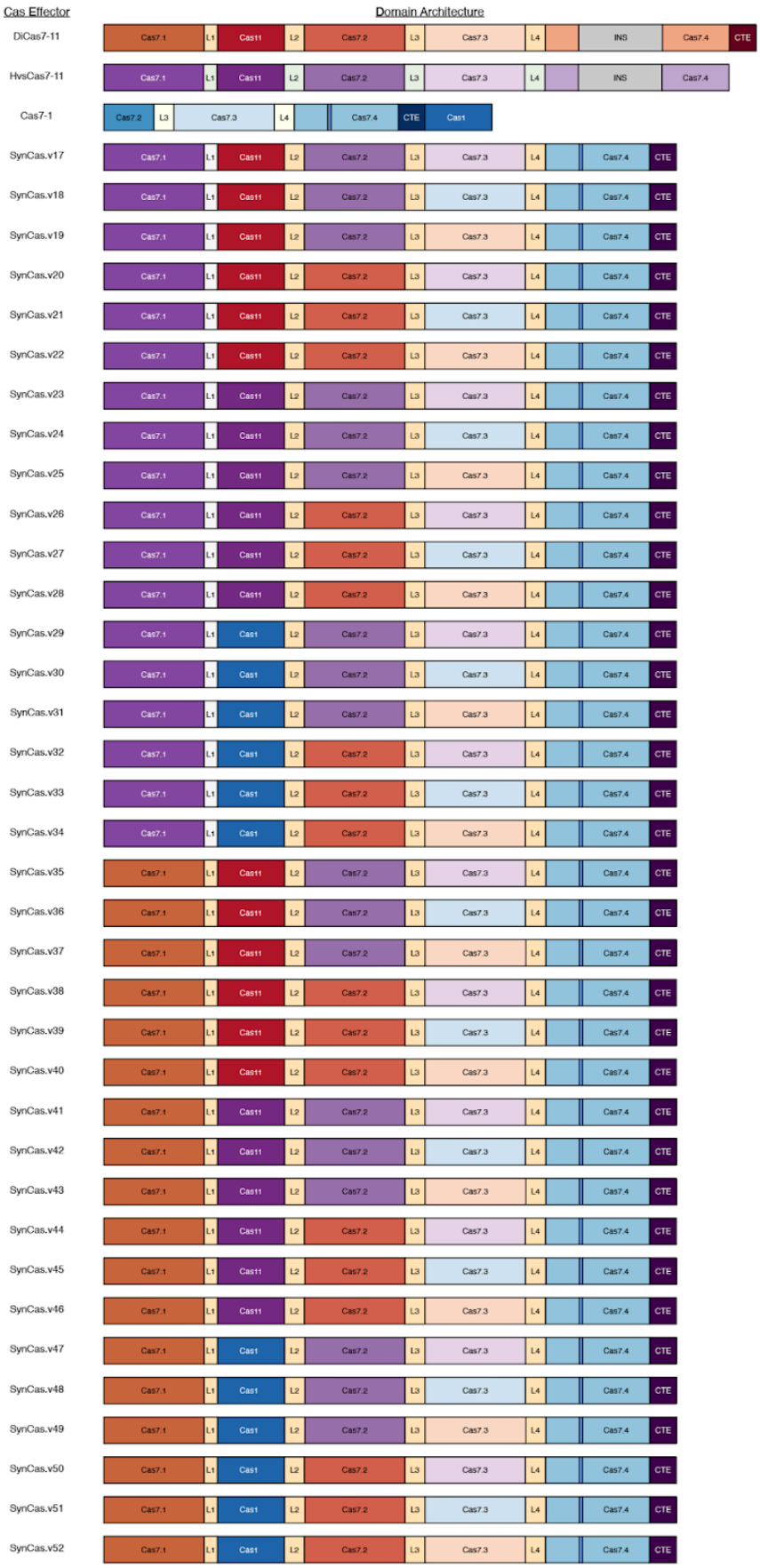


Figure S5.5: Biochemical assessment of SynCas.v10. (A) Examination of the influence of spacer length on cleavage activity of SynCas.v10. Red arrow indicates cleavage products from Cas7.2 domain, blue arrow indicates cleavage products from Cas7.3 and/or Cas7.2/Cas7.3 domain(s), and brown arrow indicates 20nt product from DR processing. The asterisk next to the 20nt spacer Cas7.2 cleavage indicates a cleavage product and processed crRNA. (B) Assessment of cleavage activity for SynCas.v10 over time (1.5 hr, 3 hr, 6 hr, and 12 hr).

Figure S5.6: Domain breakdown of compact SynCas variants. Original proteins to build the initial SynCas variants are DiCas7-11 (red), HvsCas7-11 (purple), and Cas7-1 (blue). Color scheme represents which native protein individual domains originate.


Cas Effector
Domain Architecture
DiCas7-11
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
INS
Cas7.4
CTE
HvsCas7-11
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
INS
Cas7.4
Cas7-1
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
Cas1
SynCas.v17
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v18
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v19
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v20
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v21
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v22
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v23
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v24
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v25
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v26
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v27
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v28
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v29
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v30
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v31
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v32
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v33
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v34
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v35
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v36
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v37
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v38
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v39
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v40
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v41
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v42
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v43
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v44
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v45
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v46
Cas7.1
L1
Cas11
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v47
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v48
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v49
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v50
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v51
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE
SynCas.v52
Cas7.1
L1
Cas1
L2
Cas7.2
L3
Cas7.3
L4
Cas7.4
CTE

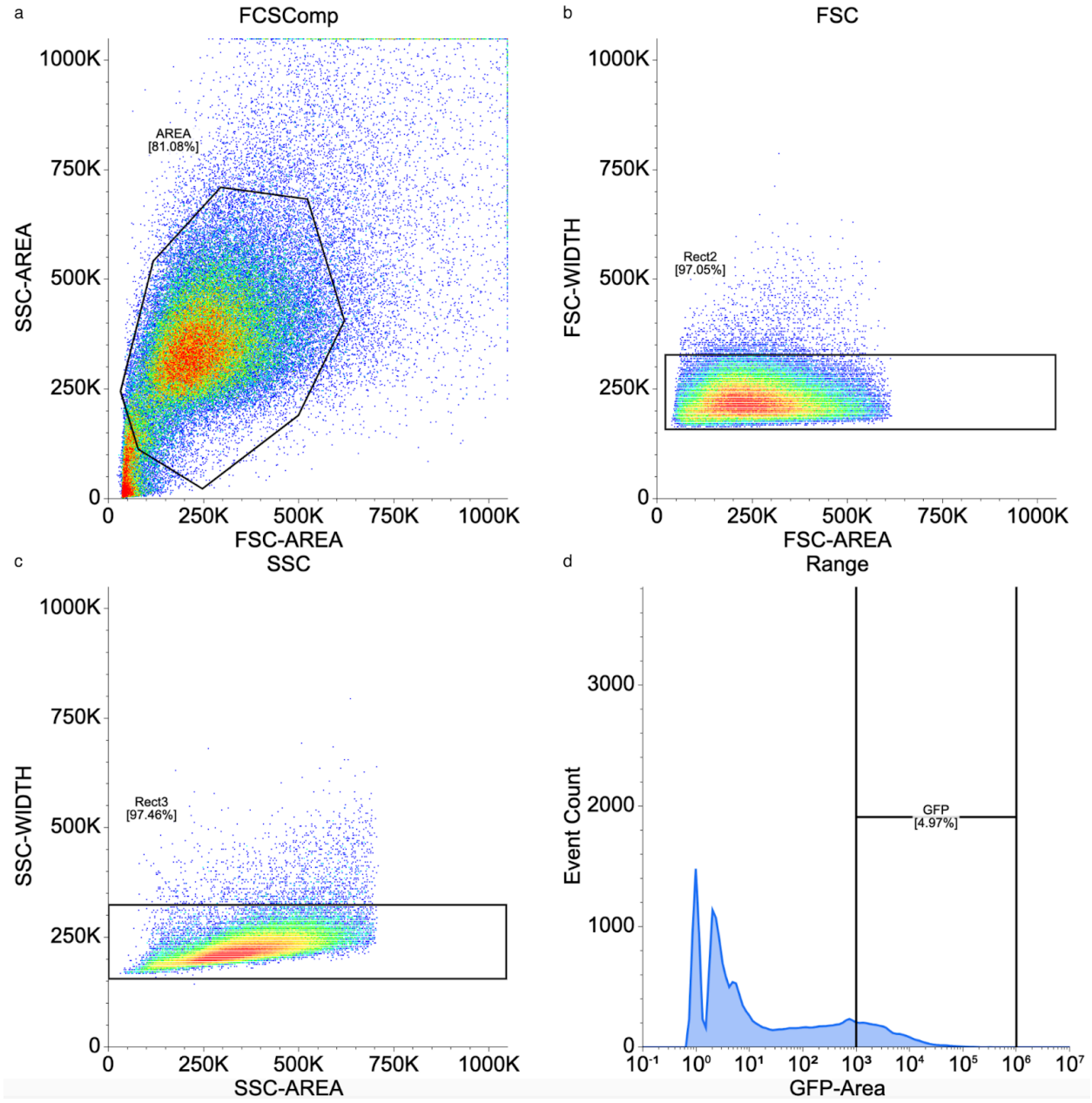


Figure S5.7: Representative gating method for flow cytometry analysis of EGFP knockdown. (A) Area of cells with polygon identifying selected cells. (B) Further refinement of selected cells – FSC width vs area. (C) Further refinement of selected cells – SSC width vs area. (D) Defined range of high EGFP population.

5.9 Supplementary Tables

Table S5.1: Breakdown of domains for type III-E effectors based on clustal omega sequence alignment.

	Cas7.1	Linker 1	Cas11	Linker 2	Cas7.2	Linker 3	Cas7.3	Linker 4	Cas7.4 (com.)	INS	CTE	Cas1	Total
DiCas7-11	238	21	106	36	181	43	199	67	300	315	95	N/A	1601
Cas7-1	N/A	N/A	N/A	N/A	72 (Partial)	48	196	72	274	6	93	143	904
HvsCas7-11	231	12	98	37	178	56	196	70	280	286	84	N/A	1528
CbfCas7-11	244	40	111	36	184	62	193	82	282	284	55	N/A	1573
CjcCas7-11	241	63	98	42	204	80	228	71	283	298	204	N/A	1812
CmaCas7-11	238	34	97	42	180	49	212	70	315	312	108	N/A	1657
CsbCas7-11	235	30	102	41	201	55	201	80	264	356	152	N/A	1717
DpbaCas7-11	N/A	N/A	42 (Partial)	40	181	50	199	74	298	292	191	N/A	1367
DsbaCas7-11	235	19	108	35	180	66	203	78	299	344	92	N/A	1659
FmCas7-11	237	37	105	39	182	46	194	72	292	225	181	N/A	1610
GwCas7-11	N/A	N/A	56 (Partial)	39	188	48	216	86	353	276	141	N/A	1403
HreCas7-11	242	51	96	40	181	50	204	78	286	279	213	N/A	1720
HsmCas7-11	236	16	94	35	181	43	192	72	283	287	109	N/A	1548
HvmCas7-11	226	57	115	40	181	50	199	74	298	292	191	N/A	1723
OmCas7-11	248	22	104	40	182	42	190	72	279	426	196	N/A	1801
SmCas7-11	239	21	102	38	183	52	194	68	283	286	109	N/A	1575
SstCas7-11	234	27	97	35	182	40	208	71	308	348	102	N/A	1652
SybCas7-11	N/A	N/A	N/A	43	181	56	207	69	286	281	206	N/A	1329
Shortest AA Sequence	226	12	94	35	178	40	190	67	264	6	55	143	

Table S5.3: List of crRNAs used in this study.

	crRNA	Spacer Target	RNA sequence (5' --> 3')	Figure(s)
1	Di-crRNA-EGFP	EGFP	gguuggaaagccgguuuucuuugaugucacggaacuucaccuugaugccguucuucugcuugucg	Figure 2, 4
2	Di-crRNA-NT	NT	gguuggaaagccgguuuucuuugaugucacggaaccuaccugguagcccuuguauuugaucaggc	Figure 2, 4
3	Di-crRNA-EGFP-22	EGFP	gguuggaaagccgguuuucuuugaugucacggaacuucaccuugaugccguucuucu	Figure 3
4	Di-crRNA-NT-22	NT	gguuggaaagccgguuuucuuugaugucacggaaccuaccugguagcccuuguauuu	Figure 3
5	Di-crRNA-CoV2-1	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaacaauuucugugguaacacuaaua	Figure 3
6	Di-crRNA-CoV2-2	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaacagaauuucugugguaacacuaa	Figure 3
7	Di-crRNA-CoV2-3	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaacguagaauuucugugguaacacu	Figure 3, 4
8	Di-crRNA-CoV2-4	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaacgguagaauuucugugguaacac	Figure 3
9	Di-crRNA-CoV2-5	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaaccugguagaauuucugugguaac	Figure 3, S4
10	Di-crRNA-CoV2-6	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaaccacugguagaauuucuguggua	Figure 3
11	Di-crRNA-CoV2-2-30	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaaccugguagaauuucugugguaacacuaauag	Figure S4
12	Di-crRNA-CoV2-2-28	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaaccugguagaauuucugugguaacacuaau	Figure S4
13	Di-crRNA-CoV2-2-26	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaaccugguagaauuucugugguaacacua	Figure S4
14	Di-crRNA-CoV2-2-24	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaaccugguagaauuucugugguaacac	Figure S4
15	Di-crRNA-CoV2-2-20	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaaccugguagaauuucuguggua	Figure S4
16	Di-crRNA-array	SARS-CoV-2	gguuggaaagccgguuuucuuugaugucacggaaccugguagaauuucugugguaacgguuggaaagccgguuuucuuugaugucacggaac	Figure 3
17	Di-crRNA-mDR-array	SARS-CoV-2	uugaugucacggaaccugguagaauuucugugguaacuugaugucacggaac	Figure 3
18	Hvs-crRNA-EGFP	EGFP	gaaacaccguuugaaggaggauaauaagguugauguuucaccuugaugccguucuucugcuugucg	Figure 4
19	Hvs-crRNA-NT	NT	gaaacaccguuugaaggaggauaauaagguugaugucuaccugguagcccuuguauuugaucaggc	Figure 4
20	Rx-crRNA-CoV2	SARS-CoV-2	caaguaaaccccuaccaacuggucgggguuugaaacguagaauuucugugguaacacuaauaguaa	Figure 4
21	Di-crRNA-mCherry	mCherry	gguuggaaagccgguuuucuuugaugucacggaaccaccuugaagcgcaugaacucc	Figure 4
22	Di-crRNA-KRAS	KRAS	gguuggaaagccgguuuucuuugaugucacggaacgguuaaaaauuuacagauugugcu	Figure 4

Table S5.4: List of select plasmids used in this study with respective Addgene IDs.

	Construct	Application	Addgene ID
1	CMV-SynCas.v10-bGH-AmpR	Mammalian expression	208589
2	CMV-SynCas.v14A-bGH-AmpR	Mammalian expression	208590
3	CMV-SynCas.v35-bGH-AmpR	Mammalian expression	208591
4	CMV-SynCas.v39-bGH-AmpR	Mammalian expression	208592
5	CMV-SynCas.v41-bGH-AmpR	Mammalian expression	208593
6	CMV-SynCas.v47-bGH-AmpR	Mammalian expression	209003
7	T7-RBS-6xHis-SUMO-SynCas.v10-term-KanR	Protein purification	208594
8	T7-RBS-6xHis-SUMO-dSynCas.v10-term-KanR	Protein purification	208595
9	T7-RBS-6xHis-SUMO-SynCas.v12-term-KanR	Protein purification	208596
10	T7-RBS-6xHis-SUMO-dSynCas.v12-term-KanR	Protein purification	208597

5.10 References

Adeola, Fowotade. 2018. “Normalization of Gene Expression by Quantitative RT-PCR in Human Cell Line: Comparison of 12 Endogenous Reference Genes.” *Ethiopian Journal of Health Sciences* 28 (6): 741–48. https://doi.org/10.4314/ejhs.v28i6.9.

Beljouw, Sam P. B. van, Anna C. Haagsma, Alicia Rodríguez-Molina, Daan F. van den Berg, Jochem N. A. Vink, and Stan J. J. Brouns. 2021. “The gRAMP CRISPR-Cas Effector Is an RNA Endonuclease Complexed with a Caspase-like Peptidase.” *Science* 373 (6561): 1349–53. https://doi.org/10.1126/science.abk2718.

Bot, Jorik F., John van der Oost, and Niels Geijsen. 2022. “The Double Life of CRISPR-Cas13.” *Current Opinion in Biotechnology* 78 (September): 102789. https://doi.org/10.1016/j.copbio.2022.102789.

Brogan, Daniel J., Duverney Chaverra-Rodriguez, Calvin P. Lin, Andrea L. Smidler, Ting Yang, Lenissa M. Alcantara, Igor Antoshechkin, et al. 2021. “Development of a Rapid and Sensitive CasRx-Based Diagnostic Assay for SARS-CoV-2.” *ACS Sensors*, October. https://doi.org/10.1021/acssensors.1c01088.

Buchman, Anna B., Dan J. Brogan, Ruichen Sun, Ting Yang, Patrick D. Hsu, and Omar S. Akbari. 2020. “Programmable RNA Targeting Using CasRx in Flies.” *The CRISPR Journal* 3 (3): 164–76. https://doi.org/10.1089/crispr.2020.0018.

Cianfrocco, M. A., M. Wong-Barnum, C. Youn, R. Wagner, and A. Leschziner. 2017. “COSMIC2: A Science Gateway for Cryo-Electron Microscopy Structure Determination.” In *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact*, 1–5. PEARC17 22. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3093338.3093390.

East-Seletsky, Alexandra, Mitchell R. O’Connell, Spencer C. Knight, David Burstein, Jamie H. D. Cate, Robert Tjian, and Jennifer A. Doudna. 2016. “Two Distinct RNase Activities of CRISPR-C2c2 Enable Guide-RNA Processing and RNA Detection.” *Nature* 538 (7624): 270–73. https://doi.org/10.1038/nature19802.

Ekundayo, Babatunde, Davide Torre, Bertrand Beckert, Sergey Nazarov, Alexander Myasnikov, Henning Stahlberg, and Dongchun Ni. 2023. “Structural Insights into the Regulation of Cas7-11 by TPR-CHAT.” *Nature Structural & Molecular Biology* 30 (2): 135–39. https://doi.org/10.1038/s41594-022-00894-5.

Goswami, Hemant N., Jay Rai, Anuska Das, and Hong Li. 2022. “Molecular Mechanism of Active Cas7-11 in Processing CRISPR RNA and Interfering Target RNA.” *eLife* 11 (October). https://doi.org/10.7554/eLife.81678.

Hu, Chunyi, Sam P. B. van Beljouw, Ki Hyun Nam, Gabriel Schuler, Fran Ding, Yanru Cui, Alicia Rodríguez-Molina, et al. 2022. “Craspase Is a CRISPR RNA-Guided, RNA-Activated

Protease." *Science* 377 (6612): 1278–85. https://doi.org/10.1126/science.add5064.

Jumper, John, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, et al. 2021. "Highly Accurate Protein Structure Prediction with AlphaFold." *Nature* 596 (7873): 583–89. https://doi.org/10.1038/s41586-021-03819-2.

Kato, Kazuki, Sae Okazaki, Cian Schmitt-Ulms, Kaiyi Jiang, Wenyuan Zhou, Junichiro Ishikawa, Yukari Isayama, et al. 2022. "RNA-Triggered Protein Cleavage and Cell Growth Arrest by the Type III-E CRISPR Nuclease-Protease." *Science* 378 (6622): 882–89. https://doi.org/10.1126/science.add7347.

Kato, Kazuki, Wenyuan Zhou, Sae Okazaki, Yukari Isayama, Tomohiro Nishizawa, Jonathan S. Gootenberg, Omar O. Abudayyeh, and Hiroshi Nishimasu. 2022. "Structure and Engineering of the Type III-E CRISPR-Cas7-11 Effector Complex." *Cell* 0 (0). https://doi.org/10.1016/j.cell.2022.05.003.

Kelley, Chase P., Maja C. Haerle, and Eric T. Wang. 2022. "Negative Autoregulation Mitigates Collateral RNase Activity of Repeat-Targeting CRISPR-Cas13d in Mammalian Cells." *Cell Reports* 40 (7): 111226. https://doi.org/10.1016/j.celrep.2022.111226.

Makarova, Kira S., Yuri I. Wolf, Jaime Iranzo, Sergey A. Shmakov, Omer S. Alkhnbashi, Stan J. J. Brouns, Emmanuelle Charpentier, et al. 2020. "Evolutionary Classification of CRISPR-Cas Systems: A Burst of Class 2 and Derived Variants." *Nature Reviews. Microbiology* 18 (2): 67–83. https://doi.org/10.1038/s41579-019-0299-x.

Özcan, Ahsen, Rohan Krajeski, Eleonora Ioannidi, Brennan Lee, Apolonia Gardner, Kira S. Makarova, Eugene V. Koonin, Omar O. Abudayyeh, and Jonathan S. Gootenberg. 2021. "Programmable RNA Targeting with the Single-Protein CRISPR Effector Cas7-11." *Nature*, September, 1–6. https://doi.org/10.1038/s41586-021-03886-5.

Silas, Sukrit, Georg Mohr, David J. Sidote, Laura M. Markham, Antonio Sanchez-Amat, Devaki Bhaya, Alan M. Lambowitz, and Andrew Z. Fire. 2016. "Direct CRISPR Spacer Acquisition from RNA by a Natural Reverse transcriptase–Cas1 Fusion Protein." *Science*. https://doi.org/10.1126/science.aad4234.

Varadi, Mihaly, Stephen Anyango, Mandar Deshpande, Sreenath Nair, Cindy Natassia, Galabina Yordanova, David Yuan, et al. 2022. "AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models." *Nucleic Acids Research* 50 (D1): D439–44. https://doi.org/10.1093/nar/gkab1061.

Wang, Qixue, Xing Liu, Junhu Zhou, Chao Yang, Guangxiu Wang, Yanli Tan, Ye Wu, Sijing Zhang, Kaikai Yi, and Chunsheng Kang. 2019. "The CRISPR‐Cas13a Gene‐Editing System Induces Collateral Cleavage of RNA in Glioma Cells." *Advancement of Science* 6 (20): 1901299. https://doi.org/10.1002/advs.201901299.

Yang, Kevin K., Zachary Wu, and Frances H. Arnold. 2019. “Machine-Learning-Guided Directed Evolution for Protein Engineering.” *Nature Methods* 16 (8): 687–94. https://doi.org/10.1038/s41592-019-0496-6.

Zetsche, Bernd, Jonathan S. Gootenberg, Omar O. Abudayyeh, Ian M. Slaymaker, Kira S. Makarova, Patrick Essletzbichler, Sara E. Volz, et al. 2015. “Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System.” *Cell* 163 (3): 759–71. https://doi.org/10.1016/j.cell.2015.09.038.